STRONG PRONOUNS IN NULL-SUBJECT LANGUAGES AND THE AVOID PRONOUN PRINCIPLE

Olga Fernández Soriano
Universidad Autónoma de Madrid

It is well known that languages like Spanish and Italian (as opposed to English and French) allow such sentences as (1a), in which there is no overt subject and that this follows from some properties of AGR under INFL, namely that it is marked [+ pronominal] (cf. Chomsky (1981), (1982), Rizzi (1982)):

(1)

- (a) estudia (Spanish)
- b. -- studia (Italian)
- c. *(He/she) studies (English)
- d. *(Il/elle) étude (French)

Nevertheless, in Pro-drop languages we have also the possibility of having an overt pronoun, that is, on a par with (1a, b), we get such structures as (2):

(2)

- a. él/ella estudia (Spanish)
- b. Lui/lei studia (Italian)

But these pronouns are not used in the same contexts in which one would use them in English or French. In fact, overt pronouns are almost never used in those contexts, That is, there is no real choice between null/overt pronoun.

We will be concerned here with this issue and try to give an explanation for this apparent alternation. We will furthermore try to clarify the cases in which a strong pronoun is permitted (or preferred) in a Pro-drop language. We will derive our explanation from the theory of pro and the nature and feature assigning properties of AGR. In doing so we make the following assumptions:

- The property of AGR that allows it to license a subject pro is that it is a (subject) clitic parallel to the ones that appear for objects and to those of Northern Italian dialects (cf. Rizzi (1986a)). Arguments for this claim are given in Fernández Soriano (in prep.)1.
- We accept Rizzi's (1986b) account of what he calls the Pro Licensing Parameter. This allows us to extend the phenomenon displayed in (1a, b) to a wider range of argument instances. Concretely, we accept that the empty category coindexed with direct and indirect object clitics is also a pro (Suñer (1982)). For our purposes, then, the set of elements that can license a pro can be reduced to (subject and object) clitics. 2

On the other hand, as the other types of empty categories in Chomsky's typology, pro has its conditions both on licensing and identification, which should be kept separated (Rizzi (1986b),
Adams (1987)). Every time both types of conditions are met, pro can appear.

1. NULL /OVERT PRONOUN ALTERNATIONS AS INSTANCES OF "DOUBLING"

   Our claim is that pro must appear whenever it is possible, that is, whenever it is licensed and fully identified. This is intended to mean that other elements in the sentence or discourse, apart from AGR or CL, may intervene in the full identification of pro.

   We are, thus, adopting a version of the so-called Avoid Pronoun Principle, which we understand as a well-formedness condition that operates in every grammar and states something like:

   \[(3) \text{pro} \text{ is obligatory when it is licensed and fully identified}\]

   In languages like English and French it happens to be the case that these conditions are never met.

   If this is true, we predict that strong pronouns never alternate with null ones. That is, if an overt pronoun appears in a context in which a pro is licensed and fully identified, this is not an alternation but a case of doubling (in the same sense as we understand clitic doubling). The pronoun appears in addition to the pro and must, therefore, have a special status and be in an extrasentential position (i.e. it is not a subject or an object).

   This will explain why one can observe that strong pronouns in null subject languages have distinctive status (see Rigau (1987), for Catalan), whose meaning is similar to the "doubling" pronouns moi, myself that appear in French and English in such constructions as:

   \[(4)a. \text{Je vais moi} \]
   \[b. \text{I go myself}\]

   And why is it, then, sentences in (2) are never equivalent to the corresponding English and French (1'a, b), as has frequently been observed, the real counterparts being (1a, b).

   \[(1') \text{He studies} \]
   \[\text{Il étudie}\]

   There are, in fact, more reasons to assume that strong pronouns in English and French are parallel to null ones in Spanish and Italian, whose strong pronouns, in turn, have a different status. That is, pro, licensed by an object clitic or by AGR, is the equivalent to English he or French lui.

   One piece of evidence comes from the fact that only null pronouns (in Pro-drop languages) and strong pronouns in non-Pro-drop languages can be bound by a quantifier (Montalbetti (1984). Strong pronouns of Pro-drop languages only allow the
correferential reading, as the following examples show:

(5)a. Muchos chicos dijeron que pro/*ellos no lo habían hecho  
   b. Many boys said that they did not do it
   c. Muchos chicos dijeron que las chicas los miraban pro/*a ellos
   d. Many boys said that the girls were looking at them

In (5a) and (5c), contrary to (5b) and (5d), the pronoun cannot be bound by the quantificational phrase muchos chicos.

Another piece of evidence is provided by Luján's (1984) data concerning the impossibility of strong pronouns in Pro-drop languages to "precede their antecedents", as opposed to null pronouns and strong pronouns in non-Pro-drop languages.

(6)a. Cuando pro/*éli trabaja, Juani no bebe
   b. When he works, John does not drink
   c. Cuando lo miran pro/*a éli, Juan se pone nervioso
   d. When they look at him, John gets nervous

As Luiggi Rizzi (p.c.) points out to me, this seems not to be a question on precedence of the pronoun but of its being part of an adverbial clause. There are two reasons for this:

a. In Italian at least, the structure is not any better if we have the adverbial clause in a non-extraposed position, that is, when the pronoun does not precede its antecedent:

   (7) Gianni non beve quando pro/*lui lavora
       Gianni does not drink when he works

   (judgements in this case are not so clear in Spanish)

b. If the embedded clause is not adverbial the pronoun can precede its antecedent without causing ungrammaticality:

   (8)a. Le persone che lui ha aiutato sono convinte che Gianni è una buona persona
        People he has helped are convinced that Gianni is a good person
   b. Las personas a las que él ha ayudado están convencidas de que Gianni es una buena persona

Nevertheless, the fact that null pronouns are parallel to strong ones in English and that strong ones behave differently still remains.

What we want to maintain is that this follows from the fact that strong pronouns in these cases are "doubled" and that their peripheral position is due to the obligatory presence of pro in subject and object position (by virtue of Avoid Pronoun).

If this is so, we expect that when there is no (strong enough) AGR the pronoun will not display the properties above. There is evidence in favor of this last claim.
1. M. Suñer (1986) analyzes some dialects of Spanish which allow overt pronouns with infinitives. In these cases, the pronoun can precede its antecedent (or be in an adverbial clause, as in (9)) and be bound by a quantifier (in fact, (10b) can be a paraphrase of (10a):

(9) Al estar ella trabajando, María no podrá venir el vienes
    Being she working, Mary will not be able to come on friday

(10)a. Muchos chicos dijeron que, para ellos hacer eso, tendrían que estar locos
    Many boys said that, for them to do that, they would have to be crazy
    b. [For many X, X a boy] [X said that, for X to do that, X would have to be crazy]

2. The same is true for the case of gerunds. This can be seen also in Italian (I owe the example to Luigi Rizzi)

(11)a. Muchos chicos dijeron que, siendo ellos felices, no les importa cómo se sientan los demás
    Many boys said that, being they happy, they do not care how the others feel
    b. [For many X, X a boy] [X said that, X being happy, X does not care how the others feel]
    c. Molti ragazi hanno detto che, essendo loro felici, ...

(12)a. Estando ella contenta, a María no le preocupa cómo estén los demás
    Being she happy, Mary does not care how the others are
    b. Essendo lei molto stanca, Maria non verrà alla festa
    Being she very tired, Mary will not come to the party

(Cf. *Siccome lei è molto stanca, Maria non verrà alla festa)

In our analysis, this will follow from the fact that there is no doubling, since there is no AGR and _pro_ is not licensed.3

3. When _pro_ is not properly identified because AGR is not strong enough, again we have that overt pronouns behave like those in English. This is what happens, for example with the obligatory _tu_ that appears in subjunctive clauses in Italian, presumably due to the fact that the verbal inflection cannot differentiate the three persons in the singular. One cannot attribute distinctive status to this pronoun:

(13) `E necessario che tu vada
    It is necessary that you go (subjuctive)

The distinctive status cannot be attributed either to the _yo/él_ that appears in constructions with the imperfect in Spanish, due to the same reason (the first and third person coincide, and, in some dialects, also the second):
(14) Iba yo/él tranquilamente andando por la calle cuando cayó un obús
   I was walking down a street when a howitzer fell

In this case one can also see that the pronoun can be bound by a quantifier:

(15) Todo testigo declaró que, él estaba sentado en un café cuando estalló una bomba
   Every witness declared that, he was sitting in a café when a bomb exploded

4. All this holds also for the cases of object pronouns in Italian, which does not allow clitic doubling (in non-colloquial speech at least). These pronouns can be bound by a quantifier and precede their antecedents:

(16) a. Molti studenti hanno detto che a loro piace giocare
    Many students said that to them it is pleasing to play
   b. Quando tocca a lei, Maria sempre gioca
    When it is her turn, Mary always plays
   c. Quando non guardano a lei, Maria si arrabia
    When they don't look at her, Mary gets mad

We can account for all these data by saying that Avoid Pronoun makes it obligatory that pro appears in subject and object position in the contexts in which it is licensed and fully identified. The strong pronoun, then, appears in addition to the null one and, thus, has a different (distinctive) status. In the contexts in which pro is not licensed (i.e. when no clitic appears) strong pronouns in Spanish and Italian behave exactly like in English or French. In the contexts in which pro is licensed but not fully identified the situation is basically the same but here there are some other facts to consider which we will try to sketch next.

The content of pro may be recovered by other element in the sentence or discourse, so the judgements about the interpretation of sentences like (14), (15) may vary depending on the speaker and the particular context. As the following examples show, the easier it is to recover the content of the subject, the less probable it will be that it is expressed by a full pronoun and the odder will be the sentences in which it is bound by a quantifier or appears in an adverbial clause:

(17) è necessario che ti vesta bene
    It is necessary that you Cl ref. dress properly

(18) Mariai siempre cuenta que, cuando (ellai) era pequeña
    Mary always says that when she was little she liked to play with dolls
Every girl said that she was sitting watching TV when someone telephoned her.

In all these sentences, there is another element apart from AGR which allows the speaker to recover the content of pro. Its appearance is, thus, much more natural and one tends to give distinctive status to the full pronoun if present. In fact, in subordinate clauses, the presence of a null pronoun is almost always preferred.

2. THE POSITION OF DOUBLED FULL PRONOUNS

In this section we will deal with the question of the position of the (subject or object) pronoun in the contexts in which it appears in addition to a pro. There are basically two positions in which it could be: either in Spec of COMP (where it has been claimed that focused or topicalized elements go) or to a position adjoined to IP. We will accept this second option for a number of reasons; here we will give just some of them for reasons of time and space.

2.1. Strong pronouns are not in Spec of COMP:

a. 'Doubled' strong pronouns are not incompatible with a Wh-element (which are in Spec of COMP).

(20) Tu hermano, al cual yo no conocía, me dijo que estabás aquí
Your brother, whom I did not know, told me that you were here

b. For the case of object pronouns one would have to postulate a special device to have a (Spec of) COMP in the right of the sentence.

2.2. Strong pronouns are not focused:

c. Strong pronouns are not incompatible with another focused element (21) and do not have focus interpretation. In fact (22b) is not a possible paraphrase of (22a):

(21) DINERO, necesitas tú 4
MONEY, need you (it is money what you need)

(22) a. Yo quiero comer
I want to eat

b. *Soy yo quien quiere comer (nadie más)
It is me who wants to eat
(The behaviour of strong pronouns is, then, parallel to that of "dislocated" elements).

I propose that the structure would be the following (irrelevant details omitted):

(23) CP
    SpecC C' C IP
    NP IP
    Yoi NP I'
    Proi I VP
    AGRi VP NP5
                   a élj
    Spec V' V+CLj NP
    Proj

3. PHI FEATURES AND AVOID PRONOUN

In this last section we will deal with the status of Avoid Pronoun (AP). We have seen so far that there is some restriction which makes the appearance of a Pro obligatory whenever it is licensed and fully identified, and we related this fact to AP, which establishes that a null pronoun is always to be preferred to an overt one (Chomsky (1981)).

Nevertheless, as it is stated, it is not clear what the status and effect of this principle would be and how we can appeal to it to obtain the desired results, that is, to ensure that the pronoun appears in addition to the Pro.

What we want to propose here is, first, that AP is a well-formedness condition that operates in every grammar. How can it be stated? It would be desirable to make it follow from some other principles of UG, and this is what we will try to do next. We will state AP as a condition against vacuous assignment of features, in line with the leading idea of Chomsky (1988). More precisely, we will derive the effect of AP from the obligatoriness of AGR to assign its (person, number and, crucially, Case) features and the impossibility for the strong pronoun in argument position to get those features.

The first question that comes to mind is whether full NP's
display any special properties in Pro-drop languages. The answer seems to be negative. One cannot attribute distinctive status to the underlined NP's in (24)

(24)

<table>
<thead>
<tr>
<th>Example</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. María quiere venir a la fiesta</td>
<td>María wants to come to the party</td>
</tr>
<tr>
<td>b. No le des eso a Juan</td>
<td>Don't CL give that to Juan</td>
</tr>
<tr>
<td>c. Mis padres no tienen dinero</td>
<td>My parents do not have any money</td>
</tr>
<tr>
<td>d. Los estudiantes pidieron que</td>
<td>The students asked for the exams</td>
</tr>
<tr>
<td>se retrasaran los exámenes</td>
<td>to be postposed</td>
</tr>
</tbody>
</table>

These NP's, thus, have the same status as the pronouns that appear in such constructions as (9) to (16), in which their content is not fully recoverable. It would be desirable to derive the effect of Avoid Pronoun from some other principle of the grammar in such a way that allows us make contrast between pronouns and full NP's (and among pronouns themselves) follow from more basic differences which distinguish both elements. What are, then, the similarities between a pronoun whose appearance is obligatory because of the lack of features in AGR to assign and a Full NP (i.e., between English and French pronouns, full NP's and Spanish and Italian obligatory pronouns). Let's take one aspect of the problem at a time.

Intuitively speaking, it seems that, if AGR is a pronominal element, the presence of another pronominal that introduces the same features is, in some sense, redundant. Pro, on the other hand, can only have the features assigned by AGR, since one can safely propose that it has no features of its own.

It seems to be plausible as well to assume that, if AGR has features to assign, it must assign them. We accept that these features are basically person, number and, crucially, Case.6

(25) Strong AGR must obligatorily assign its person number and Case features.

In Baker (1988) it is claimed that NP's can be visible, i.e., casemarked, in two ways (apart from the standard "strict adjacency"): by having a morphological marking, that is, by bearing a case morpheme, as in languages like Latin, or by having an agreement morpheme in the Case assigning head. This seems to be what happens in Pro-drop languages in the case of the subject and also in the case of objects if we accept that clitics and AGR (in Pro-drop languages) are the same element. Let us further suppose that an NP cannot be Case-marked in these two "morphological" ways.

(26) An NP cannot get Morphological Case in two different ways

Strong pronouns in Spanish and Italian are the only elements in which Latin Case marks still remain, so one can claim that they
are "intrinsically" Case-marked, i.e. they bear Case morphemes, as well as person and number. If they are assigned these same features by INFL this would result in a "vacuous assignment", and, thus a violation of (26).

In English and French one can claim that AGR has no features to assign and, crucially, that subjects do not get case by means of having an agreement element (a pronominal clitic) in the case assigner but by being governed by it. One might claim that this way of getting case is not incompatible with "morphological" Case, and we can attribute their (obligatory) appearance in subject (and object) position to the fact that they introduce features that AGR does not have: person and number. This would be the case of Spanish and Italian obligatory pronouns: AGR does not have ("enough") features and a Pro would not be identified (and not even licensed if one claims that AGR in English is not a pronominal –a clitic– and therefore not a licenser).

Pro, on the other hand, can be considered as "empty" in the sense that it does not have any features of its own (apart from thematic role, which it gets from the verb probably in Spec of VP) and can (and must) get them from AGR (which, as we said, must assign them obligatorily).

Something similar happens with full NP's, which do not have any case marking and must, thus, get it from AGR.

Evidence for this last claim might come from the fact that in standard Spanish only pronominal subjects can appear with infinitives and in Caribean Spanish this is highly preferred (M. Suñer (1986)), full NP's are impossible (or marked, in Caribean) in this construction. This can be accounted for in our analysis, since full NP's would not have any way of being PF-identified (Baker (1987)) as they do not have morphological case.7 (It is a mystery, though why strong pronouns cannot appear as preverbal subjects of infinitives under raising verbs).

In Fernández Soriano (1988) it is analyzed some data about second language acquisition which show that the fact that foreign students learning Spanish (or Italian) tend to overuse the pronouns can be related to their lack of familiarity with verbal inflection and not to the failure to incorporate the Pro-drop option. In this case, then AGR is not "distinctive' enough to fully identify a pro. Some evidence for this comes from the fact that the students almost never use expletive strong pronouns and that the use of pronouns is minimal in subordinate clauses (when the content is more likely to be recoverable) and with third person (in this case they prefer sometimes to use the full NP). Another piece of evidence for this is provided by the behaviour of Spanish students learning Italian. Here, despite of the fact that both languages are Pro-drop, they overuse strong pronouns exactly in the same contexts as before. This shows that their AGR (with which they are not familiar yet) is not rich enough and they prefer to use a full pronoun.

Sumarizing, in this paper we have dealt with the special status of strong pronouns in Null-subject languages and seen that this status is related to the fact that the pronoun is "doubled". On the other hand, it seems that this follows from the fact that,
in all the context in which an overt pronoun displays the properties analyzed, a pro is licensed and fully identified. Assuming some version of AP, we claimed that the presence of that pro is obligatory in those cases, the pronoun being, thus, in a non-argument position (concretely, adjoined to IP or VP). This accounts for its special status. Furthermore, we made AP follow from other, more general, principle of the grammar against vacuous assignment of features, specifically, against vacuous assignment of "morphological" case.

NOTES

* This paper was started during the stay of the author in Geneva with a grant from the Comunidad Autónoma de Madrid. Its further development was supported by a Research grant from the Comisión Interministerial para la Ciencia y la Tecnología (CICYT) to the Research Project Configuración y Papeles Temáticos en la Sintaxis y la Morfología (Ref. PB 85-024)

I am very grateful to L. Rizzi, I. Roberts, A. Belletti, G. Cinque, T. Guasti and G. Giusti for their very helpful comments in the previous version of this paper during my stay in Geneva.

Thanks also to Esther Torrego for her interesting comments and suggestions. Special thanks go to Violeta Demonte for her encouraging, generous and helpful comments and ideas both on the form and content of this paper.

1. What we really want to claim is that CL and AGR are the same element. In fact, there is also evidence (cf. Fernández Soriano (1984)) that shows that clitics can be considered agreement morphemes (they must attach to a root, they do not undergo deletion under identity, they sometimes trigger internal sandhi phenomena, they never appear in their own -not even as an answer to a question, etc.). We take the label clitic for both elements because it seems to capture better the idea of a pronominal element and the parallelism of AGR with subject clitics of the Italian dialects. But the real generalization is that AGR in Pro-drop languages is, for all syntactic purposes, identical to object clitics, the only differences being (morpho)phonological.

2. In fact, Rizzi (1986b) claims that in certain languages there is a slot in the verbal matrix which also licenses a pro in object position. Among other things, this comes from the fact that it can give an explanation for languages like Hindi, for example, in which there are instances of pro in object position and it does not have clitics. For our purposes, though, we can accept that clitics are the pro licensors.

3. V. Demonte points out to me that the so called "absolute participle" constructions have a special behavior, since they do not seem to admit pronominal subjects and, on the other hand, they can take full NP's and sometimes null pronomininals. This can be shown with the following contrast:
Nevertheless, this might be explained if we accept that in this cases there is an AGR node which triggers participle agreement, in line with Kayne (1987) and Chomsky (1988). We will not pursue this matter here.

4. The reason why in this sentence the subject is postverbal is the appearance of an element in COMP, which triggers verb anteposition. See Torrego (1984) for details.

5. The full pronoun which reffers to the object might also be generated adjoined to IP. We take the other option (VP-adjunction) because it can appear before certain adverbs, like in

(i) Lo vi a él en el teatro
    Lo vi a él por la mañana

I owe this suggestion to V. Demonte.

6. Some parts this idea came out of very fruitfull conversations with I. Roberts.
REFERENCES