``Arbitrary'' *pro* and the theory of pro-drop[*]

Patricia Cabredo Hofherr,
CNRS UMR 7023 / Université Paris 8
cabredo@linguist.jussieu.fr

Since Rizzi's influential work on pro-drop (Rizzi 1986), it has been generally assumed that null subject pronouns pattern together depending on their theta-role. The idea is that null pronouns bearing a full theta-role need more content identification than null pronouns bearing a quasi-theta-role or no theta-role at all (section 1).

In what follows I will discuss evidence showing that the possibility of pro-drop is not directly dependent on the theta-role borne by the pronoun (section 2). In particular I present data from different languages [1] that allow null 3rd person pronouns with so-called arbitrary reference even though anaphoric 3rd person argumental *pro* is not available.

In order to account for the split between arbitrary and anaphoric null pronouns I propose a modification of Rizzi's theory of pro-drop. I argue that anaphoric and non-anaphoric *pro* should be distinguished and I propose that the two types of pro-drop are content-identified by two different types of agreement (section 3).

Finally, in section 4, I examine some predictions of the present analysis.

1. Rizzi's (1986) theory of pro-drop

Since Rizzi (1986) it has been generally accepted that null pronouns are subject to two separate licensing requirements, namely (i) formal licensing and (ii) content identification.
formal licensing: *pro* is governed by $X_y$

(where $X$ is a governing head of type $y$)

content identification: Let $X$ be the licensing head of an occurrence of *pro*:

then *pro* has the grammatical specification of the features on $X$ co-indexed

with it.  

(Rizzi 1986: 519-20)

Rizzi further points out that languages do not necessarily allow pro-drop for all types of subject pronouns (partial pro-drop). In particular, there are languages that have null subject pronouns with weather predicates, but not with referential subjects. Rizzi proposes that the possibility of having a null pronoun is dependent on the theta-role and distinguishes three types of pro-drop:

(2) Three types of pro-drop (Rizzi 1986):

a. referential *pro*: null pronouns bearing a full theta-role,

b. quasi-argumental *pro*: null pronouns bearing a quasi-theta-role,

c. expletive *pro*: null pronouns bearing no theta-role.

According to Rizzi's theory of pro-drop, the difference between these three types of null pronouns lies in their feature content. For referential *pro* to be licensed the features [person] and [number] have to be identified. For the content identification[2] of quasi-argumental *pro* it is sufficient to identify the [number]-feature, while for expletive *pro* no features have to be identified.

(3) Content identification (Rizzi 1986):
a. referential: identify the feature [person] and [number],
b. quasi-arguments: identify the feature [number],
c. expletives: no features need to be identified.

Rizzi notes that any language capable of identifying the feature content of referential null subjects should also allow null pronouns with fewer features and therefore admit quasi-argumental and expletive pro-drop. Rizzi’s theory of content identification therefore predicts an implicational hierarchy among the three types of pro-drop.

(4) Rizzi’s pro-drop hierarchy:

referential pro -- > quasi-argumental pro -- > expletive pro

full th-role quasi-th-role no th-role

Rizzi’s proposal has been the starting point for many studies of pro-drop systems in typologically diverse languages. These studies have led to some modifications of Rizzi’s original proposal. In particular, it has been pointed out that in Finnish and Modern Hebrew the referential pronouns are split into two classes with respect to pro-drop: 1st and 2nd person pronouns can be null while the 3 person pronouns bearing a full theta-role have to be lexicalised (see e.g. Holmberg & Nikanne 1994, 1999 for Finnish and Borer 1980 for Modern Hebrew).

In order to account for these pro-drop patterns, refinements of Rizzi’s licensing conditions for referential pro have been proposed (see e.g. Holmberg & Nikanne 1994 for Finnish, Shlonsky 1997 for Modern Hebrew and Vainikka & Levy 1999 for a comparison of both languages).
Less attention has been paid to a split among the referential 3rd person pronouns with respect to pro-drop. In what follows, I will examine the behaviour of antecedentless 3rd person null pronouns (so-called `arbitrary pro'[3]) and I will present data from four languages showing that these pronouns are clearly different from anaphoric 3rd person pronouns with respect to pro-drop.

2. `Arbitrary'' pro drop in full theta positions

The data presented in this section show that arbitrary 3rd person pronouns may differ from anaphoric 3rd person pronouns with respect to pro-drop. I will discuss four languages with partial pro-drop systems, namely Modern Hebrew, Finnish, Russian and Icelandic.

All four languages have quasi-argumental pro-drop: as the following examples show, the subject of weather predicates is null.

(5)  a. kar. (MH)
cold
'lt is cold.' (ex (7-30a), Shlonsky 1997)

b. Sataa.[4] (Fi)
rains

c. segodnja xolodno/teplo. (Ru)
today cold/warm
'Today it is cold/warm.'
Icelandic and Russian do not have referential null subjects (see Holmberg & Platzack 1995 for Icelandic and Franks 1995 and references cited there for Russian[5]). Finnish, in contrast, has partial referential pro-drop: it allows 1st and 2nd person null subject pronouns while anaphoric 3rd person pronouns have to be lexical (see Holmberg & Nikanne 1994).

The most complex pattern is observed in Modern Hebrew where the possibility of pro-drop varies with the tense of the verb. The present tense, historically derived from a participial form marked for number and gender only, does not license 1st, 2nd or 3rd person pro-drop (see Borer 1980, Shlonsky 1997). The past and future forms, in contrast, are inflected for person, number and gender and allow null 1st and 2nd person subjects while 3rd person...
subjects have to be lexical (unless they are controlled by an antecedent in a superordinate clause, see Borer 1989).

In all four languages arbitrary (i.e. antecedentless) 3rd person pronouns differ from anaphoric 3rd person pronouns with respect to pro-drop.

As pointed out by Shlonsky (1997), in Modern Hebrew arbitrary 3pl pronominals are null in all tenses. Unlike the 1st and 2nd person and 3rd person anaphoric pronouns arbitrary 3pl pronouns are null in the present tense (see (7a)).

(7) a. be-Sav'u't oxlim givna. (MH)
in-Pentecost eat.benoni.mpl cheese
'One eats cheese in Pentecost.' (ex in Borer 1998)

b. moxrim Sam kartisim.
sell.benoni.mpl there tickets
'They(arb) sell tickets there.' (ex (7-30b) in Shlonsky 1997)

In the past and future, the 3pl arbitrary pronouns are null contrast with the 3pl anaphoric pronouns which have to be lexical:

(8) maxru Sam kartisim. (MH)
sell.past.3mpl there tickets
'They(arb) sold tickets there.' (ex (7-29b) in Shlonsky 1997)

The same split between arbitrary and anaphoric 3rd person pronouns can be found in Finnish.
The arbitrary pronoun in Finnish takes 3sg morphology but does not pattern with the referential 3sg pronouns. Unlike the 3sg person referential anaphoric pronouns that have to be lexical, arbitrary pronouns are null (see Hakulinen and Karttunen (1973) for more examples):

(9) Metsästä löytää helposti mustikoita. (Fi)
    forest-ela find.3sg easily    blueberries-par

    'One finds blueberries easily in the forest.' (ex (43) in Vainikka & Levy 1999)

In Russian, arbitrary 3pl subjects also have to be null differing from the other referential pronouns. In examples like the following the lexical 3pl subject pronoun oni is impossible with an arbitrary interpretation:

(10) vo Francii (*oni) edjat ulitok. (Ru)
    in France (they)  eat.3pl snails

    'In France, they(arb) eat snails.'

As discussed in Sigurdsson (1989, 1990), Icelandic also allows arbitrary null subjects under certain circumstances:

(11) Eg vissi ekki aD aetti aD fara svona oft til Graenland. (Ice)
    I knew not  that should to go so often to Greenland

    'I did not know that one should go so often to Greenland.'

    (ex (28a) in Sigurdsson 1990)
The distribution of arbitrary 3rd person *pro* as compared to referential and quasi-argumental *pro* in the languages discussed here can be summarised by the following table:

(12)

<table>
<thead>
<tr>
<th></th>
<th>1/2 pro</th>
<th>3 anaphoric pro</th>
<th>3 arb pro</th>
<th>weather-pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finnish</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>MH past/future</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>MH present</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Russian</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Icelandic</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The arbitrary pronouns in the examples discussed above occupy the subject position of predicates such as *sell tickets, eat snails, find blueberries* which are clearly assigned a theta-role. This implies that the contrast between arbitrary and anaphoric referential pronouns with respect to pro-drop cannot be captured in terms of a difference in the type of theta-role along the lines of (3) above. In the following sections I will therefore propose an alternative analysis in order to account for the split between anaphoric and non-anaphoric 3rd person pronouns with respect to pro-drop.

3 An alternative analysis

The data discussed in the previous section show that 3rd person null pronouns bearing a full theta-role do not form a homogeneous class. In order to account for this observation I will develop an alternative to Rizzi's typology of pro-drop repeated here.
Three types of pro-drop (Rizzi 1986):

a. referential *pro*: null pronouns bearing a full theta-role
b. quasi-argumental *pro*: null pronouns bearing a quasi-theta-role
c. expletive *pro*: null pronouns bearing no theta-role

I propose to replace Rizzi's distinction based on the theta-role by a distinction between (i) deictic, (ii) anaphoric and (iii) non-anaphoric null pronouns. The pronouns of 1st and 2nd person singular and plural are deictic containing reference to the speaker and the hearer; the distinction between anaphoric and non-anaphoric pronouns therefore only applies to 3rd person pronouns.

Three types of pro-drop (modified):

a. deictic *pro*: null pronouns marked [+speaker]/ [+hearer],
b. anaphoric *pro*: null 3rd person pronouns that are take up a discourse referent previously introduced in the discourse,
c. non-anaphoric *pro*: null 3rd person pronouns that do not take up a discourse referent previously introduced in the discourse.

The grammatical distinction between anaphoric and non-anaphoric pronouns is independently motivated. Languages such as Bavarian and Frisian have grammaticalised this distinction on the definite determiner. The non-anaphoric form (glossed det.NA) is used with proper names, unique entities and kind-referring NPs while the anaphoric form (glossed det.A) is used with entities introduced in the discourse (see Ebert 1970 for Frisian, Krifka et al. 1995 for Bavarian):
(15)  
a.  Da/*dea  Kare is  kema.  (Bav)  
det.NA det.A Karl has arrived.  (proper name)  
b.  Da/ *dea  Kini is  gschtoabm.  
det.NA det.A king has died.  (unique entity)  
c.  Da/ *dea  Schnaps is deia.  
det.NA det.A Schnaps is expensive.  (kind-referring NP)  
d.  I hab  a Bia  un  an Schnaps bschdait.  Dea/ *da  Schnaps war deia.  
I have a beer and a schnaps orndred.  det.NA det.A schnaps was expensive.  
(anaphoric NP)  

The class of non-anaphoric 3rd person pronouns includes quasi-arguments and arbitrary 3pl: in both cases the pronoun does not pick out an already established discourse referent. This grouping of quasi-arguments and arbitrary 3pl is consistent with the evidence from the languages examined in the previous section. The data show that with respect to pro-drop the arbitrary 3rd person pronouns are more similar to quasi-argumental pronouns than to anaphoric argumental 3rd person pronouns. This observation has been explicitly made by Shlonsky (1997) for Modern Hebrew, and this author considers arbitrary pro and quasi-argumental pro as members of a single class that he terms quasi-referential pronouns. The present analysis gives an explicit formulation of the intuition underlying Shlonsky's proposal: the similarity between the two types of pronoun is reduced the fact that both lack an antecedent.

Following Rizzi's theory of pro-drop, I assume that any instance of a null pronoun needs formal licensing. As in Rizzi's analysis, I therefore propose to derive the difference between the different types of pro-drop from differences at the level of content identification. More
precisely, I propose that anaphoric and non-anaphoric 3rd person pronouns differ with respect to the features that have to be identified by agreement:

(16) Content identification (modified)

(i) deictic pro arises with agreement that identifies phi-features including a feature [+speaker] or [+hearer] (1st/2nd person pronouns),

(ii) anaphoric 3rd person null pronouns arise with an agreement morpheme identifying the full set of phi-features person, number and gender,

(iii) non-anaphoric 3rd person null pronouns arise with agreement that only identifies a subset of phi-features.

The following sections develop the mechanisms of content identification that apply to non-anaphoric 3rd person pro.

Section 3.1 develops the clause in (16iii) identifying the subsets of features that give rise to non-anaphoric pro-drop. I will argue that two different subsets of phi-features identified by agreement account for the difference between (i) non-anaphoric non-human 3rd person pro (quasi-arguments) and (ii) non-anaphoric [+human] 3rd person pro (arbitrary 3pl pro).

In section 3.2 I will examine the case of non-anaphoric [+human] 3rd person plural pro more in detail. I will argue that (at least) 5 types of non-anaphoric [+human] 3rd person plural pro have to be distinguished (section 3.2.1).

In order to account for the different types of non-anaphoric 3pl pro I propose that for the content identification of arbitrary null pronouns the subset of phi-features identified by the agreement combines with further licensing mechanisms that yield the different arbitrary readings identified in section 3.2.1. These further licensing mechanisms are discussed in
3.1 Two types of deficient phi-features

In the previous section I have proposed that non-anaphoric 3rd person pronouns form a class with respect to pro-drop by virtue of having a partial set of phi-features. The class of non-anaphoric lexical pronouns includes (i) expletives (appearing e.g. with inverted subjects), (ii) quasi-arguments (e.g. subjects of weather-predicates) and (iii) arbitrary 3rd person subjects. The existence of null quasi-arguments and null arbitrary 3rd person pronouns is motivated by the fact that these null pronouns are syntactically active, for example with respect to control and binding. The existence of null expletives (expletive pro) is more problematic, since its principal motivation is theoretical providing an element that satisfies the requirement that every finite clause have a syntactic subject in a specific structural position (EPP). Rizzi gives German as an example of a language that only has expletive pro. I have argued for an analysis of German without a empty subject pronouns elsewhere (cf Cabredo Hofherr 1999), and I will therefore assume that expletive pro does not exist.

The distinction between quasi-arguments and arbitrary subjects corresponds to a morphological difference in languages that have non-anaphoric readings with 3rd person singular and plural pro[6]: weather-predicates appear with the 3rd person singular, while arbitrary 3pl subjects appear with the 3rd person plural.

(17)  
   a. 3sg non-anaphoric pro: quasi-arguments (e.g. subjects of weather-verbs) and

   b. 3pl non-anaphoric pro: arbitrary 3pl subjects.
In what follows I will argue that while 3sg and 3pl agreement both may identify subsets of phi-features, the subsets identified differ with respect to the number feature.

The 3sg agreement-form is uncontroversially the proto-typical agreement form that appears as a default with weather predicates (18a), extraposition (18b) and also with nominal agreement morphology, as e.g. non-agreeing participles (18c):

\[(18)\]
\[
a. \quad \text{Llueve.} \quad \text{(Sp)}
\]
\[
\quad \text{(it) rains.3sg}
\]
\[
b. \quad \text{Parece que Juan no quiere venir.}
\]
\[
\quad \text{(it) seems.3sg that Juan not wants come.inf}
\]
\[
c. \quad \text{Tus hermanas han venido.}
\]
\[
\quad \text{your sisters.3fpl have.3pl come.msg}
\]

The 3rd person singular agreement appearing in these examples does not contribute any obvious semantic content.

The 3rd person plural verbal agreement that allows the non-anaphoric readings of 3pl \textit{pro} is of a different type since it contributes to the interpretation. First, the arbitrary subject is unspecific but [+human] and secondly, in certain languages, the arbitrary subject is 3rd person excluding speaker and hearer (for discussion see section 3.2.1 below).

A pattern resembling that of arbitrary 3pl subjects can be found with the nominal agreement in the examples in (19): (i) without an antecedent the interpretation is human (see Corblin 1995 for French) and (ii) the morphological features are 3pl.
I propose that both 3rd person agreement forms appearing with non-anaphoric pro have in common that their specification for [number] is deficient.

The difference between the 3rd person singular and the 3rd person plural lies in the fact that 3sg deficient agreement is uncountable, corresponding to a mass noun, while 3pl deficient agreement is countable, corresponding to count noun. This proposal is supported by the following generalisation.

Cross-linguistically, in the interpretation of non-anaphoric forms that are not marked for gender, the distinction between singular and plural corresponds systematically to a distinction between mass/ inanimate and count/ human interpretation respectively:

(20) a. Beaucoup a été dit. (Fr)
    ‘Much has.3sg been said.’ (sg, inanimate)

    b. Beaucoup ont été tués / vus.
    ‘Many have.3pl been killed / seen.’ (pl, human)

(21) a. Poco se sabe de él. (Sp)
    little refl know.3sg of him
    ‘Little is known of him.’ (sg, inanimate)

    b. Pocos vinieron a verlo.
    little.mpl came3pl to see.him
‘Few came.3pl to see him.’

(22) a. Much has been done. (sg, inanimate)
b. Many have been killed. (pl, human)

I propose that the difference between count and mass status of the pronoun associated with the deficient agreement is due to their specification for [number] feature.

While the 3pl deficient agreement has an underspecified [number] feature, the 3sg deficient agreement has no [number] feature at all. For the 3pl deficient agreement the slot for the number feature is present but the value is not specified by the agreement morpheme and has to be retrieved from the context. For the 3sg deficient agreement the number feature is simply absent and therefore never assigned a value.

With respect to the [person] feature, the deficient 3pl agreement may vary cross-linguistically: The non-anaphoric 3pl readings in Spanish exclude speaker and hearer, while this is not the case for Russian or Modern Hebrew (see section 3.2). I propose to analyse this difference as a difference in feature content of the 3pl agreement with an underspecified [number] feature: while the deficient 3pl agreement in Spanish is marked for [3rd person], the 3pl agreement in Russian and Modern Hebrew is not marked for [person]. I will remain neutral as to the question whether the [person] feature is underspecified (i.e. filled in by the context) or completely absent.

Notice that the difference between the non-anaphoric 3pl readings coincides with a difference in the verbal morphology. Russian and Modern Hebrew both have tenses that do not mark [person] but only [gender] and [number], namely the Russian past and the Modern Hebrew present. In Spanish, in contrast, all tenses mark person distinctions.
The deficient 3sg agreement, in contrast, yields a mass interpretation: this interpretation automatically excludes speaker and hearer since mass interpretation is inanimate. I therefore propose that the 3sg deficient agreement that appears with weather-predicates is not marked for person.

Given the preceding discussion, the 3pl deficient agreement is more complex than the 3sg agreement in that it is underspecified for [number], while in the 3sg deficient agreement the number feature is absent.

Furthermore, the 3pl deficient agreement may have a value for person, namely [3rd person], while the 3sg deficient agreement does not have a [person] feature. If the 3sg deficient agreement has neither [number] nor [person] features, the only feature contributed by the agreement is the characteristic feature of nominals, presumably [+N].

Summarising, I propose that non-anaphoric pro may arise with two types of deficient agreement. These two types of agreement crucially differ with respect to their specification for the feature [number]. This difference is reflected in a difference with respect to the mass vs count status of the non-anaphoric pronouns identified by the deficient agreement.

The feature [+human] is not a specific property of non-anaphoric 3pl pronouns but falls under a wider generalisation that non-anaphoric 3rd person plural forms are interpreted by default as referring to humans, while the non-anaphoric 3rd person singular forms are interpreted as inanimate. According to the analysis proposed here this is related to the fact that mass-interpretation is related to inanimacy while countable interpretation is related to the feature [+human][7]. The following summarises the main points of the analysis:
3sg deficient agreement contains the feature: +N
no number feature => mass => inanimate interpretation by default

b. 3pl deficient agreement contains the features:
   (i) +N, underspecified number (e.g. Russian) or
   (ii) +N, 3rd person, underspecified number (e.g. Spanish)
   number feature present => count => human interpretation by default

3.2. The analysis of non-anaphoric 3pl pro

In the preceding section I have proposed that the feature [number] allows us to distinguish two types of deficient agreement yielding 2 types of non-anaphoric pro, namely 3sg quasi-argumental pro and 3pl non-anaphoric (``arbitrary'') pro.

The non-anaphoric readings of 3pl pronouns do not form a uniform class, however. As I will argue in section 3.2.1, five types of non-anaphoric 3pl pronouns have to be distinguished. In order to account for the different readings of non-anaphoric 3pl, I propose that the content identification due to the agreement morpheme combines with further interpretive mechanisms, discussed in section 3.2.2.

3.2.1 A classification of non-anaphoric 3pl readings

In the present section I summarise the classification of 3pl non-anaphoric readings proposed in Cabredo Hofherr (2003). I will briefly present the criteria that this classification is based on and present cross-linguistic data that support the distinctions drawn in the classification (see).

I then briefly summarise some properties of the non-anaphoric 3pl subjects that will
play a role in the analysis proposed in section 3.2.2.

I will assume the following classification of non-anaphoric 3pl subjects into five types:

(24)  

(I) specific existential reading (temporally anchored):

Tocan a la puerta.
``(They) knock.3pl at the door." (=someone is knocking...)

(II) vague existential reading (not temporally anchored):

Han encontrado una motocicleta en el patio.
``(They) have.3pl found a motorbike in the courtyard."

(III) inferred existential reading (inferred from a result):

Aquí han comido mariscos.
``Here, (they) have.3pl eaten seafood." (=someone)

(IV) corporate reading (predicates with a designated subject, see Kaeerde 1943):

Volvieron a aumentar el IVA.
``(They) raised the VAT again."

(V) universal/ locative reading (licensed by a locative):

En España hablan español.
``In Spain, (they) speak.3pl Spanish."[8]

This classification takes several factors into account.
The first division is between the readings (24.I-III) and the readings in (24.IV-V): while the former can be roughly paraphrased by existential quantification (by using a subject corresponding to someone), the latter cannot. Relying on this difference I will distinguish the existential readings (24.I-III) from the non-existential readings (24.IV-V).

The three existential readings can be further differentiated depending on the anchoring of the event in time. While the specific existential reading in (24.I) is anchored to a particular point in time, vague existential and inferred existential readings in (24.II) and (24.III) only imply that an event of the type described has taken place (see Casielles Suárez 1996 for a similar distinction). I will assume that in the specific existential reading (24.I) the event is anchored to a deictic point in time while the vague and the inferred existential readings (24.II/III) are obtained by existential quantification over event instantiations.

The inferred existential reading patterns with the vague existential reading in that it does not imply an anchoring of the event to a precise point in time. Nevertheless, the inferred existential reading in (24.III) has to be distinguished from the vague existential reading in (24.II) since it is subject to further restrictions. The inferred existential reading is only possible if the event can be inferred from a perceivable result (compare the discussion in Toth 2000).

Data from French further support the distinction between the existential readings. First, the French 3pl pronoun ils cannot have a specific existential interpretation while the vague existential reading is possible.

(25) a. Ils nous attaquent. (Fr)
"Someone is attacking us."

ok "They.anaphoric are attacking us." (specific exist.*)

b. Ils ont trouvé une moto dans la cour.

"They have found a motorbike in the courtyard." (vague exist.ok)

Secondly, the evidence from French suggests that the vague existential and the inferred existential reading have to be distinguished: while the vague existential reading is possible for French 3pl ils, the inferred existential reading is not[9].

(27) Ici ils ont mangé des fruits de mer. (Fr)

"Here they have eaten seafood." (inferred exist.*)

Syrian Arabic further confirms the independent status of the inferred existential reading: in Syrian Arabic the specific existential[10] and the vague existential readings are possible, while the examples corresponding to an inferred existential reading were rejected by my informant.

In Spanish, the inferred existential reading (24.III), unlike the vague existential reading, systematically appears with a locative and the perfect haber, "have" + past participle. The constellation in Spanish suggests that the inferred reading is linked to the properties of the perfect, which are beyond the scope of this paper. I will therefore leave the inferred existential reading aside in the analysis of arbitrary 3pl readings given in section 3.2.2.

Apart from the three existential readings, the classification in (24) distinguishes two non-existential readings: the corporate reading and the universal/locative reading (24.IV-V).
The corporate reading (24.IV) arises with predicates such as *deliver the mail, operate patients, raise taxes* that are associated with a designated group carrying out the activity (e.g. postmen, doctors and governments). Pesetsky (1995) coined the term *corporate* to refer to *[a] pronoun [that] picks out some socially designated group of people, prototypically governments, bosses, criminals, or shopkeepers.* Pesetsky further suggests that the existential arbitrary readings are in fact corporate readings. As pointed out by Toth (2000), this cannot be maintained since Hungarian and Spanish allow existential readings with predicates that do not have a designated subject (as for example *sing, knock on the door*).

The fact that in French corporate and the specific existential reading are dissociated provides a second argument in favour of a distinction between the two readings. As already mentioned, the French 3pl pronoun *ils* cannot have a specific existential interpretation (see (25a)); the corporate reading, in contrast, is possible (see Kleiber 1994[11]):

(27) Ils ont encore augmenté les impôts. (Fr)

‘They raised taxes again.’

(corporate reading)

(ex in Kleiber 1994)

Finally, the universal/locative reading differs from the corporate reading in two ways: (i) it does not impose a restriction on the type of predicate, and (ii) it depends on the presence of a locative expression.

The following table summarises the properties of the non-anaphoric readings discussed in this section.
Before turning to the analysis of the five readings that I have distinguished above, I will briefly summarise the main properties of non-anaphoric 3pl readings that will play a role in the discussion.
The most striking property of non-anaphoric 3pl subjects is that they necessarily refer to humans: even if the selectional properties of the predicate force a non-human subject, the only available interpretation is pragmatically anomalous with a [+human] subject.

(29) Aquí ladran en la mañana.  
    ‘Here, (they[+human]) bark in the morning.’

A second property often pointed out for non-anaphoric 3pl pronouns is the fact that they exclude speaker and hearer (see Suñer 1983 for Spanish 3pl pro, Kitagawa and Lehrer 1990 for English they, and Kleiber 1994 for French ils).

This property crucially distinguishes the 3pl arbitrary interpretation in Spanish, French, and English from arbitrary PRO and the impersonal uses of the 2nd person (English you, French tu, Spanish pro.2sg/ tú) that do not exclude the speaker (cf. Hernanz 1990, Kitagawa and Lehrer 1990).

However, the exclusion of speaker and hearer is not generally valid for 3rd person non-anaphoric readings. The non-anaphoric plural null pronoun in Russian and in Modern Hebrew need not exclude the speaker and the hearer, as indicated by the translation by one in the following examples.

(30) a. vo Francii [] edjat ulitok.  
    in France (they) eat.3pl snails

    [] ix kladut v sol' na celij den' i gotovjat s chesnokom
    (one) them put.3pl in salt on whole day and prepare.3pl with garlic (univ/ loc)
‘In France, (they) eat snails. One puts them in salt for a whole day and prepares them with garlic.’

b. be-Savu'ot [ ] 'oxlim givna. (MH)
in-Pentecost eat.benoni.mpl cheese

‘One eats cheese in Pentecost.’ (ex in Borer 1998)

The third property concerns the number feature. For the non-anaphoric 3pl plural pronouns the interpretation is not necessarily [plural]: the existential readings do not imply a plurality (see e.g. Suñer 1983 for Spanish, Cinque 1988 for Italian):

(31) a. ‘¡Que me matan!’ Así clamaba una liebre infeliz que se miraba en las garras de un águila altanera.

‘“That they kill me!” So lamented an unhappy hare that found itself in the claws of a haughty eagle.’ (Samaniego, Fábulas I,5 quoted in Bello 1847)

b. Lo hanno cercato: era un signore anziano. (It)

‘They have been looking for him: it was an elderly man.’(Cinque 1988)

The discussion in this section has dealt with the classification of non-anaphoric readings of 3pl pronouns, be they lexical or null. As the Spanish examples in (24) show, the five readings distinguished here are all attested for null 3pl pronouns. In the following section I will discuss the further licensing mechanisms that yield the different non-anaphoric readings.

3.2.2 Further content identification
According to the analysis presented in section 3.1, non-anaphoric [+human] readings arise with a deficient form of 3pl agreement. As I have further argued in the preceding section that an analysis of non-anaphoric 3pl pro has to account for five types of readings. I therefore propose that the interpretation of non-anaphoric 3pl null pronouns depends on two further factors apart from the phi-feature content.

The first factor is the semantic ambiguity of pronouns. Pronouns can either function as a definite plural, denoting a unique maximal group, or as an indefinite, introducing a variable into the semantic representation.

The translation as a definite plural is motivated by the parallel behaviour of definite plurals and plural pronouns. Both can be used anaphorically, implying contextual uniqueness and maximality of the referent group. For definite plurals anaphoricity and uniqueness can be dissociated as discussed above (see (15)), and I therefore assume that the same dissociation is possible for plural pronouns, resulting in pronouns that refer to a unique maximal group without being anaphoric.

The translation as a variable is independently motivated by bound variable readings of pronouns.

The second factor lies in further mechanisms of content identification.

A non-anaphoric 3pl pronoun that is translated as a unique maximal group receives a corporate or a universal/locative reading, depending on the mechanism of content identification of the group.

For non-anaphoric 3pl pronouns that are translated as a variable, the specific and the vague existential readings arise through existential closure.
(32) 3pl pro translated as

a. a variable ----> existential readings

b. a definite plural ----> non-existential reading

In the following sections I will discuss the means of content identification for the translation of the plural null pronoun as a maximal group and as a variable separately.

**3pl referring to a unique maximal group** In what follows I examine the content identification for the translation of the 3pl pronoun as a unique maximal group. This translation gives rise to the two non-anaphoric readings that are not existential, namely the corporate and the universal/locative reading.

The analysis proposed here assimilates the subjects of the universal/locative and the corporate reading to definite NPs. Such an analysis is further supported by the comparison of null 3pl subjects with a universal reading with lexical NPs in generic sentences. The example (33a) with a non-anaphoric 3pl pro subject has a reading comparable to (33b) with a plural definite NP in that both state a habit or a recurrent pattern. The example (33c) with an indefinite singular subject differs from both (33a/b) in that it implies a modal component of obligation. This contrast is not dependent on the contrast between singular and plural, as the example in (33d) shows: the definite singular unlike the indefinite singular does not necessarily imply a modal reading of obligation.

(33) a. En España, se acuestan temprano. (Sp)

   'In Spain (they) go to bed early.'
b. En España, los niños se acuestan temprano.

‘In Spain children go to bed early.’

c. En España, un niño se acuesta temprano.

‘In Spain a child (modal: should) go to bed early.’

d. En España, el bebé de la familia se acuesta temprano.

‘In Spain the baby of the family goes to bed early.’

As we have seen in section 3.2.1, the non-anaphoric readings of null 3pl pronouns exclude speaker and addressee in Spanish while in Russian and Modern Hebrew this restriction does not hold.

In languages where the non-anaphoric 3pl agreement is not marked for person, the non-anaphoric 3pl pro can receive and interpretation corresponding to people in general, one.

As pointed out by Kleiber for French non-anaphoric ils, in languages where the non-anaphoric 3pl pronouns exclude the speaker and the hearer such pronouns cannot get the reading of people in general, since such a reading would include the speaker and the addressee. Under the assumption that that a simple 3pl pronoun cannot refer by default to a semantically complex group like everyone but you and me, it is then necessary to restrict the reference of the unique maximal group to exclude the hearer and the speaker. The corporate and universal readings arise through different restricting mechanisms.

The first possibility is that the predicate is associated with designated subject, namely a group typically carrying out the activity: this gives the corporate reading. A second restricting mechanism is provided by the identification of the group by a locative expression that
functions like an I-level predicate (see Condoravdi 1989, Casielles 1996) -- this results in the universal/locative reading.

As pointed out by Toth (2001), not any predicate identifying a group is possible; the sentence in (34a), for example, does not have a reading comparable to (34b):

(34)  

a. De viejos necesitan más ayuda.  

of old.3mpl (they.anaphoric) need more help  (non-anaphoric reading *)  

‘In old age, they need more help.’  

b. Compare:  

Los viejos necesitan más ayuda.  

‘The old need more help.’

I propose that this is due to the fact that the adverbial expression *de viejos* in (34a) is interpreted as a restricting when-clause *when (they are) old*. However, such a restriction does not give individuals but temporal stages with a certain property that could apply to any individual, in particular the speaker and the addressee. Consequently, a restriction of this type does not provide a group that excludes speaker and addressee, and therefore such an expression cannot content-license non-anaphoric 3pl *pro*.

Extending Kleiber's (1994) analysis for French, the present analysis treats the collective and the universal arbitrary reading as definites that denote a unique maximal group. The group is unique in the context and identifiable through the locative or the designated subject associated with the predicate. The unspecific interpretation is due to the fact that the individual members of the group are not identified.
The present analysis of the subject of the corporate reading as a definite plural accounts for the fact that these examples cannot be translated felicitously using the indefinite someone as a subject. The existential flavour of the corporate reading is due to the fact that it is not necessary that the entire collective entity take part in the event for a sentence like those in (35) to be judged true (see Kleiber 1994).

(35) a. Les Américains ont débarqué sur la lune en 1969. (Fr)
   ‘The Americans landed on the moon in 1969.’

   b. Jean a touché les feuilles de l'arbre.
      ‘Jean touched the leaves of the tree.’ (exs in Kleiber 1994)

The analysis proposed here cannot explain, however, why locatives are the privileged means of defining a group that is not lexically associated with the predicate. In principle the following temporal expressions should be able to define groups of people that exclude speaker and hearer: nevertheless these examples are less acceptable than examples with a locative adjunct, even if they are clearly better than the example (34a):

(36) a. En la edad media sólo tomaban cerveza (Sp)
    porque el agua estaba contaminada.
    ‘In the Middle ages (they) only drank beer because the water was contaminated.’ (non-anaphoric: ?*)

   b. Durante el servicio militar aprenden a utilizar un arma.
      ‘During the military service (they) learn to use a gun.’ (non-anaphoric ?*)
c. Compare:

En la mili aprenden a utilizar un arma.

‘In the military (they) learn to use a gun.’ (non-anaphoric ok)

The contrast between locatives and other adjuncts may be due to the fact that locatives share properties with nominal subjects. In locative inversion structures for example locatives have been argued to occupy the subject position (see e.g. Bresnan 1994). This affinity between locatives and nominal subjects may provide a lead to an explanation for the restriction to locative adjuncts for universal readings of non-anaphoric 3pl subjects. I have to leave this question open here.

3pl translated as a variable In the present section I propose two mechanisms of content identification for the translation of the 3pl pronoun as a variable. As I have said above, I will set the inferred existential reading aside since the analysis of this reading is probably dependent on the analysis of the perfect, which is beyond the scope of the present paper.

We have seen above that the specific existential reading and vague existential reading can be dissociated in a language like French. Consequently, the two readings must rely on separate licensing mechanisms that may be available separately. I propose that both the specific and the vague existential reading arise through existential closure taking scope over the subject. The difference between the two readings is analysed as a scope difference. The specific existential reading arises through existential closure of the VP (Heim 1982). This existential closure only gives an existential reading if the subject is within VP. The vague existential reading on the other hand relies on existential quantification over the event as a whole, taking scope over Tense.

If this analysis is correct, the difference between French ils and Spanish 3pl pro with
respect to the specific existential reading can be traced back to a syntactic difference. In French subject pronouns occupy a VP-external subject position (spec IP, with subsequent Phonological cliticisation onto the verb) and therefore subject pronouns are not in the scope of the default existential closure that applies to the VP. In Spanish null pronouns occupy the VP-internal position and can stay in the scope of existential closure at the VP-level. If existential closure takes scope over VP, the value for Tense can still be specified, yielding a temporally anchored interpretation.

Given that the present analysis admits that non-anaphoric 3pl pronoun can be translated as a variable, it may seem puzzling that unselective binding by if/when-clauses or Q-adverbs cannot license a generic reading in a language like Spanish. In this respect 3pl pro contrasts with lexical indefinite NPs. Notice, however that since speaker and hearer have to be excluded the unselective binding with if/when-clauses or Q-adverbs cannot give an interpretation corresponding to people in general (see the discussion above). Since if/when-clauses and Q-adverbs do not define a group excluding the speaker and hearer, however, they cannot license a universal non-anaphoric reading.

If this argumentation is correct, this predicts that in Russian and Modern Hebrew where the 3pl does not necessarily exclude the speaker and the addressee null subjects should be licensed by if/when-clauses and Q-adverbs. I have not been able to check this prediction.

4 Anaphoric arbitrary subjects and further questions

In the preceding discussion I have only considered non-anaphoric readings in isolation. Two types of phenomena show that unspecific readings in isolation have to be kept separate from cases where several instances of unspecific readings co-occur.
The first phenomenon concerns two instances of null pronouns co-occurring within a sentence, illustrated in Modern Hebrew and in Finnish.

As shown by Borer (1989), Modern Hebrew admits null subject pronouns with 3rd person inflection that are co-referent with a c-commanding DP in the main clause as in the following example:

(37) Tal 'amar le-’itamar Se-[ ] ya-cli’ax. (MH)

Tal said.past.3ms to-Itamar that-((pro)*) 3ms-succeed.fut

‘Tal said to Itamar that he (Tal or Itamar) is going to succeed.’(ex in Shlonsky 1997)

This option is only available, as pointed out by Shlonsky, if the verb in the matrix clause is past or future (i.e. marked for person). This example shows that further licensing mechanisms for null 3rd person pronouns are available with a local antecedent. Given the analysis of non-anaphoric readings here, it is therefore possible that a null 3rd person pronoun with an unspecific interpretation is licensed by an unspecific antecedent in the matrix clause. In such a configuration the second arbitrary subject would be dependent on a local arbitrary antecedent.

A second example of locally co-occurring null 3rd person pronouns can be found in Finnish. Hakulinen & Karttunen (1973) in their discussion of generic 3sg null pronouns in Finnish identify several licensing conditions that are necessary to license a generic null 3sg. They show that the presence of a pronoun with an non-anaphoric reading in the superordinate clause may license another null pronoun with the same interpretation in a that-clause, even if the licensing conditions for the second null pronoun are not fulfilled, as in the following
example.

(38) Tässä työssä [ ] vaaditaan, että [ ] osaa ruotsia.  
This work requires of you that you know Swedish.  
(ex 31b in H & K)

A second phenomenon that may independently allow null 3rd person subjects with an unspecific reading is anaphoric dependency on an unspecific antecedent. This case is illustrated by the following fragments taken from Spanish:

(39)  
a. En esta mina pro trabajan mucho.  
pro Despiertan a las 5 y media, pro salen a las 6 de la mañana y  
pro regresan a las 8 de la noche, y si pro llegan más tarde, el capataz  
se enfada con ellos. Y a fines del mes les pagan una miseria.  

‘In this mine (they) work a lot. (They) wake up at 5 o'clock, (they) leave at 6 in the morning and come back at 8 at night, and if (they) arrive later, the supervisor gets annoyed with them. And at the end of the month (they) pay them a misery.’  

(b. Salen a las 6 de la mañana y regresan a las 8 de la noche.  
‘(They) leave at 6 in the morning and come back at 8 at night.’  

(c. En esta mina trabajan mucho.  
‘In this mine (they) work a lot.’
(40)  a. En Francia pro comen caracoles. pro Los ponen en sal durante 12 horas y después pro los preparan con mantequilla y ajo.

``In France (they) eat snails. (They = people in France) put them in salt for 12 hours and then (they) prepare them with butter and garlic.''

(b) pro Los ponen en sal durante 12 horas.

``(They) put them in salt for 12 hours.''

The locative that licenses a universal reading is only present in the first sentence of the fragments (39a)/(40a); nevertheless the unspecific reading is taken up by the boldface pronouns in the later sentences even though the necessary locative is not present locally. In isolation such examples do not receive an arbitrary reading as (39b) and (40b) show. Notice that even object pronouns (les) and strong pronouns that occur as complements of prepositions (ellos) can take up the unspecific reference of the non-anaphoric subject in the first sentence.

As the comparison of (39a) and (39b) shows, pronouns that cannot have a non-anaphoric reading in isolation can receive an unspecific interpretation in a fragment where they co-occur with a subject that satisfies the licensing conditions for an non-anaphoric reading (cf. (39c)).

The same behaviour can be observed with the corporate reading as illustrated in (41).

(41)  a. Ayer arreglaron mi computadora. pro Llegaron a las 10 de la mañana y se fueron a las 6 de la tarde.

‘Yesterday (they) repaired my computer. (They) came at 10 in the morning
and left at 6 in the afternoon.’

b.  *pro* Llegaron a las 10 de la mañana y se fueron a las 6 de la tarde.

‘(They) came at 10 in the morning and left at 6 in the afternoon.’

I propose that the boldface pronouns in (39a), (40a) and (41a) are not to be analysed as instances of non-anaphoric 3pl pronouns. The interpretation of these pronouns is unspecific since they are anaphoric to an unspecific antecedent. The same unspecific interpretation may arise with lexical pronouns in English, as in the following example:

(42) Alguien/ un sicópata asesinó a Holmes. *pro* Lo esperó aquí, le disparó y después llevó el cadáver hasta el río, donde lo encontró la policía.

‘Someone/a madman killed Holmes. (He) waited for him here, shot him and then dragged the body to the river, where the police found it.’

The boldface 3sg *pro* in the second sentence and the corresponding pronoun *he* in the gloss refer to an unspecific individual since in both cases the pronouns is co-referent with the unspecific subject *someone / a madman* of the first sentence in the fragment.

In view of the data that I have just discussed, I have based the discussion in the previous sections on isolated examples of non-anaphoric readings in order to exclude interference of this type of reading that has a non-anaphoric flavour but is not itself non-anaphoric.

Now, if the distinction anaphoric /non-anaphoric *pro* proposed here is justified, this predicts that in languages that do not have anaphoric 3rd person *pro* there should be no arbitrary 3pl *pro* by co-reference with a vague antecedent as in the Spanish examples (39),
According to my Russian informant, this prediction is correct for Russian, as the following examples show.

(43) včera (*oni) mne pochiniš kompjuter.  
    yesterday (they) me.dat repaired.3pl computer

    (*oni) prishli v 10 chasov ...  
    (they) came.3pl at 10 o'clock

(44) a. Na etom zavode (*oni) mnogo rabotajut.  
    in this factory (they) much work.3pl

    (*oni) prixdjat v 6 chasov utra ...  
    (they) come.3pl at 6 o'clock morning.gen

b. vo Francii (*oni) edjat ulitok.  
    in France (they) eat.3pl snails

    (*oni) ix kladut v sol' na celij den' i gotovjat s chesnokom  
    (they) them put.3pl in salt on whole day and prepare.3pl with garlic

According to my Russian informant, the subject of the first sentence in (44b) cannot be taken up neither by a lexical nor by a null 3pl pronoun. If the subject of the second sentence in
(44b) is a null 3pl pronoun, this pronoun does not take up the subject of the first sentence (i.e. roughly "people in France, the French"), but can only be understood as an instruction to prepare snails (corresponding to a sentence with the subject "one" in English).

In Modern Hebrew, the prediction is only partially borne out. As predicted by the present analysis, a 3pl null subject with corporate interpretation cannot be taken up by a null 3pl pronoun: in the example (45) the anaphoric subject of the second sentence has to be a lexical 3pl pronoun hem.

(45) 'etmol [ ] tiknu li 'et ha-maxSev
(MH)
yesterday (they) repaired to-me acc the-computer
‘Yesterday they repaired my computer.’

??(hem) ba'u be-'eser ve-halxu be-SeS ba-'erev
they came at-ten and-left at-six at-night (corp)

In the first example with a locative licensing the universal/locative reading in (46), the null subject in the second sentence is not interpreted as anaphoric on the 3pl pro in the first sentence. Like in Russian, the only reading for a null 3pl subject is that of a general instruction, a recipe.

(46) be-carfat [ ] 'oxlim Sablulim
(MH)
In-France (they) eat.3mpl snails.
samim 'otam be-melax le-mSex yom
(they) put.3pl them in-salt for-one day

ve-'az [ ] mevaSlim 'otam be-Sum
and-then (they) cook.3mpl them in-garlic. (univ / loc)

For the second example of a reading licensed by a locative expression, Modern Hebrew
differs from Russian, however: in the example (47), the non-anaphoric 3pl null subject of the
first sentence can be taken up by 3pl null pronouns in the subsequent discourse.

(47) ba-maxane ha-ze, [ ] 'ovdim harbe. (MH)
in-camp the-this, (they) work.3mpl much.

kamim be-SeS ba-boker, [ ] matxilim la'avod be-Seva
(they) get-up at 6 in the morning, (they) start working at-7 o'clock

lo mafsikim 'ad Seva ba-'erev
and (they) only finish at 7 in the evening,

ve-be-sof ha-xodeS [ ] mekablim saxar 'aluv
and-at the-end of the month (they) earn (a) miserable wage. (univ / loc)

The present analysis does not provide an explanation for this contrast in Modern Hebrew.
Further investigation of the role of the locative could shed light on the difference between the
two locative readings, but I have to leave this question open here.
Notice that even in languages that have anaphoric 3rd person null subjects, not all non-anaphoric subjects can be taken up by a null pronoun, as illustrated by the example with a specific existential subject in the following example.

(48) Tocan a la puerta. *pro Quieren entrar. (Sp)
    (they) knock at the door. (they) want come-in.inf

In Russian and Modern Hebrew the specific existential reading cannot be taken up by a following 3pl pronoun either, be it null or lexical.

(49) a. (*oni) stuchat. *pro/*oni xotjat voiti. (Ru)
    (they) knock. (they) want come-in

b. dofkim ba-delet. *pro/*hem rocim le-hikanes. (MH)
    (they) knock at-the door. (they) want come-in

The preceding examples show that the different readings of non-anaphoric 3rd person subjects differ with respect to discourse transparency: while the corporate readings can be taken up in the discourse in Spanish and in Modern Hebrew, the 3pl specific existential reading cannot be taken up in either of the languages considered here.

In the present section I have presented data showing (i) that null 3rd pronouns in a finite subordinate clause can be licensed by an antecedent in the matrix clause independently of the possibility of anaphoric pro-drop, and (ii) that 3pl pronouns can receive an unspecific reading by co-reference with a non-anaphoric antecedent.
I have further argued that not all readings of non-anaphoric 3pl pronouns are discourse-transparent introducing a referent into the discourse: the data from Spanish, Russian and Modern Hebrew suggest in particular that the specific existential reading does not introduce a discourse referent.

5 Conclusion

In this paper I have examined non-anaphoric readings of 3pl pronominals. I have given arguments showing that non-anaphoric 3pl do not pattern with anaphoric 3pl pronouns with respect to pro-drop. In fact, the distribution of non-anaphoric 3pl is closer to that of quasi-argumental subjects than to that of anaphoric argumental 3rd person pro.

In order to account for this observation, I have proposed to replace Rizzi’s typology of pro-drop based on the type of theta-role by a typology based on the distinction between anaphoric and non-anaphoric pro.

According to the analysis proposed here non-anaphoric pro arises with deficient agreement that only identifies a subset of phi-features. Developing this hypothesis, I have argued that there are two types of deficient agreement. The deficient 3sg-agreement does not contain number or person features and functions like a mass noun being interpreted as inanimate by default. The deficient 3pl-agreement, in contrast, may be marked for 3rd person or not, depending of the language, but is always underspecified for number and behaves on a par with count nouns, its default interpretation being [+human].

In order to account for the different types of non-anaphoric 3pl pro I have further proposed that 3pl pronouns can be translated either as referring to unique maximal groups, comparable to definite plurals, or as variables, comparable to indefinites.
As a consequence of the latter, certain non-anaphoric 3pl readings resemble indefinites (specific and vague existential reading), while others resemble definite plurals (corporate and universal/locative reading).

Finally, I have discussed some predictions of the present analysis concerning the occurrence of 3rd person null pronouns that are anaphoric on a non-anaphoric antecedent.

Bibliography


Endnotes

[*] I want to thank Ildikó Tóth for many discussions on arbitrary pronominals. I owe Jean-Marie Marandin and two anonymous reviewers many comments on an earlier version of this work that have led to significant improvements. I particularly want to thank Brenda Laca for her detailed comments. I am very grateful to Nisrine Al-Zahre, Nora Boneh, Gilles Boyé, Adil El Ghali, Ora Matushansky and Kristiina Saarinen for help with their native languages. All errors and misinterpretations are my responsibility.

[1] In the examples, the languages are indicated as follows Bav= Bavarian, Fi= Finnish, Fr= French, Ge= German, Gr= Greek, Ice= Icelandic, It= Italian, MH= Modern Hebrew, Sp = Spanish, Ru= Russian. The following abbreviations are used in the glosses: m(asculine), f(eminine), fut(ure), inf(initive), par(titive), ela(tive), neg(ation), pres(ent).

[2] See Toth (2000) for a detailed discussion of the difference between content-assignment to pro and content-identification of pro. In Toth's terms the present analysis assumes content-assignment to a formally licensed null pronoun. Since this distinction does not bear on the following discussion, I will retain Rizzi's term content-identification in what follows.

[3] Following the terminology familiar from the literature, I use the term arbitrary pronoun to refer to pronouns without an antecedent. It has to be stressed, however, that the term arbitrary pro for non-anaphoric null pronouns is misleading, since (i) non-anaphoric readings do not form a homogeneous class and (ii) non-anaphoric null pronouns clearly differ from uncontrolled PRO (arbitrary PRO), see the discussion in section 3.2 below.
In Homberg & Nikanne (1994) the use of a subject *se* is indicated as optional with weather verbs. My Finnish informant does not accept the subject *se* with weather verbs. She marginally accepts *se* if the weather verb is embedded under an emphatic expressions. The alternation zero/*se* might therefore be comparable to a phenomenon that is observed in French and German where the subject of weather predicates can alternate with a deictic expression under emphasis, *il/ça pleut* (Fr), *es/das regnet* (Ge), ‘it/that rains’.

Null subjects are possible in spoken Russian, these null surface subjects behave like null topics and not like null pronouns, cf. Matushansky (1998).

In the following discussion I will leave aside Finnish and Icelandic since in these languages only the 3rd person singular *pro* seems to be available (cf. Hakulinen & Karttunen (1973) for Finnish, Sigurdsson (1989) for Icelandic).

This correlation is visible in certain languages where plural is obligatory only for [+human] entities, such as e.g. Kriyol as discussed by Kihm (2003).

This class covers the core-cases of the quasi-universal readings in the sense of Cinque (1988). Notice, however, that depending on the locative and the predicate it combines with the group defined by the locative expression need not be quasi-universal as the following example shows.

(i) En este colegio enseñan ruso.
    in this school (they) teach.3pl Russian.
[9] The closest equivalent to an inferred existential reading involves the use of the impersonal pronoun *on*:

(i) Ici on a mangé des fruits de mer. (Fr)
    here ON has eaten seafood.

[10] My informant preferred the impersonal passive to a specific existential 3pl pro.