

Please do not quote without permission

Contents

1	Introduction	5
1.1	Variation and Invariance in Generative Grammar	5
1.2	The Null Subject Parameter	7
1.3	Variation in the Minimalist program	9
1.4	Organization of the book	10
I	What is the Null Subject Parameter?	11
2	The Null Subject Parameter. Introduction	13
2.1	Syntactic properties associated with the NSP	13
2.1.1	Null subjects	13
2.1.1.1	Typology of null expletive subjects	16
2.1.2	Free Subject Inversion	18
2.1.3	Null resumptive pronouns	19
2.1.4	Violation of the <i>that</i> -trace filter	19
2.1.5	Long wh-extraction	21
2.1.6	Section summary	22
2.2	Interpretive differences between null and overt subjects	22
2.2.1	Arbitrary and generic interpretations	23
2.2.2	Reference to clauses	24
2.2.3	Quantifier Binding	25
2.3	Typology of NSLs	26
2.3.1	Consistent null subject languages	27
2.3.2	Discourse-related null subject languages	27
2.3.3	Partial null subject languages	29
2.3.4	Section summary	31
2.4	Chapter summary	31
3	The core content of the NSP	33
3.1	Introduction	33
3.2	Property correlations	33
3.2.1	Null thematic subjects & no null expletives	35

3.2.1.1	Overt expletives in Dominican Spanish	36
3.2.1.2	A CP-related expletive	39
3.2.2	Subject inversion & no null expletives	44
3.2.3	<i>That</i> -trace effects & overt expletives	45
3.2.4	Free subject inversion and <i>that</i> -trace effect violations	46
3.2.4.1	Bavarian German and Övdalian	46
3.2.4.2	Levantine Arabic	48
3.2.4.3	Dominican Spanish	49
3.2.4.4	Brazilian Portuguese	50
3.2.4.5	Subject extraction in Lubukusu	52
3.3	Summary	53
3.4	Variation in logophoricity	53
3.4.1	Person effects in inverse copular clauses in Brazilian Portuguese and Spanish	54
3.4.2	Person effects in Shipibo switch-reference clauses	55
3.5	Summary	56
4	The nature of the EPP and the Null Subject Parameter	59
4.1	Introduction	59
4.2	<i>Pro</i>	60
4.2.1	The existence and position of <i>pro</i>	60
4.2.1.1	Presentational sentences	61
4.2.1.2	Copular sentences with nominal predicates	62
4.2.1.3	Floated quantifiers	63
4.2.1.4	Agreement in the Ancona variety	65
4.2.1.5	Section summary	66
4.3	AGR/INFL. The pronominal agreement hypothesis	66
4.3.1	Pronominal agreement and reference	67
4.3.2	Pronominal agreement and ellipsis	70
4.3.3	Arguments against the pronominal agreement hypothesis from Irish	71
4.3.4	Arguments from Finnish expletives against the pronominal agreement hypothesis	72
4.3.5	Word-order asymmetries in Agreement	73
4.4	Deletion	74
4.4.1	Finnish expletives and NSs	74
4.4.2	<i>Pro</i> as ellipsis	76
4.4.3	Null/overt asymmetries	77
4.5	Is the EPP universal?	79
4.5.1	Evidence for expletives in NSLs	79
4.5.2	Positional freezing	83
4.5.3	The Subject Criterion revisited	86
4.5.3.1	Subject and topichood in Oromo	87
4.6	The revised NSP	88
4.7	Summary	89

II On Identification 91**5 Identification and morphology 93**

5.1	The minimal morphological threshold (MMT)	95
5.2	The structure of ϕ -features	98
5.2.1	Φ -feature hierarchies	98
5.3	AGREE	101
5.4	Deriving NSs in several languages	104
5.4.0.1	MMT set to person, number and gender: Tarifit	104
5.4.1	MMT set to person and number: Italian and Spanish	106
5.4.2	MMT set to person: Bengali	109
5.4.3	Gender in Arabic	111
5.4.4	Mixed paradigm in Hebrew	111
5.4.5	Mixed paradigms in German Vernaculars	113
5.4.6	Summary	115
5.5	Identification and locality	116
5.5.1	On the locality of head-agreement: Standard Germanic vs. C-agreeing German Vernaculars	116
5.5.2	Local agreement vs. distant reference: First and second person shifted readings	119

6 Discourse identification 121

6.1	Topics and <i>pro</i>	121
6.2	Typology of Topics	123
6.3	Topic-identification in an inflection-rich NSL	125
6.3.1	Absolute clauses in Spanish	125
6.3.2	Adjunct clauses in Hebrew	126
6.3.3	Subjunctive clause NSs	127
6.3.4	Topic drop	130
6.3.5	Constraining the system	131
6.4	Subject identification in discourse-NSLs	133
6.4.1	Two types of nominals	135
6.4.2	Shipibo and Capanawa	137
6.5	Summary	139

7 Null/overt subject contrasts 141

7.1	<i>Pro</i> as a weak pronoun	142
7.2	On the difference between a null and overt pronouns	143
7.2.1	Binding effects	144
7.2.2	Backwards coreference in temporal adjunct clauses	146
7.2.3	Backwards anaphora in conditional sentences	148
7.3	Deriving expletives	151
7.3.1	Full and default agreement with associate DPs	152

7.3.2	Reference to clauses	154
8	The status of preverbal subjects in NSLs	155
8.1	Preverbal subjects as CLLD phrases	156
8.1.1	Scope asymmetries	156
8.1.2	Extraction facts	159
8.1.3	Adverb intervention effects	161
8.1.4	PS incompatibility with quantified CLLD phrases	162
8.1.5	The distribution of Bare NPs	163
8.1.6	Section summary	164
8.2	Against PS-as-CLLD	165
8.2.1	Quantified PS and CLLD phrases	165
8.2.2	Discourse status of PS and CLLD phrases	166
8.2.3	Possible antecedents	169
8.2.4	Section Summary	170
8.3	The status of PS	170

Chapter 1

Introduction

1.1 Variation and Invariance in Generative Grammar

Perhaps one of the most revolutionary tenets that Generative Grammar assumed from its outset was the idea that all languages share common underlying grammatical features. This belief stemmed from different sources, initially from observations about the way in which children learn languages. In particular, as Chomsky has repeatedly observed (cf. Chomsky, 1975, 1988, for example), children learn grammatical rules and patterns for which they have no overt evidence (the so-called poverty of stimulus or Plato's problem) and they do not assume certain grammatical rules that could be generalized from the available data.¹ Chomsky has concluded from these observations that much of the grammatical knowledge we have must be innate, and that the grammars of language are much closer to each other than it would seem at first sight. Both of these assumptions together make the process of language acquisition (or language discovery, to be more precise) easier: if children come with an innate predisposition for languages that contains fairly specific and delimited principles, then in contact with the environment, children will simply make sense of what they are hearing guided by such innate principles. The guiding principle that languages are at core very similar led to important discoveries of how what looks like very different patterns on the surface turn out to have very similar representations.

The other side of the universalist tenet is the need to explain how surface variation is realized across individual grammars. While languages may look similar in having noun-like words, or verb-like words, they also differ with respect to other categories. So, for example, English's meager representation of aspectual distinctions on its verbs pales in comparison to Russian's very complex aspectual system. This tension between underlying similarity and surface variation has been formalized in many different ways over the years. In the 1980s framework of GG, invariance was supposed to be a matter of principles, whereas variation was a matter of parameters. A principle was conceived as a rigid statement about grammar that was common to all languages, and should therefore be instantiated in some way in every language. A parameter, on the other hand, was seen as the delimited range of variation that accounts for the observed differences across

¹Needless to say, the poverty of stimulus argument remains a controversial issue in the broader cognitive-science community.

languages.

To take a particular example of this approach to variation~invariance, certain languages like English or Spanish displace the *wh*-words in questions to the beginning of the clause, whereas others (like Chinese) leave those *wh*-words in the same position as in their non-question counterparts, as illustrated in (1)-(2).

- (1) a. ¿**A quién** piensa Pedro que viste? (Spanish)
 to whom thinks Pedro that saw
 ‘Who does Pedro think you saw?’
 b. Pedro piensa que viste a María
 Pedro thinks that saw to maría
 ‘Pedro thinks that you saw María’
- (2) Zhangsan yiwei Lisi mai-le **shinwe** (Chinese)
 Zhangsan thinks Lisi bought what?
 ‘What does Zhangsan think Lisi bought?’

Closer examination of these languages suggested that both types had similar properties. In particular, Huang (1982) argued that languages with *wh*-words in situ have scope properties that are very similar to those of displaced *wh*-word languages. Within the original formulation, the differences between *wh*-in situ languages and displaced *wh*-languages was formalized as a difference in the level of application of the movement rule: In Spanish, this rule applies in overt syntax, in Chinese, it would apply after overt syntax, at LF. Thus, an invariant principle (*wh*-words must take scope over the clause) can be expressed in two ways: through overt or covert movement.² While the validity of the generalization proposed by Huang has been subsequently challenged, it exemplifies well the overall research strategy within the generative paradigm, as well as a specific formulation of how an underlyingly common rule can yield superficial variation.

Ideally, parameters should show at least the following characteristics: a) determine the range of variation within a given principle, b) delimit a cluster of properties related to the phenomenon under analysis, and c) identify, from among the cluster of properties, which one is the primary, necessary one that drives the parameter.

As noted, having a common underlying principle with a limited range of variation facilitates the process of L1 acquisition: If the child’s knowledge includes the notion that *wh*-words must take scope, determining whether it moves overtly or covertly becomes a matter of processing the available input.

Within the Principles and Parameters version of the 1980s, the Null Subject Parameter (NSP) was one of the most studied and formalized instances of invariance~variation. This parameter tried to account for the fact that in certain languages, subjects must be obligatorily overt, whereas in others, subjects can be overt or null. As we will see below, in the classical formulation of the parameter, null subject languages (NSL) allow the subject to be null because their potential reference can be recovered through the morphology of the verb.

²In more recent theoretical approaches, the difference can be cast in terms of where the copy of the *wh*-word surfaces overtly.

1.2 The Null Subject Parameter

The NSP attempts to provide a unified analysis for the observation that clauses require obligatorily overt subjects in some languages like French and English but not in other languages (Spanish, Italian, Chinese), as illustrated in (3)-(4).

- (3) a. Chuirfidís. (Irish)
 put.COND.3.PL
 ‘They would put.’
 b. Chuirfeadh siad.
 put.COND. they
 ‘They would put.’ (from McCloskey and Hale, 1984, 491, ex. 8)
- (4) a. We left. (English)
 b. *Left

Given these two types of languages, one could simply state a typological description such as (5). While this is a reasonable description, it raises several other issues, for example, given that the verb *chuirfeadh* ‘would put’ is interpreted as having an agent or theme regardless of whether the subject is present or not, how is this thematic property formalized in a unified way?

- (5) Languages of the world vary among those that have obligatory subjects and those that have optional subjects.

An alternative approach, which is the one adopted by GG, was to assume that all languages have subjects and the variation comes from whether the overt expression of that subject is obligatory or not. The explicit formalization of this general idea was initially derived from two distinct principles, the first one is a general principle (the **Projection Principle**, originally formulated in Chomsky (1981, 38)), which ensures that if a given head is lexically specified as assigning a theta role, that role must be assigned to a syntactically realized constituent, and this constituent must be present at all levels of representation. For example, a transitive verb like *eat* with two theta roles, requires two syntactic arguments (subject and object) that receive the theta role and they must be realized at all times.

The Projection Principle makes thematic subjects obligatory in all languages. Thus, assuming that most verbs usually have a theta role to assign, they will require at least one argument, and this argument will frequently surface as a subject.³ It follows that whenever a language does not show an overt subject, a null argument must be present to receive the verb’s theta role, otherwise the Projection Principle will be violated.

This formulation does not account for why syntactic subjects seem to be obligatory in languages like English even when the subject is semantically vacuous, as in (6)-(7). Since *seem* in (6a) does not assign a theta role to its subject, the Projection Principle does not require an overt argument, however, its absence renders the example ungrammatical, as seen in (6b).

³The Projection Principle does not directly derive the requirement that the unique verb of an argument be the subject in the sense of the argument in Spec, IP. In the Principles and Parameters framework, this is a by-product of the fact that nominative case was assigned in Spec, IP.

- (6) a. It seems to be raining. (English)
 b. *Seems to be raining.

The situation is slightly different for the expletive in (7a): the verb *surprise* does assign two thematic roles, one of them to the indirect object *me*, the other to the clause *that you couldn't finish your meal*, but the expletive in the preverbal subject position arguably does not receive a separate thematic role. In fact, when the clausal subject appears initially, as in (7b), the expletive is no longer possible. This suggests that the clause is the thematic subject both in (7a) and b), and that the expletive is somehow doubling that subject in (7a). Likewise, it shows that the presence of the expletive in (7a) is not related to the Projection Principle, since the subject theta role is assigned to the clause.

- (7) a. It surprised me [that you couldn't finish your meal]. (English)
 b. [That you couldn't finish your meal] surprised me.

The facts just described regarding expletives led to the principle in (8), which essentially captures the fact that in English and other languages, even non-theta-assigning verbs require an overt expletive subject. The qualification that the subject must be in Spec, IP is meant to derive the difference between (7a) and b). In the first case, the clausal subject is not in Spec, IP, hence an expletive is required.

- (8) All clauses must have a subject (in Spec, IP).

Both the Projection Principle (the requirement that thematic arguments be present throughout the derivation) and the requirement that clauses have subjects constitute the **Extended Projection Principle** (EPP) (cf. Chomsky, 1986, 116 and Svenonius, 2002, 9 for a summary), although EPP is frequently used with the more restricted meaning that clauses require subjects.

Once one assumes a principle like the EPP, it follows that languages are much more similar than what (5) would suggest, and it also follows that whenever a clause lacks an overt subject, it really is syntactically present, but not seen. Thus, the examples in (3a, b) and (4a) above have similar underlying representations, perhaps along the lines of (9). In (9a) *pro* represents a null subject in Irish.

- (9) a. *chuirfidís pro*.
 b. We left.

Assuming some invariant principle like the EPP shifts the research questions in a completely different direction: Given that English-like and Irish-like languages are substantially identical at the level of semantic interpretation, what is it that allows Irish to have a null subject? As we will see below, many answers have been given to this question, but one traditional and influential intuition was that the inflectional information on the verb identifies the null subject in the Irish example but not in the English ones. Irish is particularly suggestive in this respect, because it has two types of tenses, those that have distinctive person and number morphology and those that do not. Null subjects are only possible with those that have distinctive morphology (cf. McCloskey and Hale, 1984).

1.3 Variation in the Minimalist program

With the advent of the Minimalist Program (MP) in the early 1990s (cf. Chomsky, 1993 and much subsequent work), much of the theoretical machinery that made it possible to express parametric variation was eliminated. Given the stated goal of simplifying the overall theoretical model, parameters no longer have an independent place in accounting for variation across grammars. Rules or principles that apply only at certain levels of representation are also avoided. As a result, the analysis for overt/covert *wh*-movement suggested in previous sections would need to be retooled, since rules can no longer apply at different levels. Rather, the MP restricts variation to two, perhaps related sources: a) Differences in specifications of lexical items and b) where copies of items surface. For example, in early minimalist formulations, the difference between verb movement in French and English was related to the strength of a lexical feature of inflection, which resulted in attracting the verb in French but not in English. Feature strength is a property of individual lexical entries (in this case inflection). In the case of *wh*-movement in English vs. Chinese, the difference is not whether the *wh*-word moves earlier or later: It moves in both languages at the same time, but in English the lower copy of that movement is deleted, whereas in Chinese, the lower copy remains.

Over the past several years, other theoretical constructs have changed radically. The notion of agreement has become central in articulating syntactic relationships between constituents, but in recent formulations, agreement is disconnected from movement, so that two elements can agree at a distance. As a result, the triggers for constituent movement are now even less obvious than before, and in general they have been subsumed under the notion of EPP features, a short-hand term for describing that a constituent has moved.

All of these changes have forced researchers to undertake a radical revision of the NSP. From proposals that attempt to eliminate *pro* as an independent theoretical construct to revisions and parametrization of the EPP (cf. Alexiadou and Anagnostopoulou, 1998; Tomioka, 1999; Ordóñez and Treviño, 1999; Manzini and Savoia, 2002; Holmberg, 2005; Saab, 2009; Barbosa, 2009; Biberauer et al., 2010; Sigurðsson, 2011 among many others), the interest in how to account for null subjects within minimalist guidelines has triggered many papers. Extending empirical coverage of the null subject phenomenon outside of the better-known cases has been less of a concern, with some important exceptions. One of them concerns the evolution of NS varieties into overt-subject grammars in cases such as European and Brazilian Portuguese (cf. Duarte, 2000; Kato, 1999, 2000; Kato and Negrão, 2000; Modesto, 2000; Barbosa et al., 2005; Modesto, 2008; Camacho, 2008, 2010a), and Caribbean and non-Caribbean Spanish (cf. Cabrera, 2007; Camacho, 2008, 2010a, as well as other languages or constructions (cf. Cole, 2000, 2009; Nicolis, 2008; Shlonsky, 2009; Camacho and Elías-Ulloa, 2010; Camacho, 2011; Roberts, 2010b).

In this book, we review the current status of the NSP. First, I review the properties that were ascribed to the NSP in its original formulation and how subsequent research has sharpened our understanding of those properties and their connection with null subjects. Then, I will propose revisions to the original NSP that take into account those findings and changes in the overall theoretical framework. As a result, I propose a NSP much in the spirit of the original proposal but which avoids some of the shortcomings that have been noted over the years.

1.4 Organization of the book

The first part of this book introduces the core properties of the NSP. Chapter 2 presents the syntactic and interpretive properties originally associated with the parameter, as well as the general typology of null subject languages. Chapter 3 reviews what some of the leading proposals on NSs predict regarding the association of those core properties. I will suggest that many of the originally proposed properties do not correlate with the availability of NSs. Chapter 4 introduces existing proposals on how the EPP is satisfied in NSLs. Specifically, I will review proposals that endorse the independent availability of a null pronoun *pro*, those that assume that *pro* does not exist, but rather subject-properties are encoded in verbal inflection (the pronominal agreement hypothesis), and those that suggest that *pro* is the product of ellipsis. By the end of the first part, what I hope is a clear picture will emerge of the essential properties of the NSP and also of the minimal syntactic primitives required to account for them.

In part II, I address the issue of how NSs are identified. Chapter 5 discusses how NSs are identified in a clause in connection with overt morphological inflection, and chapter 6 looks at discourse identification when morphology fails. Chapter 7 considers a number of instances where the null/overt contrast determines the distribution of subjects, and finally, chapter 8 examines the dislocated properties of overt subjects.

Part I

What is the Null Subject Parameter?

Chapter 2

The Null Subject Parameter. Introduction

2.1 Syntactic properties associated with the NSP

As mentioned in the introductory chapter, the NSP derives from the idea that all languages have an subjects (the EPP, cf. (8) in 1.2). Languages that apparently lack subjects have null versions (both thematic and expletive), and this parametric setting correlates with a cluster of syntactic properties. The five initial properties that were proposed to be related to the NSP included a) having null subjects, b) having free inversion in simple sentences, c) availability of “long wh-movement” of subjects, d) availability of empty resumptive pronouns in embedded clauses and e) presence of overt complementizers in **that*-trace contexts (cf. Perlmutter, 1971; Chomsky and Lasnik, 1977; Kayne, 1980; Jaeggli, 1982; Jaeggli and Safir, 1989; Safir, 1985; Taraldsen, 1980; Rizzi, 1982). In addition, languages with null subjects have been argued to display certain interpretive differences (cf. Montalbetti, 1984, among others). For example, overt pronouns in a NSL are usually emphatic and/or focused, whereas null subjects usually signal topic continuity. In this section, we will review and exemplify the different properties associated with the NSP.

2.1.1 Null subjects

The first property, namely the availability of null subjects, was illustrated earlier, and it can be seen again in examples (1)-(3). Certain languages allow subjects to be null: In languages like Chamorro and Irish, when person subject agreement is present, pronouns must be null, as seen in (1) for Chamorro. In other languages like Spanish or Quechua, overt pronouns are optional, as seen in (1)-(3).

- (1) Hu-fahan (*yu') i lepblu. (Chamorro)
AGR-buy I the book
'I bought the book.' (from Chung, 2003, 552)
- (2) a. Salimos. (Spanish)
left
'We left.'

b. Nosotros salimos.
we left
'We left.'

(3) a. Huwan-mi papa-ta mikhu-n. (Quechua)
Huwan-FOC/EVID potato-ACC eat-3.SG
'Huwan eats potatoes (attested).'

b. Papa-ta mikhu-n-mi.
potato-ACC eat-3.SG-FOC/EVID
'S/he eats potatoes.' (from Sánchez, 2010, 24)

Subjects can be thematic or expletive, and some languages allow both types of subjects to be null, or just one of them. Thus, in principle we find four possible combinations of null/overt and expletive/thematic, although languages with null thematic subjects and overt expletives are rare.

a) Null thematic subjects/null expletive subjects. Spanish, Chamorro and Quechua have both null thematic subjects (illustrated above) and null expletives, seen in (4)-(6). These examples illustrate an asymmetry between expletive and thematic subjects: Whereas thematic subjects can be null or overt in Spanish, expletives must be null (cf. (4a) vs. (4b)).¹

(4) a. Nieva. (Spanish)
snows
'It is snowing.'

b. *Ello nieva.
it snows
'It is snowing.'

(5) Kao [put fin] um-uchan? (Chamorro)
Q by end AGR-rain
'Did it rain in the end?' (from Chung, 1998, 337)

(6) a. pro para-chka-n-mi. (Quechua)
pro rain-PROG-3.SG-FOC/EVID
'It is raining.'

b. *Pay/*Kay para-chka-n-mi.
(s)he/this rain-PROG-3.SG-FOC/EVID
'It is raining.' (from Sánchez, 2010, 24)

b) Overt thematic subjects/null expletive subjects. In Cape Verdean Creole, thematic subjects must be overt but expletives are null, as shown in (7) and (8).

(7) Sta livri. (Cape Verdean Creole)
is/are free
'I/you/he/she/we/they are free.' (from Baptista, 2002, 255)

¹Example (4b) is grammatical in El Cibao varieties of Dominican Spanish, see sect 3.2.1.1.

- (8) a. Sta faze kalor oji. (Cape Verdean
is make heat today
Creole)

‘It’s hot today.’

- b. Ma gosi nés tenpu, parse ki ta nase mas.
but now in-this time seem COMP TMA be born more
‘But it would seem that in these times, more are being born.’

(from Baptista, 2002, 64-65)

c) Null thematic subjects/Overt expletives. In Finnish, thematic subjects can be null whereas some expletives are obligatorily overt. Thus, the subject *minä* ‘I’ is optional in (9), but the overt expletive *sitä* is obligatory in (10), if nothing else occupies the Spec, IP position. In (10c-d), another constituent has moved to that position, so the expletive is optional (cf. sect. 4.3.4).

- (9) (Minä) puhun englantia. (Finnish)
(I) speak.PL English
‘I speak English.’ (from Holmberg, 2005, 539-540)

- (10) a. *Leikkii lapsia kadulla. (Finnish)
play children in-street
‘There are children playing in the street.’
b. Sitä Leikkii lapsia kadulla.
EXPL play children in-street
‘There are children playing in the street.’
c. Kadulla Leikkii lapsia.
in-street play children
‘There are children playing in the street.’
d. Lapsia Leikkii kadulla.
children play in-street
‘There are children playing in the street.’ (from Holmberg and Nikkane, 2002, 71-72)

d) Overt thematic subjects/overt expletive subjects. This last possibility is illustrated by English or French, where both types of clauses require overt subjects:

- (11) a. Elle arrive demain. (French)
she arrives tomorrow
‘She’s arriving tomorrow.’
b. *Arrive demain.
arrives tomorrow
‘She/He’s arriving tomorrow.’

- (12) a. Il pleut beaucoup ici. (French)
 EXPL rains a lot here
 ‘It rains a lot here.’
 b. *Pleut beaucoup ici.
 rains a lot here
 ‘It rains a lot here.’

Thus, in principle all combinations of null/overt and thematic/expletive are possible. However, there are further distinctions in the typology of both expletive and thematic subjects, which I will explore in turn.

2.1.1.1 Typology of null expletive subjects

The paradigm given in the previous section is more complicated than I have presented it. Svenonius (2002, 5-7) points out that several authors have proposed that some expletives may be referential, so that the full taxonomy can be divided into two major classes: Those that have some kind of referential capability, and those that are purely expletive. Among the first, he quotes weather-verb expletives, extraposition expletives, and among the second, presentational/impersonal expletives, illustrated in (13)-(16).

- (13) Aquí pro siempre llueve. Quasi-referential expl. (Spanish)
 here pro always rains
 ‘Here it always rains.’
- (14) It surprised me [that Santos won] Extraposition expl. (English)
- (15) pro parece que sabe la verdad. Raising expl. (Spanish)
 pro seems that knows the truth
 ‘It seems that s/he knows the truth.’
- (16) Hay unos visitantes en la puerta. Existential expl. (Spanish)
 are some visitors in the door
 ‘There are some visitors at the door.’

Chomsky (1981, 323-5) argues that weather-verb expletives are quasi-argumental because they can control an infinitival subject, as illustrated in (17). Expletive *there*, on the other hand, cannot control, as seen in (18a). Notice, however, that this contrast is independent of the construction, since control is possible with *this* in a similar context as in (18a) (cf. (18b)). The contrast between (18a-b) suggests that the referential properties of the expletive determine control, not necessarily the type of expletive construction.

- (17) a. It often clears up here right after snowing heavily. (English)
 (from Svenonius, 2002, 6, 5a)

- b. Korai szürkületet okozva befelhösödött. (Hungarian)
 early twilight.ACC causing clouded
 ‘It clouded, causing early twilight.’ (from Kiss, 2002, 119, ex. 39a)
- (18) a. *There is often a party here right before being a wake. (English)
 (from Svenonius, 2002, 6, ex. 5b)
- b. This is often a gathering place before being a hiring place.

It is not clear to what extent one can conclude from the examples in (17) that a theta role is assigned to the expletive, in particular because it is not clear what that theta-role would be. Although there is some intuitive connection between the events of snowing and the event of clearing up in (17a), it is less clear that the subject of *clearing up* controls (i.e. has the same reference as) the subject of *snowing heavily*. The same objection holds for the relationship between *clouding* and *causing early twilight* in (17b).

Another dimension of variation relates to the availability of different types of null expletives within the same language, as pointed out by Nicolis (2005, 2008) and Sheehan (2007, ch. 5). As a particularly interesting example, Sheehan (2007, 240) presents the table below (from Nicolis, 2005, adapted for current purposes), where different Romance-based creoles show different patterns of null expletives.

- (19) Null expletives in Romance-based Creoles.

	Quasi-referential	Extraposition	Raising	Existential
Kriyol	overt	overt?	optional	overt
Haitian Creole	optional	overt	null	null
Mauritian Creole	DP	optional?	overt	overt
Cape Verdean	null	null	null	null
Papiamentu	null	null	null	null

Finnish is somewhat similar to Haitian Creole, since expletives are optional with weather verbs and extraposed subjects and possibly existential ones as well. Specifically, Holmberg and Nikkanen (2002) note that next to (10a) above, where the expletive is obligatory if nothing else appears in first position, we find examples like (20), where the verb is clause-initial and no expletive is required.

- (20) a. On ilmennyt ongelmia. (Finnish)
 have appeared problems
 ‘Problems have appeared’
- b. Sattui onnettomuus.
 occurred (an) accident
 ‘An accident occurred.’
- c. Tuli kiire.
 came haste
 ‘We/they are in a hurry.’

- d. Oli hauskaa että tulit käymään.
 was nice that came-2.SG visiting
 ‘It was nice that you came by.’ (from Holmberg and Nikkane, 2002,
 ex.19)

Holmberg and Nikkane (2002) account for this difference by making the EPP feature optional, so that the structures in (20) do not require overt movement of any constituent to clause-initial position. This movement, which is required (but does happen) in (10a) above, licenses [-focus] constituents. In their analysis, Finnish is a topic-oriented language, which triggers the need to have an overt constituent preceding the verb in most circumstances.

German presents a slightly more complicated picture: Expletives are required with meteorological verbs but not with existentials (cf. Sheehan, 2007, 239-240).

- (21) a. Getern regnete es. (German)
 yesterday rained EXPL
 ‘Yesterday it rained.’
 b. *Getern regnete.
 yesterday rained
 ‘Yesterday it rained.’ (from Sheehan, 2007, 239-240)

- (22) a. Es steht ein Mann vor der Tür. (German)
 there stood a man in-front-of the tower
 ‘There stood a man in front of the tower.’
 b. Vor der Tür steht (*es) ein Mann.
 in-front-of the tower stood EXPL a man
 ‘There stood a man in front of the tower.’ (from Sheehan, 2007, 239-240)

From these data, we can conclude that the availability of null argumental subjects is neither a necessary nor a sufficient condition for null expletives: Finnish has null argumental subjects in certain circumstances, but overt expletives are obligatory in certain cases.

2.1.2 Free Subject Inversion

Languages like Spanish or Italian have a default, non-marked word order SVO, however, they can also display an alternative VS word order in declarative sentences, as illustrated in (23).

- (23) a. Habló Marta. (Spanish)
 spoke Marta
 ‘Marta spoke.’
 b. Mangia Gianni. (Italian)
 eats Gianni
 ‘Gianni eats.’

- (27) a. *Who do you think that will meet Jim?
 b. Who do you think that he will meet?
- (28) a. Who do you think will meet Jim?
 b. Who do you think he will meet?

- (29) *That-trace Filter*
 * $[S_i \text{ that } [NP^e] \dots]$,
 unless S_i (or its trace) is in the context $[NP _ NP \dots]$

In NSLs, subject extraction is possible even with a full complementizer, as discovered by Perlmutter (1971). Thus, (30) is grammatical, and in fact, the complementizer cannot be deleted.

- (30) a. ¿Quién piensa-s que vendrá?
 who think.2.SG that will come
 ‘Who do you think that will come?’
 b. *¿Quién piensa-s vendrá?
 who think.2.SG will come
 ‘Who do you think that will come?’

The account for the contrast between English and Italian/Spanish regarding *that*-trace effects was later subsumed under the Empty Category Principle. In particular, NLSs allow long-distance subject extraction because the trace of the subject is properly governed. The properties of agreement in NSLs allow for proper government of subject traces in Italian/Spanish, but not in English (cf. Jaeggli, 1980; Rizzi, 1982 and Chomsky, 1981). One instantiation of this intuition takes subjects in these languages to be generated in postverbal position. For example, Jaeggli (1980, 261) argues that the grammaticality of (29a) follows from the fact that Spanish has postverbal subjects, hence extraction takes place from the postverbal subject position.

In favor of this hypothesis, Rizzi provides evidence from the Northern Italian dialects Trentino and Fiorentino, where it is possible to show that subject *wh*-movement takes place from a postverbal position (cf. Rizzi, 1986a; Brandi and Cordin, 1989; Giupponi, 1988; Poletto, 1993, quoted in Barbosa, 1995, 31-33). In Fiorentino subject clitics obligatorily agree with a preverbal DP (cf. (31a-b)), whereas postverbal subjects take a default, 3.SG.MASC clitic (cf. (31c)).

- (31) a. La Maria l' è venit-a. (Fiorentino)
 the Maria 3.SG.FEM is come-FEM
 ‘Maria has come.’ (from Barbosa (1995, ex. 53a))
- b. *La Maria gli a telefonà.
 the Maria 3.SG has telephoned.
 ‘Mary has telephoned.’ (from Barbosa (1995, ex. 53b))
- c. Gl' è venit-o la Maria.
 3.SG is.3.SG come.MASC the Maria
 ‘Maria has come.’ (from Brandi and Cordin (1989, ex. 14))

This agreement pattern provides a marker for subject position, and within this context, extracted subjects obligatorily have neutral agreement, suggesting that the subject originates postverbally, as shown in (32). In addition, Brandi and Cordin (1989, 125) show that the same holds for subject extraction from embedded clauses:

- (32) a. *Quante ragazze le sono venute? (Fiorentino)
 how many girls 3.PL.FEM are come (from Brandi & Cordin (1989 ex. 39a))
 b. Quante ragazze gli è venuto con te?
 how many girls 3.SG.MASC is come with you
 ‘How many girls came with you?’ (from Brandi & Cordin (1989 ex. 37a))

Note, however, that with the generalization of the VP-internal subject hypothesis (cf. Koopman and Sportiche, 1991), and the adoption of the Copy theory of movement (cf. Chomsky (1993) in its most recent incarnation), the explanation for these asymmetries based on the availability of a postverbal subject position for NSLs can no longer be maintained: Both English and Italian/Spanish are now assumed to have a VP-internal (i.e. postverbal) position, and that position will have deleted copies of the overt subject in preverbal position in both types of languages, as depicted in (33).

- (33) a. English: [_{IP} DP [_I [_{vP} $\overline{\text{DP}}$ V]]]
 b. Italian: [_{IP} DP [_I [_{vP} $\overline{\text{DP}}$ V]]]

Thus, under current minimalist assumptions, the account of the subject-object long-distance wh-extraction differences between English and Spanish/Italian cannot be based on the availability of the postverbal subject position in Spanish/Italian (cf. Pesetsky and Torrego, 2001 for an alternative analysis of *that*-trace effects).⁵

A separate but important question is the following: In its original formulation, the *that*-trace filter eliminated configurations with an overt complementizer in subject extractions, but as Barbosa et al. (2005, fn. 7) point out, what seems to be more relevant is “the fact that subjects are extracted from post-verbal position”. This conception raises a few practical issues: How do we know what counts as a *that*-trace filter violation in any given language? Second, how do we know whether extraction has happened from a postverbal position? Is it just a matter of free V-S inversion? I will return to these issues in chapter 3.

2.1.5 Long wh-extraction

Long wh-extraction entails movement of a wh-word across a clausal boundary in contexts where another wh-operator occupies potential intermediate landing sites. In the relative clause example in (34a), the relative pronoun has been extracted from the subject position of *abbia visto* ‘has seen’

⁵Szczegielniak (1999) attempts to derive *that*-trace effects within a Minimalist theory of phases, and specifically argues that the connection between lack of *that*-trace effects and NSs arises because NSLs do not project a Spec, TP, hence the A-subject remains in situ, whereas English/French does project Spec, TP (and the subject must move obligatorily, one must assume). It also follows that in NSLs, the preverbal subject is not in Spec, TP. This approach is highly compatible with the pronominal agreement analysis that I will present in 4.3.

over the wh-word *chi* ‘who(m),’ as schematized in (34b). In (35a), from Suñer (1994, 354), the subject wh-word has been extracted over the object wh-word.⁶

- (34) a. L’uomo che mi domando *chi* abbia visto. (Italian)
 the-man that CL wonder who has seen
 ‘The man who I wonder who he has seen.’ (from Chomsky, 1981, 240)
- b. L’uomo che mi domando [*chi* t abbia visto]
 ↑-----↓
- (35) a. ¿Quién₂ dijiste (tú) a quiénes₃ (no) les había e₂ prestado un montón
 who said you to whom not CL had lent a lot
 de dinero e₃? (Spanish)
 of money
 ‘Who did you say to whom had (not) lent a lot of money?’
- b. ¿Quién dijiste tú [a quiénes no les había e prestado un montón de dinero?]
 ↑-----↓

2.1.6 Section summary

To summarize, out of the six initially proposed core properties of the NSP (null subjects, null expletives, free subject inversion, null resumptive pronouns, long wh-extraction and absence of *that*-trace effects) at least four were tied to the availability of a postverbal subject position, which, in turn, was connected with the existence of *pro* in preverbal position. NSs were possible because of the rich-agreement properties of the NSLs. The existence of null expletives was also indirectly connected: Subjects can be postverbal in these languages because null expletives satisfy the EPP, although if the EPP were not a universal principle, the reason to postulate null expletives would disappear, as we will see.

We have also seen potential objections to some of these analysis, in particular given changes in theoretical assumptions, the availability of the postverbal subject position only in NS languages becomes problematic.

2.2 Interpretive differences between null and overt subjects

One essential asymmetry between NSLs and overt-subject languages is that NSLs may potentially contrast null vs. overt pronouns in subject position. This contrast can be used to convey several interpretive differences. Put in a different way, null subjects are not simply the null counterpart of overt pronouns (but see 4.4). Thus, several interpretive distinctions have been argued to correlate with the null/overt distinction.

⁶I find ungrammatical both (35a) and the Spanish translation of (34a) in (i).

(i.) *El hombre que me pregunto a quién haya visto/vio. (Spanish)
 the man that CL ask to whom had seen/saw
 ‘The man who I wonder whom (he) saw.’

2.2.1 Arbitrary and generic interpretations

In Spanish and in many other languages, overt pronominals cannot be interpreted as having arbitrary reference (cf. Suñer, 1983; Jaeggli, 1986). This is illustrated in (36). Thus, (36b) cannot be interpreted as ‘someone or other said that they had come’, but (36a) can.

- (36) a. Dijeron que habían venido. (Spanish)
 said that had come
 ‘They (specific or arbitrary) said that they had come’
 b. Ellos dijeron que habían venido.
 they said that had come
 ‘They (specific only) said that they had come.’

In a related matter, Holmberg and Sheehan (2010, 129) point out that languages like Brazilian Portuguese, which has arguably lost null subjects, contrasts with European Portuguese (and Spanish as well) in the fact that (37) can be interpreted as generic, whereas in EP/Spanish, it must be interpreted referentially. In order to obtain a generic interpretation for this type of example in these NS varieties, one must add a generic clitic *se* must be added, as (37b-c) illustrates.

- (37) a. É assim que faz a doce. (Brazilian Portuguese)
 is thus that makes the sweet
 ‘This is how one makes the dessert.’
 b. É assim que se faz a doce. (European Portuguese)
 is thus that CL makes the sweet
 ‘This is how one makes the dessert.’
 c. Es así que se hace el dulce. (Spanish)
 is thus that CL makes the sweet
 ‘This is how one makes the dessert.’

Holmberg and Sheehan (2010, 129) point out that the importance of distinguishing between arbitrary and generic null subject pronouns. Arbitrary ones denote people in general, but exclude the speaker and the addressee, like *they* does in *they like spicy food in Peru*, whereas generic ones denote people in general including the speaker and the addressee, much like English *one*. Consistent NSLs have an arbitrary null subject, but resort to a distinct strategy to express a generic statement. For example, in many varieties of Spanish, the 2nd person is productively used in generic contexts, although the subject can be overt or null, as in (38). This speaker is comparing the impossibility of making copies of videogames with the impossibility of copying a movie shown in a movie-theater (as opposed to copying a book for personal use). In this example, the use of the 2nd person does not imply participation of the addressee, but is rather intended as a generic statement (cf. Hernanz, 1990).

- (38) Es como en el cine ... tu puedes mirar la película por la que has pagado pero
 is like in the movies you can watch the movie for the that have paid but
 no te la puedes llevar. [sic] (Spanish)
 not CL CL can take
 ‘It’s like the movies, you can watch a movie you have paid to see, but you can’t take it.’

- b. *Claro, y *pro*_i ha convencido a su padre de que compre un carro.
of course and has convinced to his father of that buy a car
‘Of course, and that has convinced his father to buy him a car.’
- c. Claro, y eso_i ha convencido a su padre de que compre un carro.
of course and that has convinced to his father of that buy a car
‘Of course, and that has convinced his father to buy him a car.’

Iatridou and Embick (1997) point out that in certain cases a null pronominal expletive can “refer” to a clause, as seen in (43). These null categories are different from the preceding ones: They are arguably expletives because the clause is extraposed, whereas in (41) the relevant clause is not extraposed (cf. Iatridou and Embick, 1997, 61). The same effect can be seen in Spanish (cf. (44)).

- (43) a. An o Kostas argisi *pro* tha ine dropi. (Greek)
if the Kostas is late FUT be shame
‘If Kostas arrives late, it will be a shame.’
- b. *pro* ine dropi pu o Kostas tha figi.
be shame that the Kostas FUT leave
‘It is a shame that Kostas will leave.’ (from Iatridou & Embick (1997, ex. 6))
- (44) a. Si Gabriel se retrasa, *pro* será una lástima. (Spanish)
if Gabriel CL is late, will be a shame
‘If Gabriel is late, it will be a shame.’
- b. Es una lástima que Gabriel se retrase.
is a shame that Gabriel CL is late
‘It is a shame that Gabriel is late.’

Iatridou and Embick (1997) argue that the phenomenon observed in (41) is syntactic in nature, that is, *pro* cannot be coindexed with the category IP/CP due to a ϕ -feature mismatch: IP/CP lacks number, gender and possibly person, whereas thematic *pro* requires all three of them.

2.2.3 Quantifier Binding

The second interpretive difference was first formalized by Montalbetti (1984, 95) as the **Overt Pronoun Constraint**, in (45). This constraint suggests that whenever there is a contrast between overt and null pronouns in a language, the overt version cannot be bound as a variable. The effects of this constraint can be seen in (46).

- (45) **Overt Pronoun Constraint (OPC)**
Overt pronouns cannot link to formal variables iff the alternation overt/empty obtains.
- (46) a. Todo estudiante_i cree que *pro*_i es inteligente. (Spanish)
every student thinks that is intelligent
‘Every student_i thinks that he_i is intelligent’

- b. Todo estudiante_i cree que él_{*i/j} es inteligente.
 every student thinks that he is intelligent
 ‘Every student_i thinks that he_j is intelligent.’

The OPC has been argued to exist in a number of unrelated languages, among them Chinese, Japanese and Korean (cf. Kanno, 1997; Yamada, 2005), Tarifit (Ouhalla, 1988). However, in some work on the acquisition of Spanish as an L2, ? has not found a clear pattern for the native control group. To further confound the situation if the overt pronominal is focused the bound reading becomes possible (cf. Sánchez, 1994; Lozano, 2002).

2.3 Typology of NSLs

A primary typological distinction among NSLs is whether the verbal paradigm displays overt person/number morphology or not. Thus, languages like Spanish, Italian or Quechua encode grammatical person and number in most of the verbal paradigm (cf. (47)), whereas others like Japanese have morphemes that do not encode person/number (cf. (48)), and others like Chinese have no verbal morphemes whatsoever (cf. (49)).

(47) Quechua morphological paradigm

a. Tiya-ni	live-1.SG	‘I live’
b. Tiya-nki	live-2.SG	‘you live’
c. Pay tiya-n	live-3.SG	‘he/she lives’
d. Tiya-nchik	live-1.PL.INCL	‘we (inclusive) live’
e. Tiya-yku	live-1.PL.EXCL	‘we (exclusive) live’
f. Tiya-nkichik	live-2.PL	‘You(pl) live’
g. Tiya-nku	live-3.PL	‘They live’

(from <http://facultad.pucp.edu.pe/ciencias-sociales/curso/quechua/gramatica.html#stop3>)

(48) Japanese morphological paradigm

a. Yom-ru	read-PRES	‘read’
b. Yom-ta	read-PAST	‘read’
c. Yom-anai	read-NEG	‘not read’
d. Yom-eba	read-CONDIT	‘not read’
e. Yom-oo		‘let’s read’
f. Yom-itai		‘want to read’
g. Yom-are	read-PASS	‘was read’
h. Yom-ase	read-CAUS	‘make read’

(from Jaeggli and Safir, 1989, 29)

- (49) Chinese verbal paradigm
 Shi 'be (all persons)'

2.3.1 Consistent null subject languages

The Quechua/Italian/Spanish-type of NSL shows fairly systematic a use of null subjects. In some sense, these are the default, whereas overt pronouns are typically used to indicate change of topic or contrast. For this reason, they have been called consistent NSLs. The traditional intuition has been that null subjects are possible due to rich agreement (cf. Taraldsen, 1980; Rizzi, 1986a and much subsequent work, as well as sect. 4.3 below). This hypothesis works well for some rich-agreement languages, but formalizing the notion of rich agreement in a precise, predictive and accurate way has proven to be very challenging. The main challenge involves determining the necessary and sufficient conditions for rich agreement. Jaeggli and Safir (1989) make such an attempt, defining the notion of a **morphologically uniform** paradigm in (50). Thus, they argue, morphological verbal paradigms can be divided into three types across languages: Those whose forms are all morphologically derived (the Quechua type in (47) above and the Japanese type in (48)), those that are wholly underived, such as Chinese, and those that are mixed, like English and French.

(50) **Morphological Uniformity**

An inflectional paradigm P in a language L is morphologically uniform iff P has either only underived inflectional forms or only derived inflectional forms. (from Jaeggli and Safir (1989: 30))

Based on this partition between derived and non-derived paradigms, they argue that only morphologically uniform languages can have null subjects (cf. (51)). This predicts that Quechua, Spanish, Italian, Chinese and Japanese will have null subjects, whereas English and French will not.⁸

(51) **The Null Subject Parameter**

Null subjects are permitted in all and only languages with morphologically uniform inflectional paradigms.

2.3.2 Discourse-related null subject languages

For the Chinese/Japanese type, Huang (1984) proposed that null subjects are licensed through topics (hence their categorization as discourse-related null subject languages). Thus, in (52), the embedded subject can be coreferential with the matrix subject *Zhangsan* or with "some other person whose reference is understood in discourse (i.e. the discourse topic)" (Huang, 1989, 187-188).

⁸In the case of German and Icelandic, which are morphologically uniform (and arguably have agreement-rich), but do not have null thematic subjects, they argue that the main difference between them and Italian/Spanish stems from the fact that German/Icelandic, being V2 languages, have T and AGR in separate nodes, whereas Italian/Spanish do not. Jaeggli and Safir assume that AGR must Case-govern the empty category to identify it. Since German/Icelandic have T and AGR split, and T is in C, AGR does not case-govern the null category, hence it is not identified.

- (52) Zhangsan shuo [*pro* hen xihuan Lisi]. (Chinese)
 Zhangsan say very like Lisi
 ‘Zhangsan said that he liked Lisi.’ (from Huang, 1989, 187, ex. 4a)

More recently, Tomioka (2003) has proposed linking the availability of null subjects in discourse NSLs with properties of nouns in those languages. Tomioka notes that Japanese null pronouns have a wide range of semantic interpretations (referential uses, bound variable uses, E-type pronoun uses, indefinite uses, property anaphora, etc.). He correlates this range of meanings with another property of Japanese, namely the wide availability of bare NPs with many semantic interpretations in that language. Thus, for example, (53) shows a bare NP that can be interpreted as an indefinite or definite singular or plural.

- (53) Ken-wa ronbun-o yon-da (Japanese)
 Ken-TOP paper-ACC read-PAST
 ‘Ken read a paper/papers/the paper/the papers.’ (from Tomioka, 2003, 328, ex. 19)

Tomioka proposes deriving the diverse meanings of *pro* through the same mechanisms proposed for the meanings of NPs. In particular, he argues that *pro* in languages like Japanese is the phonologically null version of bare NPs. If this view is correct, it can be extended to other discourse NSLs, yielding the generalization in (54). This generalization applies to Japanese, Korean, Mandarin Chinese, Thai, Hindi, Turkish, and Brazilian Portuguese, among others.

(54) **Discourse Pro-drop Generalization**

All languages which allow discourse pro-drop allow (robust) bare NP arguments.
 (from Tomioka, 2003, 336)

Under this view, discourse NSs are the by-product of *N'*-Deletion/NP ellipsis, an operation that applies crosslinguistically, but whose results depend on the independent structure of nominals. Thus, in English, *N'*-Deletion/NP ellipsis leaves an overt D, as in (55a) in Japanese it does not, as in (55b).

- (55) a) English-type NP-ellipsis (DP-language) b) Japanese-type NP-ellipsis (NP-language)



As Tomioka points out, this approach opens up several questions, among them, why languages with robust bare NPs (like English) do not allow NSs.

Another question raised by the analysis relates to partial NSLs like Shipibo (see immediately below), which have robust bare NPs in Tomioka’s sense, but allow NSs only in 3rd person, not in 1st or 2nd.

In Shipibo, the partition is the opposite, 1st/2nd person are obligatorily overt, whereas 3rd is optionally null (cf. Camacho and Elías-Ulloa, 2010).

- (59) a. Lima-n-ra noko-ke. (Shipibo)
 Lima-DIR-EVID arrive-PERF
 ‘He/she went to Lima.’
 b. Ea-ra Lima-n noko-ke.
 1-EVID Lima-DIR arrive-PERF
 ‘I arrived in Lima.’
 c. Mia-ra Lima-n noko-ke.
 2-EVID Lima-DIR arrive-PERF
 ‘You arrived in Lima.’ (from Camacho and Elías-Ulloa, 2010, 72, ex. 13)

A third type of person restriction is illustrated by Marathi, which only allows NSs in 2nd person, according to Holmberg et al., 76. Interestingly, verbal morphology only distinguishes the 2nd person singular unambiguously.

Another source of variation relates to whether the null subject is in main or embedded clauses. In Shipibo, for example, 1st and 2nd subjects can be null so long as their antecedent is in the immediate clause (cf. Camacho and Elías-Ulloa, 2010). In Brazilian Portuguese, Finnish and Marathi, embedded subjects can be null if controlled (cf. Vainikka and Levy, 1999; Holmberg et al.; Modesto, 2008). Finally, indefinite subjects can be null in Brazilian Portuguese, Finnish and Marathi (cf. Holmberg et al.). Both of these properties are illustrated in (60). In the first example, the null subject is interpreted as indefinite, whereas in the second example, it is controlled by the main-clause subject.

- (60) a. Unahlyat lavkar utthavla jato. (Marathi)
 summer-in early wake go-PRES-3.SG.MAS
 ‘In summer one wakes up early.’
 b. Ram mhanala ki ghar ghetla.
 Ram say-PAST.3.SG.MAS that house buy-PAST.3.SG.NEUT
 ‘Ram said that he bought a house.’ (from Holmberg et al., 60, ex. 1)

Barbosa (2010) extends Tomioka’s (2003) proposal for discourse NLSs to partial *pro*-drop languages. In particular, she argues that partial NSLs have properties in common with discourse NSLs that set them apart from consistent NSLs: In partial NSLs, the null subject is optional in certain contexts where it is obligatory in consistent NSLs (i.e. no Avoid Pronoun Principle effects). This is illustrated by the contrast between European and Brazilian Portuguese.

- (61) a. O João_i disse que ele_i comprou um computador. (#EP) (√BP)
 the Joao says that he bought a computer
 ‘Joao says that he bought a computer.’
 b. O João_i disse que *pro*_k comprou um computador. (√EP) (√BP)
 the Joao says that he bought a computer
 ‘Joao says that he bought a computer.’ (from Barbosa, 2010, ex. 3)

Additionally, an intervening antecedent blocks a NS in certain contexts in partial NSLs but not in consistent NSLs:

- (62) a. O João_i disse que os moleques acaham [que *pro*_k é esperto] (✓/EP) (*BP)
 the Joao says that the children believe that is smart.
 ‘Joao says that the children believe that he is smart.’ (from Barbosa, 2010, ex. 8)
- b. *Talila_i ’amra le-Itamarj še *pro*_k tavo. (Hebrew)
 Talila said to-Itamar that will-come.FEM.SG
 ‘Talila said to Itamar that she will come.’ (from Holmberg, 2005, 554, ex. 41b)

A third difference relates to the possibility of interpreting a 3rd person null subject as an indefinite in partial NSLs, a possibility that consistent NS do not have. Instead, these languages resort to other strategies, for example having a clitic.

- (63) a. É assim que faz o doce. (BP)
 is so that make.3.SG the cake
 ‘That’s how one makes the cake.’
- b. É assim que se faz o doce. (EP)
 is so that CL make.3.SG the cake
 ‘That’s how one makes the cake.’ (from Barbosa, 2010, ex. 18-19)

Based on these patterns, she argues that in consistent NSLs, agreement is pronominal, and preverbal subjects are CLLD phrases (see 4.3.1 and 8.1). In partial NSLs and discourse NSLs, agreement is not pronominal and null subjects are NP anaphora.

2.3.4 Section summary

In this section, I have presented a typological distinction between consistent, agreement-based NSLs, discourse-based NSLs and partial NSLs. The latter type involves variation across several domains: NSs allowed depending on person and tense, or depending on whether they appear in main or embedded clauses. In this last respect, partial NSLs seem to have properties more consistent with discourse-based NSLs, where the NS is recovered through discourse. Finally, partial NSLs show interpretative effects, for example lack of interpretive contrasts between overt and null pronouns.

2.4 Chapter summary

In this chapter, I have presented the cluster of properties proposed for the NSP, namely null subjects, null expletives, long wh-extraction, *that*-trace effect violations and alternative VS word order, as well as the interpretive properties: OCP effects, absence of arbitrary readings with overt pronouns (in NSLs), the need to have additional strategies for generic readings with *pro* in consistent NSLs, and the impossibility of referring back to CPs through *pro* in languages like Spanish/Greek.

Finally, I have presented two typologies of NSLs: First, the distinction between consistent NSLs (those that have rich agreement and where NSs appear without systematic restrictions for person, number, etc.), discourse-related NSLs (those with no agreement where NSs are identified in discourse) and partial NSLs (those that have NSs in limited contexts, depending on person/tense, etc.). Among partial NSLs, we find those whose NSs correlate with explicit inflection (Hebrew, Finnish), and those that do not (Shipibo).

The second typology relates to whether the NS are restricted to a main or embedded clause. Thus, Brazilian Portuguese, Finnish and Marathi restrict NSs to controlled instances in embedded clauses.

Chapter 3

The core content of the NSP

3.1 Introduction

As observed in the previous chapter, the notion of a parameters as articulated particularly in the Principles and Parameters framework, makes the claim that several grammatical properties cluster together around one core property. This conception facilitates language acquisition (cf. Hyams, 1986, 1992; Hyams and Wexler, 1993 and Licerias (1989) among others) and makes fairly strict predictions about sequence of acquisition, possible and impossible grammars, etc.

In the case of the NSP, there have been several proposals regarding the core property of the NSP (cf. Gilligan (1987, 78-96) for a summary). Taraldsen (1980), for example, proposes that the subject position is potentially bound by AGR, which gives rise to the possibility of null non-thematic subjects. That availability of binding by AGR underlies *that*-trace filter violations: Since the subject position is bound in NSLs, extraction is licensed.

Rizzi (1982) modifies this notion, arguing that rich AGR binds the post-verbal subject position, so the availability of *that*-trace effect violations in NSLs is related to the possibility of extracting from that position. Long *wh*-extraction and null resumptive pronouns also follow from this basic asymmetry: in languages like Italian, the postverbal subject position is licensed (technically, governed) in a way that the preverbal position is not in English. The availability of postverbal positions, in turn, is overtly manifested in free inversion (cf. 2.1.2 and 2.1.4 above). Gilligan (1987) tests the typological predictions of Rizzi's early account, and reaches certain implicational correlations between having null subjects, free inversion and *that*-trace violations, which Roberts and Holmberg (2010) further refine.

In this section we will review the different proposals and contrast them against empirical data from a number of languages, particularly Brazilian Portuguese, Dominican Spanish and Lubukusu and other Bantu languages.

3.2 Property correlations

Gilligan's (1987) study surveys the property cluster associated with the NSP in over 100 languages

to establish whether the combinations predicted in (1) are the only attested ones.¹

- (1) Gilligan's (1987) typology of NSLs based on Taraldsen (1980), Rizzi (1982) and Safir (1985).

Null thematic subjects	Null nonthematic subjects	Subject inversion	<i>That</i> -trace filter violations
+	+	+	+
-	+	+	+
-	-	+	+
-	-	-	-

Although many of the studies he relies on are missing data on one or more of the properties, Gilligan (1987, 147) tentatively concludes that his survey supports the implications in (2).

- (2)
- a. If a language has a thematic *pro*, it has expletive null subjects.
 - b. If a language has free inversion, it has expletive null subjects (statistically significant correlation).
 - c. If a language has free inversion, it has *that*-trace effects violations.
 - d. If a language has *that*-trace effects violations, it has null expletive subjects.

These implicational statements predict certain possible and impossible combinations. Specifically, the combinations in (3) should not be attested.

- (3) Combinations predicted not to exist in Gilligan's (1987) typology.

	Null thematic subjects	Null nonthematic subjects	Subject inversion	<i>That</i> -trace filter violations
a)	Yes	No		
b)		No	Yes	
c)			Yes	No
d)		No		Yes

Based on these results, Roberts and Holmberg (2010, 22) propose the implicational scale in (4), which defines the three types of languages represented in (5).

¹(1) corresponds to Safir's analysis, which establishes less direct correlations between the different phenomena, in particularly taking into account languages where *that*-trace effect violations exist without null subjects, such as Dutch. Gilligan tests two additional patterns, both more restrictive, corresponding to Taraldsen (1980) and Rizzi (1982). For illustrative purposes, (1) should suffice.

- (4) Subject inversion → (allow *that*-trace violations → expletive null subjects)
- (5) I: Has all three properties (Italian, Spanish, Greek, etc.)
 II: Allows *that*-trace violations and expletive null subjects (Cape Verdean, Berbice Dutch, Mauritian, Papiamentu, Saramaccan)
 III: Only allows expletive null subjects (Haitian, basilectal Jamaican)

It is difficult to assess the empirical validity of both Gilligan's predicted impossible combinations and of Holmberg & Roberts' implicational scale, in part because in the absence of a specific analysis, it is not clear what constitutes a *that*-trace filter violation or free inversion (cf. Roberts and Holmberg, 2010 on the limitations of typological predications like these). Be that as it may, in the following sections, I will review some of the existing evidence.

3.2.1 Null thematic subjects & no null expletives

The first correlation states implies that one should not have languages like Spanish but with overt expletives. The issue that arises is whether languages or varieties that have partial thematic NSs and overt (but not obligatory) expletive subjects count as potential counterexamples. As already discussed in 2.1.1, Finnish has null subjects in 1st and 2nd person, as seen in (6), and it also has what seems to be optional expletives, as seen in the contrast between (7a) and b). In the first example, the expletive is obligatory, but in the second one, it is not required, even though the verb appears in 1st position. Because *kiire* 'haste' cannot satisfy the EPP, it need not (and cannot) raise, hence the EPP is suspended.

- (6) (Minä) puhun englantia. (Finnish)
 (I) speak.PL English
 'I speak English.'
- (7) a. *(Sitä) Leikkii lapsia kadulla.
 EXPL play children in-street
 'There are children playing in the street.'
- b. Tuli kiire. "We/they are in a hurry."
 came haste

Like Finnish, certain varieties of Dominican Spanish have NSs and also an overt expletive, which is nevertheless not obligatory, as we will see in the next section. By comparing this variety with other Romance and Spanish varieties, we will establish at least two types of expletives, one that competes with overt preverbal subjects, and one that is compatible with them. This distinction suggests that there are at least two positions for expletives, a CP-related one, and an IP-related one.

3.2.1.1 Overt expletives in Dominican Spanish

Dominican Spanish (henceforth DS) has a high proportion of overt subjects, particularly when compared with other varieties of Spanish. Thus, for example, in Otheguy et al.'s (2007) study of speakers of different Spanish varieties recently arrived in New York, Dominicans had 41% of overt pronouns vs. 27% for Ecuadorians, 24% for Colombians and 19% for Mexicans. In a separate study conducted in the Dominican Republic, Cabrera (2007) found a preference for overt subjects over null ones in 70% of sentences tested. In the following example, we find a sequence of several overt pronominals even though the topic does not change.²

- (8) Hay unas muchachitas que están juntas conmigo que **ellas** viven pa' fuera, entonces **ellas** vinieron a estudiar en la escuela del Pino, entonces **ellas** saben mucho inglés . . . yo no me acuerdo en el país que **ellas** vivían. (DSEC)
 'There are some girls that are together with me, that they live outside, so they came to study in the Pino School, so they know a lot of English . . . I don't remember the country in which they lived.'
 (from Bullock and Toribio, 2009, 56, ex. 12b)

However, overt subjects are not obligatory, as the following example shows. In this sequence, all subjects that maintain the topic are null, even in the case of *hablaron* 'they spoke', which retrieves two separate topics (the referent of the first two subjects and his friend).

- (9) Se fue, cogió un bastón, se encontró con su compadre y hablaron tanto que le
 CL went, got a stick, CL met with his friend and talked so much that CL
 salió raíz al ba'tón. (DSEC)
 came out root to-the stick
 'He left, got a (walking) stick, he ran into a friend, and they talked so much that the stick sprung roots.'
 (from Bullock and Toribio, 2009, 63, ex. 26b)

As a conclusion, while DS has an increasing use of subjects, it is by no means an obligatorily-overt subject language. In the variety spoken in El Cibao (henceforth DSEC), there is an optional expletive, illustrated in (10) (cf. Henríquez Ureña, 1939, 1940/1975; Toribio, 1993; Silva-Villar, 1998; Hinzelin and Kaiser, 2006, 2007; Bullock and Toribio, 2009, among others). This expletive covers at least the range of *there* and *it* in English, so it can appear with existential/presentational verbs (cf. (10a, d)), weather verbs (cf. (10b)), unaccusative verbs (cf. (10c, e)) and transitive verbs (cf. (10f)).³

- (10) a. Ello hay personas que lo aprenden bien (el inglés) (DSEC)
 EXP is people.PL that CL learn.PL well (the English)
 'There is people who learn it well (English).'

²The examples below are from the Dominican variety spoken in El Cibao, which will be relevant below. Whenever an example is explicitly from El Cibao, it will be marked as DSEC, otherwise, Dominican examples will appear as DS.

³DSEC presents a particular methodological challenge, because many of its features have a highly negative social connotation, and there is strong social pressure to eliminate them, particularly among educated, urban speakers. Since many studies rely on data from urban speakers, conclusions must be tentative. One particularly valuable study that relies on rural and/or uneducated speakers from El Cibao is reported in Bullock and Toribio.

- b. Ello no está lloviendo aquí pero allá sí.
 EXP not is raining here but there yes
 ‘It’s not raining here, but there, it is.’
- c. Ello vienen haitianos aquí.
 EXP come.PL Haitians here
 ‘There come Haitians here.’
- d. Ello queda mucho tiempo todavía.
 EXP remains much time still
 ‘There still remains much time.’
- e. Porque si ello llega una gente de pa’ fuera ...
 because if EXP arrives some people.SG of outside
 ‘Because if there arrive people from the outside ...’
- f. Ello se hace bollos con coco.
 EXP CL make cakes with coconut
 ‘One makes cake with coconut.’
- g. Ello casi no pasa transporte a esta hora.
 EXP almost not go-by transportation at this time
 ‘Almost no transportation goes by at this time.’
- h. Ello puede cantar un gallo por donde quiera y ellos lo reconocen.
 EXP can sing a rooster anywhere and they CL recognize
 ‘A rooster can sing anywhere and they recognize it.’

(from Bullock and Toribio, 2009, 57, ex. 14a-h)

However, unlike in languages like English or French, expletives in DSEC are not obligatory in the contexts just illustrated. Thus, next to the examples in (10), we find (11).

- (11) a. Había un viejo que tenía un gallo. (DSEC)
 was an old man that had a rooster
 ‘There was an old man that had a rooster.’

(from Bullock and Toribio, 2009, 70, ex 37)

- b. Como hubiera llovidos ...
 since had rained
 ‘Since it had rained ...’

(from Bullock and Toribio, 2009, 64, ex. 30)

Furthermore, as Hinzelin and Kaiser’s (2007) study shows, given the choice between an overt and a null expletive, urban speakers from the capital of El Cibao, as well as rural speakers from other regions in the Dominican Republic chose the null expletive most of the time (between 67% and 93% of the time, depending on the construction).

Expletive optionality has led Bullock and Toribio (2009, 56, fn. 9) to suggest that they may be a left-peripheral discourse marker (cf. Silva-Villar, 1998). In this sense, examples such as (12)-(13). These examples illustrate that *ello* can be used as a single response to a yes/no question, or in conjunction with *sí/no* ‘yes/no’.

- (12) a. ¿Quiere bailar? (DS)
 want to-dance
 ‘Do you want to dance?’
 b. Ello.
 EXP
 ‘Yes.’ (from Hinzelin and Kaiser, 2006, 65)
- (13) Ello sí/no. (DS)
 EXP yes/no
 ‘Yes/no.’ (from Hinzelin and Kaiser, 2006, 65)

Notice, however, that all the examples of *ello* in the literature involve postverbal subjects (cf. in particular Hinzelin and Kaiser’s 2006 extensive corpus of examples from different sources). Furthermore, (14) is ungrammatical. Since DS has an overwhelming SV word order tendency, the ungrammaticality of (14) is surprising if *ello* is strictly a discourse-related expletive.⁴

- (14) *Ello unas personas llegaron ayer. (DS)
 EXP some people.PL arrived.PL yesterday
 ‘Some people arrived yesterday.’ (from Jacqueline Toribio, p.c.)

Verbal agreement shows two patterns when *ello* appears. On the one hand, agreement can be 3rd person singular, particularly with existential *hay* ‘(there) is’. This pattern can be seen in (10a) above. A second pattern involves agreement with the postverbal subject as seen in (10c) above and in (15) below.⁵

- (15) Ello llegan guaguas hasta allá. (DS)
 EXP arrive.3.PL buses until there
 ‘Buses reach there.’ (from Toribio, 1993, 95 quoted in Hinzelin and Kaiser (2006, 69))

In the case of the 3rd person singular agreement pattern, the relevant question is whether *ello* is responsible for it, or whether there is an additional expletive that triggers agreement (cf. (16a) and b) respectively). The second option would predict the existence of overt preverbal subjects, which, as we have suggested, are not attested in the literature. The first option, on the other hand, suggests that no overt preverbal pronouns should co-appear with *ello*, particularly if they occupy the same position.

⁴This makes these expletives crucially different from the ones described in for Northern Iberian varieties in sect. 3.2.1.2.

⁵There are examples with *ello*, a plural verb and a null subject, but as Hinzelin and Kaiser (2006, 69) point out, it is not possible to know whether *ello* is an expletive, or an instance of a referential pronoun *ellos* that has its final /s/ dropped.

- (i.) Ello lo dijeron por radio. (DS)
 EXP CL said for radio
 ‘They said it on the radio.’ (from Toribio, 1993, 94 quoted in Hinzelin and Kaiser (2006, 69))

- (16) a. [CP [IP *ello* hay]]
 b. [CP *ello* [IP *pro* hay]]

To summarize, the data from Finnish and DSEC suggests that the implicational (2a) ‘thematic NS → expletive NS’ does not hold.

3.2.1.2 A CP-related expletive

Certain Romance varieties show an overt expletive with a slightly different distribution from the one just presented.⁶

Like DS, these varieties are NSLs, and like DS, they have overt expletives of different types: Weather-expletives (cf. (17)), clausal expletives (cf. (18)) and existential ones (cf. (19)).⁷

- (17) a. *Êle já orvalha.* (EP)
 EXP already drizzle
 ‘It is already drizzling.’
- b. *El chove.* (Galician)
 EXP rains
 ‘It is raining.’
- c. *Ello moja mucho.* (Leonese)
 EXP soaks a lot
 ‘It is soaking wet.’
- d. *Ello lloverá sidra, cigarrillos, corbatas, un epatante solomillo.* (19th c. Spanish)
 EXP rain-will cider, cigarettes, ties a dazzling sirloin
 ‘It will rain cider, cigarettes, ties, a dazzling sirloin.’
 (from Silva-Villar, 1998, 252, ex. 8)
- e. *Eso no está lloviendo tan duro.* (CCS)
 EXP not is raining so hard.
 ‘It’s not raining so hard.’
- (18) a. *Êle é certo que muitos se envergonhan de . . .* (EP)
 EXP is true that many CL shamed of
 ‘It is true that many people are ashamed of . . .’
- b. *E el non em fermoso percorrer mundo?* (Galician)
 QM EXP not was beautiful to-travel world
 ‘Wasn’t it beautiful to wander the world?’

⁶These varieties include three distinct clusters: One, languages spoken in the Northwestern and Western Iberian Peninsula (Galician, spoken EP, Leonese), another, 17th and 19th century Spanish (cf. Silva-Villar, 1998 and references quoted there), and the third one, Central Colombian Spanish, henceforth CCS (as observed by myself). The precise geographical or social distribution of this expletive in these varieties remains unclear, although in CCS it seems to be an archaic feature in the speech of older speakers.

⁷In the examples that follow, it is important to keep in mind that the expletive morpheme can have a separate referential interpretation, which should be discarded.

- c. Ello es necesario indagar que vida lleva. (18th-c. Spanish)
 EXP is necessary find-out what life leads
 ‘It is necessary to find out the type of life s/he is living.’
- d. Ellu foi que nun chegarun a casa-si. (Leonese)
 EXP was that not ended-up-3pl. to marry-CL
 ‘It (just so) happened that they never ended up getting married.’
 (from Silva-Villar, 1998, 252-3, ex. 9)
- e. ¿Eso no le parece que Gloria debe descansar? (CCS)
 EXP not CL seem that Gloria should rest
 ‘Don’t you think that Gloria should rest?’
- (19) a. Pois ele haverá castelhanos honrados? (EP)
 thus EXPL will-be Castilians honest
 ‘Are there honest Castilians?’
- b. Il hai cecais outro problema. (Galician)
 EXPL is perhaps another problem
 ‘There is perhaps another problem.’
- c. Ello hay por medio no sé qué papel de
 EXPL is in the middle not know.1.SG what document of
 matrimonio. (18th-c. Spanish)
 marriage
 ‘I don’t know what kind of marriage document is involved therein.’
 (from Silva-Villar, 1998, 253, ex. 10)
- d. Eso allá hay mucha gente. (CCS)
 EXPL there is many people
 ‘There is a lot of people there.’

These expletives appear productively in clause-initial position, but not in inversion contexts, for example in yes/no questions, as illustrated in (cf. (20)-(21)).

- (20) *Choverá el hoije[sic]? (EP)
 will-rain EXP today
 ‘Will it rain today?’ (from Silva-Villar, 1998, 254, ex. 12c)
- (21) a. ¿Eso estará lloviendo? (CCS)
 EXP will-be raining
 ‘(I wonder:) is it raining?’
- b. *¿Estará eso lloviendo?
 EXP will-be raining
 ‘(I wonder:) is it raining?’

In CCS, the expletive *eso* cannot appear either immediately after the verb (cf. (22b)) or clause-finally, which is a possible position for subjects in yes/no questions.

- (22) a. ¿Eso sabrá encontrar el restaurante? (CCS)
 EXP will-know find the restaurant
 ‘(I really wonder:) will s/he know how to find the restaurant?’
 b. *¿Sabrá encontrar eso el restaurante?
 will-know find EXP the restaurant
 c. *¿Sabrá encontrar el restaurante eso?
 will-know find the restaurant EXP

In embedded contexts, the varieties seem to split. Thus, Silva-Villar (1998, 255) reports that expletives are not possible in Galician and EP (cf. (23)), but in Central Colombian Spanish they are, under certain circumstances. Thus, (24a) is not very good in this variety, but (24b) is. Although the difference seems to be the presence of an overt subject in the first case vs. the second, it is not generally the case that overt subjects are not possible in that context, as seen in (25).

- (23) a. *Xa sei que el chove. (Galician/EP)
 of-course know.1.SG that EXP rains
 ‘Of course I know that it rains.’
 b. *Foi que ellu non legaran a casa-si. (Leonese)
 happened that EXP not ended-up.3.PL to marry-SE
 ‘It so happened that they ended up not getting married.’
 (from Silva-Villar, 1998, 255, ex. 13)
- (24) a. *A mi me parece que eso Gloria no debe salir. (CCS)
 to me CL seems that EXP Gloria not should go-out
 ‘I think that Gloria should not go out.’
 b. A mi me parece que eso no debe salir.
 to me CL seems that EXP not should go-out
 ‘It seems to me that she/he should not go out.’
- (25) Yo creo que eso la gente vive muy bien en esta ciudad. (CCS)
 I think that EXP the people live very well in this city
 ‘I think that people live very well in this city.’

As the preceding two examples show, the expletive is compatible with an overt nominative preverbal subject. This is also the case for EP and 17th-century Spanish, as seen in (26) for other varieties. This has led Silva-Villar (1998, 267) to suggest that the expletive in these examples is discourse-related and appears in Spec, CP, as “the spellout of the deictically or contextually bound event argument referring to the given point in space and time.”

- (26) a. Ele os lobos anda-n com fame. (EP)
 EXP the wolves go-3.PL with hunger
 ‘Wolves are hungry.’

- b. Ello yo no se porque mi padre no me llamó la torda o la
 EXP I not know why my father not CL.1.SG.ACC called.3.SG the thrush or the
 papagaya. (17th-c. Spanish)
 parrot
 ‘I don’t know why I was not called either thrush or parrot by my father.’
 (from Silva-Villar, 1998, 256-7 ex. 16)

This conception of the expletive as discourse-related is confirmed by word order patterns with overt subjects: In CCS, the expletive must precede the overt subject, as seen in (27), but it can precede or follow a dislocated object, as in (28).

- (27) a. Eso Gloria no debe salir. (CCS)
 EXP Gloria not must go-out
 ‘Gloria should not go out.’
 b. *Gloria eso no debe salir.
 Gloria EXP not must go-out
- (28) a. Los amigos, eso Pedro no los abandona nunca. (CCS)
 the friends, EXP Pedro not CL abandons never
 ‘His friends, Pedro doesn’t ever abandon them.’
 b. Eso los amigos, Pedro no los abandona nunca
 EXP the friends, Pedro not CL abandons never
 ‘His friends, Pedro doesn’t ever abandon them.’

This suggests that the expletive is located higher than the overt subject, as in (29), assuming, an expanded CP-structure, for example, as in Rizzi (1997, 257). Additionally, one would need to assume that the subject is fixed in a lower position.

- (29) [_{ForceP} [_{TopP} *eso*_{+Top} [_{FocP} [_{FinP} Subj [_{IP} [_I . . .]]]]]]

The expletive *eso* in CCS is not discourse neutral. Specifically, to be felicitous, it requires some kind of pre-supposed discourse situation. Consider the exchanges in (30)-(32). The first two examples show that *eso* is compatible with a subject-focus question or an object-focus question. In both cases, the question induces a presupposition (*x bought the creole potatoes* and *Transito bought x respectively*), and *eso* is possible in the answers.

- (30) a. ¿Quién compró las papas criollas? (CCS)
 who bought the potatoes ‘creole’
 ‘Who bought the ‘creole’ potatoes.’
 b. Eso compró las papas criollas Transito.
 EXP bought the potatoes ‘creole’ Transito
 ‘Transito bought the creole potatoes.’

- (31) a. ¿Qué compró Tránsito? (CCS)
 what bought Tránsito
 ‘What did Tránsito buy?’
 b. Eso Tránsito compró las papas criollas.
 EXP Tránsito bought the potatoes ‘creole’
 ‘Tránsito bought the creole potatoes’

(32a), on the other hand, does not allow for an answer with *eso*, as in (32b-c). Although the question also induces a presupposition (x happened), the presupposition is not specific enough to identify the situation, hence *eso* is not appropriate. Notice that the counterpart of (32b) without *eso* is grammatical, as seen in (32d).

Finally, if we compare (32) with (33), the main difference is that some extralinguistic element sets up a situation in discourse that alters the assertion structure in such a way that one can vaguely presuppose *someone broke a glass in the kitchen*, hence the answer with *eso* is possible.

- (32) a. ¿Qué pasó? (CCS)
 what happened
 ‘What happened?’
 b. #Eso Tránsito compró las papas criollas.
 EXP Tránsito bought the potatoes ‘creole’
 ‘Tránsito bought the creole potatoes’
 c. #Eso compró las papas criollas Tránsito.
 EXP bought the potatoes ‘creole’ Tránsito
 ‘Tránsito bought the creole potatoes.’
 d. Tránsito compró las papas criollas.
 Tránsito bought the potatoes ‘creole’
 ‘Tránsito bought the creole potatoes’

(33) [Context: There is noise of breaking glass coming from the kitchen]

- a. ¿Qué pasó? (CCS)
 what happened
 ‘What happened?’
 b. Eso Olga rompió un vaso.
 EXP Olga broke a glass.
 ‘Olga broke a glass.’

These contrasts provide us with a fairly specific intuition about the content and role of *eso*: it anchors the clause to a situation topic, an issue we will return to below.

To summarize the previous two sections, we have suggested that expletives can target at least two distinct positions: In DSEC, the expletive is incompatible with a preverbal subject, hence it must target a low position (IP or another subject-related position, cf. (16) above), whereas in CCS, Galician and other Northwestern Iberian languages, it targets a higher, CP-related topic position (cf. (29)).

3.2.2 Subject inversion & no null expletives

According to Gilligan's implication in (2b), if a language has free inversion, then it should have null expletives. Thus, it follows that no language should be like Spanish or Italian (with SV/VS orders) and have overt expletives. In Rizzi's original analysis, if a language has free inversion, this implies that some other category satisfies its EPP requirement. Assuming some version of the Avoid Pronoun Constraint (one that states that overt pronouns should be avoided whenever possible), then such a language should not show overt expletives.

Overt expletives are a rarity among the world's languages. In Gilligan's (1987) survey of over 100 languages, he finds no case in which the expletive cannot be null.⁸ Even in languages where overt expletives are present, we find a wide range of variation as to whether they are obligatory in all contexts. For example, in Icelandic, expletives are only obligatory in clause-initial position (as are impersonal expletives in German and Yiddish), but must be null in non clause-initial. By contrast, in Norwegian, Danish, Swedish or English, they are obligatory regardless of clausal position, as shown in (34)-(36).

- (34) a. þaðan var (*það) skammt til bæja. (Icelandic)
 from-there was there short-way to the-farms
- b. Derfra var *(der) ikke langt til gårdene. (Danish)
 from-there was there not long to the-farms
- c. Derifra var *(det) ikke langt til gårdene. (Norwegian)
 from-there was there not long to the-farms
- d. Därifrån var *(det) nära till gårdarna. (Swedish)
 from-there was there short-way to the-farms
 'From there it was a short way to the farmsteads.' (from Platzack, 1987, 387 ex. 17)
- (35) a. Nú er (*það) augljóst að Jón hefur barið Maríu. (Icelandic)
 now is it obvious that John has beaten Mary
- b. Nu er *(det) helt klart, at John har slået Maria. (Danish)
 now is it very clear that John has beaten Mary
- c. Nå er *(det) åpenbart at John har slått Maria. (Norwegian)
 now is it obvious that John has beaten Mary
- d. Nu är *(det) uppenbart att John har slagit Maria. (Swedish)
 now is it obvious that John has beaten Mary
 'It is obvious now that John has beaten Mary.' (from Platzack, 1987, 387 ex. 18)
- (36) a. Í dag hafa (*það) komið margir málvísindamenn hingað. (Icelandic)
 in today has there come many linguists here
- b. I dag er *(der) kommet mange lingvister hertil. (Danish)
 in today has there come many linguists here

⁸Although in his survey, several languages do not have data on this particular area.

- c. I dag har *(det) kommet mange lingvister hit. (Norwegian)
 in today has there come many linguists here
- d. Idag har *(det) kommit många lingvister hit. (Swedish)
 today has there come many linguists here
 ‘Today there have arrived many linguists.’ (from Platzack, 1987, 388 ex. 19)

It should be noted that none of these languages have null thematic subjects, although Icelandic has a fairly rich verbal paradigm (cf. Platzack (2003, 331)).

3.2.3 *That*-trace effects & overt expletives

The final implication from (2) predicts that one should not have languages in which *that*-trace effect violations are attested by expletives are overt. Cabrera (2007) presented El Cibao speakers were presented with a situation context and subsequently given the choice between an overt and a null complementizer in questions with a subject *wh*-word, as illustrated in (37a-b)). Under those conditions, 80% of El Cibao speakers preferred the null complementizer version (37b) (cf. Cabrera, 2007, table 4.16).

- (37) a. ¿Quién dijiste que trabajó en el turno de noche?
 who said that worked in the shift of night
 ‘Who did you say worked the night shift?’
- b. ¿Quién dijiste \emptyset trabajó en el turno de noche?
 who said \emptyset worked in the shift of night
 ‘Who did you say worked the night shift?’ (from Cabrera, 2007)

However, other native speakers of DS do allow for overt complementizers in those contexts (cf. (38), thanks to Jacqueline Toribio, p.c. for judgements).⁹

- (38) ¿Quién tú dices que compró mangos? (DS)
 who you say that bought mangoes
 ‘Who do you say bought mangoes?’ (from Jacqueline Toribio, pc.)

This suggests that this implication does not hold universally. Although it is not clear why these conflicting results arise, I take them to mean that at least for some speakers, deletion is not obligatory in these contexts.

Recall, however, that the original *that*-trace violation associated with the NSP should not be taken literally, but rather as the possibility of extracting from a postverbal position, an issue we will return to below.

⁹DS, like Puerto Rican Spanish does not show VS inversion in questions. Cf. Ordóñez and Olarrea (2006).

3.2.4 Free subject inversion and *that*-trace effect violations

This implication follows in the traditional analysis, *that*-trace effect violations (i.e. subject extraction in certain contexts) result from the availability of a postverbal subject position, also manifested in VS word orders (free inversion). The paradigmatic cases of this correlation come from Italian and Spanish. In these languages, it is possible to extract both subjects and objects without any changes in the complementizer system.

- (39) Chi credi che chiamerà questa sera? (Italian)
 who think.2.SG that will call.3.SG this night
 ‘Who do you think will call tonight?’

One can find several examples of this implication, for example in Bavarian German, Övdalian, DS and Lubukusu, but as I will argue below, both theoretical and empirical arguments argue against a direct connection between availability of a postverbal extraction site and *that*-trace effect violations. Additionally, Roberts and Holmberg (2010, 22, fn. 15) point out themselves that their proposed implicational seems to be counterintuitively backwards: Free inversion is not a prerequisite for *that*-trace effect violations, but rather, *that*-trace effects imply free inversion. As we will see below, they suggest that this apparent puzzle can be resolved if Rizzi and Shlonsky’s (2007) reformulation of what constrains subject extraction is correct.

3.2.4.1 Bavarian German and Övdalian

Bavarian illustrates a pattern where lack of inversion goes hand-in-hand with *that*-trace effects (cf. Bayer, 1984). Bavarian German, like German in general has a fairly rigid V2 word order, so Italian-style free subject inversion is not possible, and the complementizer is dropped in subject extraction (cf. 40a)).

- (40) a. Wer moanst du [mog d’Emma]? (Bavarian German)
 who think you loves Emma
 ‘Who do you think loves Emma?’
 b. Wer moanst du [daß d’Emma mog]?
 who think you that Emma loves
 ‘Who do you think loves Emma?’
 c. Weam moanst du [daß d’Emma mog]?
 who think you that Emma loves
 ‘Who do you think that Emma loves?’ (from Bayer, 1984, 210, ex. 2)

It should be noted that Bavarian German is a partial NSL: Only 2nd person subjects can be null (cf. (41a-b) vs. c)). This shows that the possibility of *that*-trace filter violations is not directly connected with the availability of NSs.¹⁰

¹⁰The 2nd person morpheme is the only one that uniquely encodes person and number. Additionally, NSs are only possible in embedded clauses as seen in (41c vs. (i.)), cf. Bayer (1984).

- (41) a. *Kumm-t *pro* noch Minga? (Bavarian German)
 come.3.SG to Munich
 ‘Will he/she/it come to Munich?’
- b. *Kumm *pro* noch Minga . . .
 come.1.SG to Munich
 ‘If I come to Munich . . .’ (from Bayer, 1984, 239, ex. 68c-d)
- c. Kummst *pro* noch Minga, dann muaßt *pro* me b’suacha
 come.1.SG to Munich then must.2.SG me visit
 ‘If you come to Munich you must visit me.’ (from Bayer, 1984, 211, ex. 7)

A similar situation can be observed in Övdalian, a variety of Mainland Scandinavian spoken in Sweden (cf. Rosenkvist, 2010b). In this language, 1st and 2nd plural subjects can be null (cf. (42)), but as (43) shows, VS word order is ungrammatical and the language has *that*-trace effects, illustrated in (44). Thus, the complementizer *at* must be dropped (cf. (44a)), or a resumptive pronoun must be inserted in the embedded clause (cf. (44b)). Alternatively, *so* can be used, but this complementizer is also a conjunction.¹¹

- (42) a. Byddjum i Övdalim. (Övdalian)
 live.1.PL in Älvdalen
 ‘We live in Älvdalen.’
- b. Ulið frarra nu.
 shall.2.PL leave now
 ‘You ought to leave now.’ (from Rosenkvist, 2010b, 231 ex. 1)
- (43) *Ar kumið Lasse. (Odvalian)
 has arrived Lasse
 ‘Lasse has arrived.’ (from Rosenkvist, 2010b, 239 ex. 10b)
- (44) a. Ukin truo’d d du (*at) uld kumå?
 who thought you would come
 ‘Who did you think would come?’
- b. Ukin truo’d d du at an uld kumå?
 who thought you that he would come
 Who did you think would come?
- c. Ukin truo’d d du so uld kumå?
 who thought you so would come
 ‘Who did you think would come?’ (from Rosenkvist, 2010b, 240, ex. 12)

(i.) *In Bayern redt *pro* Bairisch
 in Bavaria speaks.3.SG Bavarian
 ‘In Bavaria, he speaks Bavarian.’ (from Bayer, 1984, 211, ex. 6)

¹¹Övdalian verbs have identical forms for 1st, 2nd and 3rd person singular and 3rd person plural present tense and infinitive (*bait* ‘bite’). In the plural, 1st and 2nd person are distinct (*baitum*, *baitið* respectively), cf. Rosenkvist (2010b, 237).

3.2.4.2 Levantine Arabic

A similar situation to Bavarian and Övdalian arises in Levantine Arabic. Kenstowicz (1989, 264-265) points out that this variety has a *that*-trace effect like Germanic, as shown in (45). In the first example, the object wh-word can be extracted from the embedded clause, whereas in the second one, the subject wh-word cannot. As (45c) shows, dropping the complementizer allows for subject extraction.

- (45) a. ʔayy fuṣṭaan_i Fariid kaal innu l-bint iṣṭarat (Levantine Arabic)
 which dress Fariid said that the girl bought
 ‘Which dress did Fareed say that the girl bought?’
- b. *ʔayy bint_i Fariid kaal innu e_i iṣṭarat l-fuṣṭaan
 which girl Fariid said that bought the dress
 ‘Which girl did Fareed say that bought the dress?’
- c. ʔayy bint_i Fariid kaal e_i iṣṭarat l-fuṣṭaan
 which girl Fariid said bought the dress
 ‘Which girl did Fareed say that bought the dress?’
 (from Kenstowicz, 1989, 264, ex. 3a-c)

In this variety, NSs are restricted to main clauses (cf. (46a vs. b)); in embedded contexts, they must be obligatorily realized as a clitic on the complementizer (cf. 46c)).

- (46) a. (hiy) iṣṭarat l-fuṣṭaan. (Levantine Arabic)
 (she) bought the dress
 ‘She bought the dress.’
- b. *Fariid kaal innu iṣṭarat l-fuṣṭaan.
 Fariid said that bought the dress
 ‘Fareed said that s/he bought the dress.’
- c. Fariid kaal inn-ha iṣṭarat l-fuṣṭaan.
 Fariid said-she that bought the dress
 ‘Fareed said that she bought the dress.’ (from Kenstowicz, 1989, 264, ex. 4a-c)

In main clauses, the subject can be postverbal (cf. (47a)), but not in embedded clauses (cf. (47b)), so while free subject inversion may be possible in main clauses, it is not in embedded clauses. Since this is the environment crucial for licensing wh-extraction, then *that*-trace effects are expected.

- (47) a. Fariid kaal innu l-bint iṣṭarat l-fuṣṭaan. (Levantine Arabic)
 Fariid said that the girl bought the dress
 ‘Fareed said that girl bought the dress.’
- b. *Fariid kaal innu iṣṭarat l-bint l-fuṣṭaan.
 Fariid said-she that bought the girl the dress
 ‘Fareed said that she bought the dress.’ (from Kenstowicz, 1989, 264, ex. 5a-c)

3.2.4.3 Dominican Spanish

Whereas Italian represents the case where inversion coexists with *that*-trace violations, Bavarian, Övdalian and Levantine Arabic represent examples of where no inversion correlates with *that*-trace effects. Dominican Spanish (DS), on the other hand, seems to illustrate a case increasingly rigid word order (i.e. absence of free subject inversion) coexists with *that*-trace effect violations.

In his study of Caribbean Spanish, Ortiz López (2009, 88) finds that DS, Cuban and Puerto Rican speakers all overwhelmingly prefer the order SV (over 80% of the time) regardless of verb type or whether the subject is focused or not.¹² However, we do find evidence of alternative word orders in DS. In Bullock and Toribio's (2009) data, this order arises with presentational/existential sentences, such as (48). In (48a), the verb *haber* shows 1st person plural morphology, suggesting agreement with a null *pro*.¹³ In (48b), we see plural agreement between the verb *hacer* and *tres meses* 'three months'. *Hacer* is typically used in two contexts: To denote a period of time in the past (as in the example), and to denote weather-related states (such as *hace frío* 'it is cold'). In other varieties, it does not agree with the DP: *Hace 25 grados* 'it is.3.SG 25 degrees.

- (48) a. *Habemos pocas familias en Los Compos.* (DSEC)
 be.1.PL few families in Los Compos.
 'There are/we are a few families in Los Compos.'
 (from Bullock and Toribio, 2009, 54, ex. 5d)
- b. *Hacían como tres meses que no llovía.*
 be.3.pl approximately three months that not rain.
 'It was was approximately three months since it hadn't rained.'
 (from Bullock and Toribio, 2009, 54, ex. 5a)

There are a few other instances of VS word order in non-presentational contexts, for example in (9) above, repeated below. In this example, the subject *raiz* is a bare noun in postverbal position.

- (49) *Se fue, cogió un bastón, se encontró con su compadre y hablaron tanto que le*
 CL went, got a stick, CL met with his friend and talked so much that CL
salió raíz al ba'tón. (DSEC)
 came out root to-the stick
 'He left, got a (walking) stick, he ran into a friend, and they talked so much that the stick
 sprung roots.'
 (from Bullock and Toribio, 2009, 63, ex. 26b)

Given this increasingly fixed word order, one might expect DS to lose the ability to extract subjects over a filled complementizer and show canonical English *that*-trace effects, but this is not the case. Thus, subject extraction is acceptable in this variety (cf. (38)).

¹²By contrast to Ortiz López (2009), Cabrera's (2007) study, finds much lower frequencies of SV. In that study, speakers were presented with a contextually situated sentence with three possible variants, one option with a null subject, another one with a preverbal one, and a third option with a postverbal subject. Urban El Cibao speakers picked the null subject 33% of the time, the preverbal option 38% and the postverbal one 29%. It is not clear why Cabrera and Ortiz López's results vary so much. Part of it may be a test effect, but Ortiz López's results are more in line with general intuitions about Caribbean Spanish than Cabrera's. I take this to mean that DS is indeed losing free inversion.

¹³Cf. the discussion of similar examples in other Spanish varieties the examples in (23)-(24) in ch. 4.

- (50) *¿Quién tú dices que compró mangos?* (DS)
 who you say that bought mangoes
 ‘Who do you say bought mangoes?’ (from Jacqueline Toribio, p.c.)

3.2.4.4 Brazilian Portuguese

Like DS, Brazilian Portuguese (BP) is an interesting variety to consider because it is increasingly losing many of the NS properties (cf. Duarte, 1993, 1995, 2000; Barbosa et al., 2005). The European variety (EP) can have free inversion and null expletives, as shown in (51). Additionally, inversion is unrestricted, so for example, it is possible to have inverted definite subjects, as in (51a). By contrast, as Kato (2000, 250-1) points out, in BP the VS order is highly restricted: It is only productive with unaccusative verbs with indefinite subjects and in existential constructions (cf. (52a-b)). As seen in (52c), unaccusatives with definite subjects are ungrammatical and transitive sentences can only appear in the SVO word order (cf. (53)). All of the ungrammatical examples in (52)-(53) are possible in EP (cf. also Figueiredo Silva, 1994).¹⁴

- (51) a. *Telefonou o João.* (EP)
 called the Joao
 ‘Joao called.’
 b. *Chove.*
 rains
 ‘It is raining.’
- (52) a. *Tinha chegado muitas cartas.* (BP)
 had.3.SG arrived many letters
 ‘There arrived many letters.’ (from Kato, 2000, 250, ex. 46a)
 b. *Tem un gato embaixo da mesa*
 has a cat under of table
 ‘There is a cat under the table.’ (from Kato, 2000, 251, ex. 46b)
 c. **Cantam os pássaros/muitos pássaros.*
 sing the birds/many birds
 ‘The birds/many birds sing.’ (from Kato, 2000, 251, ex. 47c)
- (53) a. **Respondeu a pergunta a Maria.* (BP)
 answered the question the Maria
 ‘Maria answered the question.’
 b. **Respondeu a pergunta uma aluna.*
 answered the question a student
 ‘A student answered the question.’ (from Kato, 2000, 251, ex. 47a-b)

¹⁴Note, additionally, that in the examples in (52a-b), the postverbal subject does not agree with the verb, which shows 3.SG default agreement.

Given this situation, one might consider whether BP shows evidence of *that*-trace effects. According to Menuzzi (2000), it does, but only when the NS is extracted from a preverbal position. To show this, he uses a floated quantifier *todos* ‘all’ as a way of tracking extraction sites. With this in mind, consider (54). In the first example, *todos* ‘all’ is located postverbally, and extraction is possible, whereas when the floated quantifier is located preverbally, extraction is not, as shown in (54b). The corresponding structures appear in (55).¹⁵

- (54) a. Que rapazes, o Paulo desconfia [que gostem **todos** de Maria]? (BP)
 which boys the Paulo suspects that like all of Maria
 ‘Which boys does Paulo suspect all like Maria?’
 b. *Que rapazes, o Paulo desconfia [que **todos** gostem de Maria]?
 which boys, the Paulo suspects that (*all) like of Maria?
 ‘Which children does Paulo suspect all like Maria.’
 (from Menuzzi, 2000, 29, ex. 33)

- (55) a. [CP [que rapazes] ... [CP ~~que rapazes~~ [C que [IP gostem [vP [todos ~~que rapazes~~] ...]]]]
 ↑-----|-----↓
 b. *[CP [que rapazes] ... [CP ~~que rapazes~~ [C que [IP [todos ~~que rapazes~~] [Igostem. . .]]]]
 ↑-----|-----↓

Menuzzi (2000) further asserts that extraction from a preverbal position becomes possible in contexts of inflected infinitives, as shown in (56). For these cases, he assumes that the verb has moved to C. He claims that in this example, *nada* is dislocated to a position immediately dominating I', because in finite clauses, the only possible order would be S-*nada*-V. Assuming that V is in I, and the subject is in Spec, IP, it follows that *nada* must be adjoined to I'. Transposing this idea to inflected infinitives, the result is the structure in (57), where the inflected infinitive is higher than the subject and the dislocated object.

- (56) Que rapazes, o Manuel afirma terem todos nada oferecido de presente pr'a
 which boys the Manuel claims to-have.3.PL all nothing offered of present to-the
 Maria no aniversário dela?
 Maria in her birthday
 ‘Which boys does Manuel claim to all have offered nothing as a gift to Maria for her
 birthday?’
 (from Menuzzi, 2000, 35, ex. 40a)

- (57) *[CP [que rapazes] ... [CP ~~que rapazes~~ [C terem [IP [todos ~~que rapazes~~] nada [I' terem . . .]]]]
 ↑-----|-----↑-----↓
 (after Menuzzi (2000, 36, ex. 40b))

Below we will offer a slightly different account of these facts, but for now, note that if Menuzzi's analysis is correct, BP would be a case in which an increasingly fixed SV word order still allows for extraction from the postverbal subject position if the verb remains in I, or from either position if the verb moves to C.

¹⁵Menuzzi (2000, 30) points out that (54a) has a ‘recherché feeling about it’, but according to him, the contrasts are clear.

3.2.4.5 Subject extraction in Lubukusu

Several Bantu languages show extraction asymmetries that bear a resemblance to *that*-trace effects. Among them, Lubukusu has NSs (cf. Diercks, 2010, 172) and also alternative word orders (particularly with respect to locative inversion), but also has subject/object extraction asymmetries. This is unexpected under the implicational conditional being considered.

When a subject is extracted (both in *wh*-questions and relatives), it triggers obligatory agreement not only with the subject, but with the complementizer (cf. (58)). Thus, (58b), shows both subject and complementizer agreement (*ba* in both cases). By contrast, when objects are extracted, complementizer agreement is not possible, whether the *wh*-word is in situ or fronted (cf. (59)). Additionally, when the *wh*-object is fronted, an obligatory Comp-like element (glossed as PRED below) appears, as seen in (59c) vs. d).¹⁶

- (58) a. *ba-ba-ana ba-a-tim-a.* (Lubukusu)
 2-2-child 2.2-PST-run-FV
 ‘Children ran.’
- b. *naanu ba-ba-a-tim-a?*
 2.who 2.C-2.S-PST-run-FV
 ‘Who ran?’
- c. **naanu ba-a-tim-a?*
 2.who 2.S-PST-run-FV
 ‘Who ran?’ (from Wasike, 2007, 236 quoted in Diercks (2010, 85, ex. 6))
- (59) a. *Nafula a-a-siim-a Wafula.* (Lubukusu)
 1.Nafula 1.S-PRS-love-FV 1.Wafula
 ‘Nafula loves Wafula.’
- b. *Nafula a-a-siim-a naanu?*
 1.Nafula 1.S-PRS-love-FV who
 ‘Who does Nafula love?’
- c. **naanu Nafula a-a-siim-a?*
 who 1.Nafula 1.S-PRS-love-FV
 ‘Who does Nafula love?’
- d. *naanu ni-ye Nafula a-a-sim-a?*
 who PRED-1 1.Nafula 1.S-PRS-love-FV
 ‘Who is it that Nafula loves?’
 (from Wasike, 2007, 234 quoted in Diercks (2010, 85, ex. 5))

While most of these facts are consistent with Gilligan’s implicational universal that *that*-trace effect violations imply null expletives, and that free inversion implies *that*-trace effect violations (cf. (2) above), we find at least some languages that show what could be argued to be free subject

¹⁶Anti-locality restrictions on subject extraction in Kĩĩtharaka (cf. Muriungi, 2010) can also be seen as part of the broader range of subject/object extraction asymmetries.

inversion and no *that*-trace violations. Additionally, as noted above, the implication seem to be backwards: If *that*-trace effect violations are possible because a language can have postverbal subjects, and postverbal subjects are possible whenever there is a null subject to license the EPP, then null subjects should be the condition for free inversion and for *that*-trace effects. As suggested, Roberts and Holmberg (2010) propose that Rizzi and Shlonsky's account provides a solution to these shortcomings, and this is also the solution proposed by Diercks (2010), as we will see in sect. 4.5.2.

3.3 Summary

The data from the preceding sections suggests several conclusions:

1. Obligatory overt expletives are typologically very rare.
2. Expletives may target an extended range of IP-CP positions.
3. Implications related to *that*-trace effects and free inversion are not directly connected with null subjects.
4. Implications related to *that*-trace effects and free inversion require a different formalization.

The first generalization leads to the conclusion that the requirement for obligatorily overt expletives is not directly connected to the NSP. This conclusion is supported by Sigurðsson's (2010) proposal that EPP effects fall under two distinct categories: NP-movement and a Filled Left Edge Effect (FLEE). Recall that in Icelandic (and in German and Yiddish impersonal constructions), an overt expletive is only available in clause-initial position, but not otherwise. In the other Scandinavian languages (and in English), expletives are possible in other positions as well. For Sigurðsson (2010), this suggests that the syntax of certain expletives is driven by person agreement, which triggers their presence in several positions. Left-edge effects then separate languages that have obligatory expletives only in root contexts from languages where expletives appear in non-root-contexts.

I will return to *that*-trace violations in sect. 4.5.2, although the proposal there will not be directly connected with the NSP.

3.4 Variation in logophoricity

Another important feature of the NSP that will have to be accounted for is variation in person. As I mentioned in the presentation of the typology of NSLs (cf. sect. 2.3.3), several languages are sensitive to the 1st/2nd vs. 3rd person distinction. Specifically, we saw that Finnish and Hebrew allow for NSs under certain circumstances in 1st/2nd but not in 3rd person, and others like Shipibo allow the opposite pattern (overt in 1st/2nd and possibly null in 3rd). This has lead several researchers to argue that 1st and 2nd person head a syntactic projection in the left periphery (cf. Baker, 2008; Shlonsky, 2009; Camacho and Elías-Ulloa, 2010; Sigurðsson, 2010 among others). In addition

to those patterns, we also find evidence that subject extraction can be affected by person, and in this section we will consider three instances, inverted copular predicate in BP and in Spanish and locality of NSs in Shipibo.

3.4.1 Person effects in inverse copular clauses in Brazilian Portuguese and Spanish

Lima (2005) points out that subject-verb agreement, which is usually obligatory in BP, allows for two possibilities in inverted copular constructions, as illustrated in (60). If the predicate appears in preverbal position, agreement can target either the postverbal subject or it can be default, as in (60a), but if the predicate is postverbal, agreement with the subject is obligatory, as seen in (60b).

- (60) a. A causa da briga foi /foram as meninas. (BP)
 the cause-of-the fight was /were the girls
 ‘The cause of the fight was the girls.’
- b. As meninas foram /*foi a causa da briga.
 the girls were /*was the cause-of-the fight
 ‘The girls were the cause of the fight.’ (from Lima, 2005, 1, ex. 1-2)

By contrast, Lima (2005) points out that 1st and 2nd persons trigger obligatory agreement with the subject, regardless of word order. In particular, (61a) does not allow 3.SG agreement, in contrast to (61a). The same patterns can be observed in Spanish, cf. (62).¹⁷

- (61) a. A causa da briga *e /sou eu. (BP)
 the cause-of-the fight *is /am I
 ‘The cause of the fight is me.’
- b. Eu sou /*e a causa da briga.
 I am /*is the cause-of-the fight
 ‘I am the cause of the fight.’ (from Lima, 2005, 7, ex. 26-27)
- (62) a. Yo soy /*es la causa de la pelea. (Spanish)
 I am /*is the cause-of-the fight
 ‘I am the cause of the fight.’
- b. La causa de la pelea soy /*es yo.
 the cause-of-the fight am/is I
 ‘The cause of the fight is me.’
- (63) a. Tú eres /*es la causa de la pelea. (Spanish)
 you are /*is the cause-of-the fight
 ‘You are the cause of the fight.’

¹⁷Spanish has distinctive 2nd person agreement, unlike BP.

- b. La causa de la pelea eres /*es tú.
 the cause-of-the fight are /*is you
 ‘The cause of the fight is you.’

If one adopts Moro’s (1993) proposal for copular sentences (as Lima, 2005 does), these cases involve movement of either the subject or the predicate to preverbal position. In (60a), the version where the verb agrees with the postverbal subject can be analyzed as agreement probing downwards and finding the subject. The default case, in turn, can involve either a null expletive, or simply default agreement. In (60b), on the other hand, the closest agreeing category is the subject, so either *pro* has no place, or default agreement is dispreferred (because it is a last resort option, see chapter 7).

In the case of 1st and 2nd person, we can follow a suggestion by Lima that “(first) person licensing may be obligatorily related to a position high in the clause”, and adopt recent proposals (cf. Baker, 2008; Shlonsky, 2009; Sigurðsson, 2010) that a logophoric projection is located in the left periphery. As we will see later, if *pro* must be identified by an operator, and a 1st and 2nd person operator is situated in the left periphery, then the only option will be for *pro* to be interpreted as 1st and 2nd person, hence the agreement patterns in (61)-(63).

A similar pattern is seen in cleft-focus constructions in Spanish (cf. Camacho, 2001). The copular verb in these constructions must obligatorily agree with 1st and 2nd person, but when the focused predicate is not a subject, agreement is a default 3rd person, as seen in (64).

- (64) a. Fuiste tú quien no vino. (Spanish)
 were.2.SG you who not came
 ‘It was you who didn’t come.’
 b. *Fue tú quien no vino.
 was.3.SG you who not came
 ‘It was you who didn’t come.’
 c. Fue a ti a quien vimos.
 was.3.SG to you to whom saw
 ‘It was you who we saw.’

3.4.2 Person effects in Shipibo switch-reference clauses

A third case where extraction is sensitive to logophoricity is Shipibo NSs. Recall from sect. 2.3.3 that Shipibo is a mixed NS language that only marks plural on the verb in the 3rd person, and that overt subjects are obligatory in 1st and 2nd person but optional in 3rd person (cf. (65)).

- (65) a. Lima-n-ra noko-ke. (Shipibo)
 Lima-DIR-EVID arrive-PERF
 ‘He/she went to Lima.’
 b. Ea-ra Lima-n noko-ke.
 1-EVID Lima-DIR arrive-PERF
 ‘I arrived in Lima.’

- c. Mia-ra Lima-n noko-ke.
 2-EVID Lima-DIR arrive-PERF
 ‘You arrived in Lima.’ (from Camacho and Elías-Ulloa, 2010, 72, ex. 13)

Embedded clauses in Shipibo are most frequently switch-reference clauses with a switch-reference morpheme, as illustrated in (66). As that example shows, 3rd person subjects can be null, whereas 1st/2nd person subjects must appear overtly at least in one of the clauses or in both (cf. (67a-c)). If no subject appears in those sentence (as in 67d), it can only be interpreted as 3rd person.

- (66) Jawen atsabo oroşon-ra pei-bo ate-ke. (Shipibo)
 POSS yucca grow-PRIOR.SAME.SUB.TR-EVID leave-PL cut-PERF
 ‘He grew yucca and cut the leaves.’
- (67) a. En westiora ipo chachi-şon-ra en Quique
 1P a carachama catch-PRIOR.SAME.SUB.TR-EVID 1P Quique
 kena-ke. (Shipibo)
 call-PERF
 ‘I caught a carachama (type of fish) and called Quique.’
- b. En westiora ipo chachi-şon-ra Quique kena-ke.
 1P a carachama catch-PRIOR.SAME.SUB.TR-EVID Quique call-PERF
 ‘I caught a carachama (type of fish) and called Quique.’
- c. Westiora ipo chachi-şon-ra en Quique kena-ke.
 a carachama catch-PRIOR.SAME.SUB.TR-EVID 1P Quique call-PERF
 ‘I caught a carachama (type of fish) and called Quique.’
- d. #Westiora ipo chachi-şon-ra Quique kena-ke.
 a carachama catch-PRIOR.SAME.SUB.TR-EVID Quique call-PERF
 ‘I caught a carachama (type of fish) and called Quique.’

In Camacho and Elías-Ulloa (2010), we analyze the 1st/2nd person paradigm in a similar way as we have proposed for BP and Spanish copular structures: They involve a logophoric head that must be overtly realized. In that analysis, the overt pronoun is the overt copy of a moved pronoun, which can surface in either position. 3rd persons, on the other hand, are recovered through the discourse, so they need not be overt.

3.5 Summary

In this chapter, I have reviewed the empirical predictions of the different correlations of NSP properties, as proposed by Gilligan (1987). First, I have suggested that the correlation between null thematic and expletive subjects does not necessarily hold if one takes optionally overt expletives as counterevidence. Languages like Dominican Spanish from El Cibao has overt expletives despite having partial NSs. I have also argued that the DSEC expletive is best seen as an IP-expletive, by comparison to other Romance varieties that have a CP-expletive. Regarding the correlation between subject inversion and null expletives, I have argued that the relative rarity of overt expletives

in the world's languages suggests that these are independent properties. Furthermore, comparison within Scandinavian languages, some of which allow for expletive inversion and some of which do not also suggests that the correlation between inversion and expletives is not tenable. Regarding *that*-trace effects, I have argued that DS shows a case of a language where both *that*-trace effect violations do not hold, whether they are viewed as complementizer deletion and as subject extraction over a filled complementizer, and overt expletives still exist.

In the second part of the chapter, I have turned to the correlation between free subject inversion and *that*-trace effect violations, and I have suggested that it does not hold for at least some Bantu languages, and that *that*-trace effect violations may require an alternative explanation unconnected to the NSP. This alternative will be explored in 4.5.2.

Finally, I have argued that in addition to the EPP-related properties, the NSP must also account for the person effects that oversee the distribution of NSs in many languages, and in particular, the way they interact in cases of movement.

Chapter 4

The nature of the EPP and the Null Subject Parameter

4.1 Introduction

If one assumes the EPP, it follows that clauses without overt subjects must have a null one. One question this assumption raises is what category can satisfy the EPP in a NSL. Related to this question is the status of overt subjects: Are they in an A or A' position? Do they have properties similar to dislocated phrases?¹

Several researchers have proposed a parametric version of the EPP. For example, Alexiadou and Anagnostopoulou (1998) argue that in languages like Greek or Spanish, the EPP is satisfied by the agreement-rich verb moving to I. However, Holmberg and Nikkane (2002) point out that in Finnish, the verb also has rich inflection which licenses null subjects in many contexts, it moves to I, but overt expletives are still required in certain contexts. This leads them to propose two distinct requirements: One that can be satisfied a verb moving to I, and another one that is active in some languages but not in others, that forces the Spec, IP to be filled.

Others like Manzini and Savoia (2002) propose that the EPP entails a D feature in IP. Languages like English lack a specialized element to realize D, hence a DP must surface. Consistent NSLs, on the other hand, have a weak D, so they need not realize it at PF. Finally, Northern Italian dialects, which have a subject clitic, have a strong D which surfaces as the subject clitic.

In the following sections, I will explore the three most prominent answers to the question what fulfills the EPP in a NSL? After that, I will review the evidence for separate EPP requirements, and in chapter 8, I will discuss the position of preverbal subjects in NSLs.

As Saab (2010) points out, the question of what satisfies the EPP in a NSL has been answered in three ways. The first and probably most prolific answer, suggests that an empty category, *pro* does (cf. Chomsky, 1981; Rizzi, 1982, 1986a; Rizzi and Shlonsky, 2007 and many others). The second

¹The advent of bare phrase structure proposals cast some doubts on the theoretical viability of a unified EPP. Formulations of the EPP as a feature that attracts a category have the same problem: Given the assumption that constituents that AGREE can do so at a distance, there is no principled way to account for why subjects must appear in preverbal position. See sect. 4.5. In Rizzi and Shlonsky's analysis, the equivalent of the EPP, the Subject Criterion assumes a Spec, head relationship, which is no longer considered a theoretically well-defined configuration.

one assumes that pronominal subjects are deleted versions of overt pronouns (cf. Perlmutter, 1971; Saab, 2009; Roberts, 2010a, among others). The third one assumes that the inflectional properties of the verb can be pronominal, hence they can satisfy the EPP (cf. Jelinek, 1984; Borer, 1986; Ordóñez, 1997; Ordóñez and Treviño, 1999; Alexiadou and Anagnostopoulou, 1998; Kato, 1999; Barbosa et al., 2005; Barbosa, 2010; Sigurðsson, 2011 among others).

4.2 *Pro*

Assuming a formulation of the EPP in terms of a requirement to have a certain structural position filled, one is committed to the existence of a null pronominal (*pro*). For some researchers, *pro* is a syntactically distinct category, following Chomsky (1981). In this line, Rizzi (1986a) articulates a very influential theory of how *pro* fit into the general typology of referential elements. In his account, *pro* is subject to two independent conditions, presented in (1).

- (1) a. *Licensing of pro*: *pro* is governed by X^0 .
 b. *Identification of pro*: Let X be the licensing head of an occurrence of *pro*: then *pro* has the grammatical specification of the features on the X coindexed with it.
 (from Rizzi, 1986a, 519-20)

In a NSL, a thematic subject *pro* is governed and identified by I^0 . In English or French, on the other hand, I^0 is not rich enough to identify *pro*, so the identification condition fails and subjects must be overt. These same conditions could conceivably account for expletive *pro*, if one assumes that they must also be identified. However, given their non-thematic, non-referential status, one wonders why identification would be a requirement of expletives.

4.2.1 The existence and position of *pro*

In some sense, the crucial debate with respect to *pro* has to do with whether there is a separate category that satisfies certain syntactic requirements, or whether these are satisfied by pronominal agreement, as Holmberg (2005) points out.² In this section I will present several arguments that are compatible with both proposals, although for the sake of presentation I will use the term *pro*. I will also point out where the two theories make distinct predictions, generally favoring an independent *pro* over pronominal agreement.

The two main theoretical arguments in favor of *pro* stem from Theta Theory and the EPP. If one assumes that theta roles must be discharged, it follows that there must be some syntactic category that receives the subject theta-role, even when there is no overt subject, hence this category must be either *pro* or a pronominal agreement. On the other hand, if all clauses must have a subject (regardless of their thematic structure), then even verbs that do not assign an external theta role must have an implicit expletive subject.

Other more empirically-oriented reasons to propose *pro* are linked to the comparative distribution of overt and null subjects in a number of languages, a matter to which I turn below.

²The notion that inflection has argumental properties will be developed in detail in sect. 4.3 below. Essentially, the idea is that inflection has pronominal properties and can receive a theta role and satisfy the EPP in certain contexts.

4.2.1.1 Presentational sentences

Burzio (1986, 129-130) notes that in Italian presentational sentences (cf. also Cardinaletti, 1997, 36), overt subjects must be postverbal (cf. (2)) and in that context, NS are not possible (cf. (3)).

- (2) a. *Io ci sono alla festa. (Italian)
 I CL am at-the party (from Cardinaletti, 37, ex. 8a)
- b. Ci sono io alla festa.
 CL am I at-the party
 'I am there at the party.' (from Burzio, 1986, 130, ex. 106a)
- (3) *Ci sono alla festa. (Italian)
 CL am at-the party (from Burzio, 1986, 130, ex. 106b)

A similar situation exists in Spanish. For example, if someone knocks at the door and I ask *¿Quién es?* 'who is it?', the answer could be (4a), with a postverbal overt subject, but not (4b) with a preverbal one, or (4c), with a null subject.

- (4) a. Soy yo. (Spanish)
 am I
 'It's me.'
- b. #Yo soy.
 I am
 'I am.'
- c. #Soy.
 am
 'I am.'

In Spanish at least, this paradigm seems to follow from the assertion structure triggered by the question: *Who is it?* is only compatible with focus on the subject, which in that case must be clause-final (accounting for the contrast between (4a-b)). The ungrammaticality of (4c) can follow either because *pro* must be preverbal (as Burzio and Cardinaletti assume), or because it must receive nuclear stress in order to be focused, and since it is null, it cannot.

Given that all of these forms have identical inflectional endings, these data present a challenge for the pronominal agreement analysis. One possibility consistent with pronominal inflection would explain the distributions just observed from nuclear stress placement, as suggested in the preceding paragraph. However, in Spanish, certain verb forms are stress-final and stress falls on the vowel that indicates person/number. So if the inappropriateness of (4) is due to the impossibility of assigning nuclear stress to the focused subject, a simple-minded extension of this analysis would predict that (5a) should be appropriate in that context, because the syllable that carries stress also conveys inflectional properties. However, these examples remain ungrammatical.

- (5) a. ¿Quién fue el que rompió la puerta? (Spanish)
 who was the that broke the door
 'Who was it that broke the door?'

- b. Fue él.
was.3.sg he
'It was him.'
- c. #Fue.
was.3.sg
'(S/he) was.'

4.2.1.2 Copular sentences with nominal predicates

In Italian and Spanish, the order of the subject and the predicate can be reversed in copular sentences, as seen in (6). However, only the preverbal element can be null, whether it is the subject (cf. (7a)) or the predicate (cf. (7b)). Neither constituent can be null in postverbal position (cf. (8)). This leads Cardinaletti (1997) to conclude that the null element (*pro*) is preverbal.

- (6) a. Io sono il presidente (Italian)
I am the president
'I am the president.'
- b. Il presidente sono io.
the president am I
'I am the president.' (from Cardinaletti, 1997, 40, ex. 27)
- (7) a. Sono il presidente. (Italian)
am the president
'I am the president.'
- b. Sono io.
am I
'It's me.'
- (8) a. *Io sono. (Italian)
I am (from Cardinaletti, 1997, 40, ex. 30)
- c. *Il presidente sono.
the president am (from Cardinaletti, 1997, 40, ex. 29a)

Although these facts argue for the existence of *pro*, I can think of two objections to using them as evidence for the preverbal position of *pro*. First, by definition, the null category in (7a) is not the subject, since the verb agrees with *io* 'I'. It may be an expletive in Spec, IP required for independent reasons, but it doesn't agree with the subject. Second, Cardinaletti (1997, 40) acknowledges a potential counteranalysis to her own quoting Solà (1992, 156). Solà points out that the copular construction "is precisely used to focalize the subject, hence it cannot be dropped." Cardinaletti counters that this explanation does not apply to (8a), because the postverbal predicate is not focalized. The facts described until now are identical in Spanish, but at least in this language, the postverbal constituent would receive nuclear stress, if Zubizarreta (1998) is correct. This automatically discards *pro* from a postverbal position in the Spanish counterpart of this example.

Note, for example, that in Spanish the counterpart of (8a) would be acceptable if the subject is emphatic. Thus, a question like (9a) can be answered either with (9b) or with (9c), with a null predicate. In this example, nuclear stress is no longer on the last constituent of the clause, and the result becomes grammatical.

- (9) a. ¿Quién es el presidente? (Spanish)
 who is the president
 ‘Who is the president?’
- b. Soy yo.
 am I
 ‘It’s me.’
- c. YO soy.
 I am
 ‘I am.’

One could claim that in (9c) the pronoun, by virtue of being emphatic, is in a different position than in (8a). However, this does not explain why the predicate, which must be postverbal (cf. the ungrammaticality of (10a)), can be null. On the other hand, if the distributions in (6)-(9) follow from nuclear stress placement, then the grammaticality of (9c) also follows.

- (10) a. *YO el presidente soy. (Spanish)
 I the president am
- b. YO soy el presidente
 I am the president
 ‘I am the president.’

As with the data on presentational sentences, these facts are not straightforwardly compatible with a pronominal agreement theory, since in all cases we have the same type of inflection.

4.2.1.3 Floated quantifiers

Roberts (2010a, 72) quotes Rizzi (1987) as showing that only preverbal subjects license a floated quantifier (cf. (11a) vs. b). In Sportiche’s (1988) account, this follows from a requirement that the floated quantifier be c-commanded by the element it is associated with. Since a floated quantifier is grammatical with a null subject, it follows that *pro* must c-command it, hence it must be preverbal.³ This contrast is not easily explainable if inflection is pronominal, since in both (11a) and (11b) AGR/INFL c-commands the quantifier.

- (11) a. I bambini sono andati tutti via. (Italian)
 the children are gone all away

³One additional assumption is that the quantifier is located in the lower subject position.

- b. *Sono andati tutti via i bambini.
 are gone all away the children
 ‘All the children have gone away.’ (from Roberts, 2010a, 72, exs. 20b-c)

- (12) Sono andati tutti via. (Italian)
 are gone all away
 ‘They have all gone away.’ (from Roberts, 2010a, 72, exs. 20b-c)

Cardinaletti (1997, 50-11) extends the floated quantifier test one step further to argue that *pro* occupies the specifier closest to I. She notes that in (13a), the preverbal quantifier is ungrammatical, even though the example has *pro*. This result follows if the structure of that example is (13b), where *pro* does not c-command *tutti* and not (13c), where it does.⁴

- (13) a. *Non so se gli studenti, questo libro, *tutti* l’ hanno comprato. (Italian)
 not know if the students, this book, all CL have bought
 (from Cardinaletti 1997, 50, ex. 63)
- b. Non so se gli studenti, questo libro, [_{Agrs1P} *tutti* [_{Agrs1P} *pro* l’ hanno comprato]]
- c. Non so se gli studenti, questo libro, [_{Agrs1P} *pro* [_{Agrs1P} *tutti* l’ hanno comprato]]

The floated quantifier evidence for *pro* is not as compelling as it might seem at first sight. In a sentence like (11b), by Cardinaletti’s assumptions, *pro* occupies Spec, IP, but in such case, it should be able to c-command the floated quantifier.

Alternatively, one could argue that *pro* does not occupy Spec, IP. In such a case, something else must fulfill the EPP requirement, and the natural candidate would be for the overt subject *i bambini* ‘the children’ to have a deleted copy in that position. This assumption is consistent with Zubizarreta’s (1998, 136) analysis of VOS word order in Italian, which she takes to be derived from the underlying SVO order, “where S is in the specifier of a Focus projection above TP” (p. 136) and TP is further adjoined to FP. However, if this is correct, no *pro* appears in (11b), hence it is not evidence of its position. The ungrammaticality of that example would then follow from the fact that the overt subject does not c-command the floated quantifier.

In Spanish, the floated quantifier evidence does not support the status of *pro* as adjacent to Spec, IP. In this language, the example corresponding to (11a) is also ungrammatical (cf. (14a)), however the one corresponding to (13) is not (cf. (14b)).

- (14) a. *Salieron todos/cada uno a la calle los niños. (Spanish)
 went-out all/each one to the street the children
- b. No sé si los estudiantes, este libro, todos/(cada uno) lo han comprado.
 not know if the students, this book, all/(each one) CL have bought
 ‘I don’t know if the students have all/each bought this book.’

Following Cardinaletti’s logic, if *pro* is a weak pronoun in Spanish, (14b) should be ungrammatical, since the floated quantifiers would not be c-commanded by *pro* (cf. the structure in (13b)).

⁴In Cardinaletti’s proposal, Agrs1P and Agrs2P are different subject positions for strong pronouns, R-expressions and weak pronouns respectively.

4.2.1.4 Agreement in the Ancona variety

Another argument for the preverbal position of *pro* stems from the Central Italian dialect spoken in Ancona (cf. Cardinaletti (1997, 39-40)). In this variety, when the subject is postverbal, full agreement is not obligatory, and the verb will show 3rd singular morphology. Preverbal agreement, on the other hand, must be matched fully. Thus, in (15a) a 3rd person plural postverbal subject appears with a 3rd person singular verb, but the preverbal counterpart must agree in full (cf. (15b-c)).

- (15) a. Questo, lo fa sempre i bambini. (Ancona variety)
 this.ACC CL do.3.SG always the.PL children.PL
 'The children always do this.'
- b. *Questo, i bambini lo fa sempre.
 this.ACC the.PL children.PL CL do.3.SG always
- c. Questo, i bambini lo fanno sempre.
 this.ACC the.PL children.PL CL do.3.PL always
 'The children always do this.' (from Cardinaletti, 1997, 39, ex. 21)

Null subjects pattern like preverbal ones, i.e. they require full plural agreement, so that (16a) cannot be interpreted as 'they always do this', but (15b) can (and must). Even if some element like a floated quantifier signals plurality, *pro* cannot be interpreted as plural, as seen in (17a). If *pro* is always preverbal, agreement must be full as it is with overt subjects, hence both *pro* and the verb in (17a) are plural and combining a singular *pro* with a plural floated quantifier would result in mismatching number.

- (16) a. #Questo, lo fa sempre. (*pro* = 3rd pl.) (Ancona variety)
 this.ACC CL.ACC do.3.SG always
 'They always do this.'
- b. Questo, lo fanno sempre.
 this.ACC CL.ACC do.3.PL always
 'They always do it.' (from Cardinaletti, 1997, 39, ex. 22)
- (17) a. *Questo, lo dimostra tutte. (Ancona variety)
 this.ACC CL.ACC demonstrate.3.SG all.3.FEM.PL
- b. Questo, lo dimostrano tutte.
 this.ACC CL.ACC demonstrate.3.PL all.3.FEM.PL
 'They always demonstrate it.' (from Cardinaletti, 1997, 39, ex. 23)

Once again, this distribution poses a challenge to the pronominal agreement hypothesis. First of all, if agreement is always pronominal, one expects a referential clash between singular agreement and the plural DP in (15a). Second, the word order asymmetries observed in (15a-c) are unexpected, if agreement is always argumental. As a consequence, the ungrammaticality of (17a) becomes surprising, given the grammaticality of (15a).

4.2.1.5 Section summary

This section has summarized arguments in favor of the existence of *pro* and the hypothesis that *pro* is preverbal. In particular, it has provided data from presentational sentences in Italian and Spanish, copular sentences with nominal predicates, floated quantifiers and agreement patterns in SV and VS orders in the Italian dialect spoken in Ancona. Although these arguments suggest the existence of *pro*, I noted that an alternative based on the idea that many of the ungrammatical cases follow because *pro* cannot be emphasized by stress challenge the preverbal nature of *pro*. This same alternative would have to be invoked to save the pronominal agreement alternative, but the data from copular and presentational sentences and word order asymmetries in the Ancona variety raise a challenge for that option. Finally, the agreement facts in the Ancona variety presents a solid argument in favor of the preverbal location of *pro*. The floated quantifier data does not immediately translate to languages like Spanish, and likewise can have an alternative analysis.

4.3 AGR/INFL. The pronominal agreement hypothesis

Several researchers have argued that the existence of *pro* as an independent category is undesirable, mostly for theoretical reasons. Holmberg (2005, 537-538), for example, argues that given certain minimalist assumptions about the way agreement works, *pro* can no longer be considered an independent category. In most versions of this framework, empty categories are assumed to be deleted copies of overt categories, so strictly speaking, there are no null categories specified in the lexicon. This means, according to him, that either AGR/INFL is interpretable (i.e. pronominal), or *pro* is the result of deleting an overt pronoun. In this section, we will consider the main proposals that have advanced the first alternative, namely that agreement has pronominal properties that satisfy both the Theta Criterion and the EPP.

The pronominal agreement hypothesis was initially rooted in Hale (1983); Jelinek (1984) and Borer (1986), and could be stated as in (18).

(18) Pronominal Agreement Hypothesis

- a. AGR/INFL verbal morphology may license the EPP.
- b. Morphological affixes can receive theta roles.

The specific instantiation of (18a) depends on how one formulates the EPP, and whether one assumes a checking theory of features or an AGREE version of syntactic concord. For current purposes it does not matter. The second statement of the hypothesis in (18b) is explicitly proposed by Jelinek (1984, 44) and more recently by Holmberg (2005, 537). Thus, an example like (19a) would have the underlying structure in (19b-c). In (19b), the theta role is assigned by the verb to its pronominal person/number morpheme, and subsequent movement to I checks the EPP.⁵

- (19) a. Despertaron. (Spanish)
 awake-3PAST.PL
 ‘They woke up.’

⁵Alternatively, I AGREES with the verb in an EPP feature, and movement is independently triggered.

- b. [_{VP} despert-aron_{+D}, θ]
 c. [_{IP} [_{I'} I+ despertaron_{+D}, θ [_{VP} t_v]]]

Regarding the first part of the pronominal argument hypothesis (EPP licensing), its most recent incarnation has been developed by Alexiadou and Anagnostopoulou (1998) (henceforth A & A).⁶ Specifically, they make the claim that the EPP is parametric: In languages like English, with poor verbal morphology, it must be satisfied either by inserting an overt expletive in Spec, IP (cf. (20)), or by merging a DP in that position. In languages like Spanish or Greek, rich verbal morphology enables the verb itself to check the EPP by moving to I, as in (21).

- (20) a. There arrived a man. (English)
 b. [_{IP} there_{EPP} I arrived a man]
- (21) a. Llegó una mujer. (Spanish)
 arrived a woman
 ‘A woman arrived.’
 b. [_{IP} [_{I'} I+ llegó_{+D} [_{VP} una mujer t_v]]]

Following Chomsky (1995), Alexiadou and Anagnostopoulou (1998) argue that the EPP should be formalized as a D(eterminer) feature present on the inflectional node that needs to be checked. Since the verb inflectional endings are pronominal, then they can check the D feature on I. In English or French, on the other hand, the verb lacks this pronominal property, hence the D feature on I (the EPP feature) must be checked by a full XP raising to the Spec, IP.⁷ As we will see below in more detail, three consequences follow from this assumption: First, VS word orders in NSLs do not involve a null expletive, second, overt subjects do not occupy an A-position and third, arguably, only postverbal subjects should also be considered in an A-position.

4.3.1 Pronominal agreement and reference

If agreement is pronominal, one expects to find cases in which it acts as the source of reference. Jelinek (1984, 43), following Ken Hale’s insights, proposes that the node AUX (i.e. inflectional, INFL) is “a constituent containing case-marked, fully referential clitic pronouns that serve as verbal arguments.” As an example of how these clitics and not the overt nominal control reference, she points to example (22). The agreement clitic is *-rlipa* ‘1.PL., inclusive, NOM’, but the nominal *yapa* ‘person’ is 3rd person. The reflexive clitic *-nyanu* is also interpreted as a 1st person plural, suggesting that the agreement clitic and not the overt nominal controls co-reference.

⁶Manzini and Savoia (2002) have a different version of how the EPP is satisfied. According to them, the EPP is a D feature in IP, which is realized or checked by a DP in English, not overtly realized in Italian/Spanish or instantiated as a clitic in Northern Italian dialects.

⁷A & A argue that the locus of EPP checking is AGRSP, and that one crucial difference between NSLs and NNSLs is that in the former, AGRSP lacks a Specifier, following Contreras (1991).

- (22) Puyukuyuku-puru, kula-lpa-rlipa-nyanu yapa- \emptyset
 fog-WHILE NEG-IMPERF-L.PL(INCL)-NOM-REFL person-ABS
 nya-ngkarla. (Warlpiri)
 see-irrealis
 ‘We (plural inclusive) cannot see one another (as) person (s) (i.e., our shapes or figures)
 when it is foggy.’ (from Hale, 1983, 33 quoted in Jelinek (1984, 46, ex. 17))

Jelinek (1984, 48) also quotes examples from Spanish that seem to pattern in a similar way as Warlpiri (cf. also Ordóñez and Treviño, 1999, 58). In an example like (23), a plural DP can appear as the subject of a 1st, 2nd or 3rd person verb and the subject’s reference is interpreted depending on inflectional values. Thus, in (23a), the subject includes the speaker among the set of students, whereas in (23b) it does not.

- (23) a. Los estudiantes tene-mos mala memoria. (Spanish)
 the students have-1.PL bad memory
 ‘We students have bad memory.’
 b. Los estudiantes tien-en mala memoria.
 the students have-3.PL bad memory
 ‘The students have bad memory.’

This example suggests that the referential properties associated with the external theta role in (23) are determined by the verb’s inflectional morphology and not by the overt DP. Example (24) leads to the same conclusion. In that case, coreference between the main clause subject and the the object in the adjunct clause depends on the main-clause verb’s inflection: If they match (as in (24a, c)), coreference is possible, if they don’t (as in (24b)), no coreference obtains.

- (24) a. [Los estudiantes]_i sali-mos de la reunión después de que nos_i
 the students left-1.PL of the meeting after of that CL.3
 acusara. (Spanish)
 accused.3.SG
 ‘[The students]_i left the meeting after s/he accused us_{i/*j}.’
 b. *[Los estudiantes]_i sali-mos de la reunión después de que los_i acusara.
 the students left-1.PL of the meeting after of that CL.3 accused.3.SG
 ‘[The students]_i left the meeting after s/he s/he accused them_{*i/j}.’
 c. [Los estudiantes]_i sali-eron de la reunión después de que los_i acusara.
 the students left-3.PL of the meeting after of that CL.3 accused.3.SG
 ‘[The students]_i left the meeting after s/he s/he accused them_{i/j}.’

On the other hand, if the DP is not associated with an agreeing element, it recovers its ability to determine co-reference, as seen in (25). In this example, *los estudiantes* ‘the students’ is an object not associated with any functional category with overt inflectional morphology and it cannot be coreferential with *nosotros* ‘us’ because of the person mismatch.

- (25) Acusaron [a los estudiantes]_i, después de que él se pelease con
 accused to the students after of that CL.3 he fight.3.SG with
 nosotros_{*i/j}. (Spanish)
 us
 ‘They accused [the students]_i after s/he fought with us_{*i/j}.’

These paradigms suggest that Spanish, like Warlpiri fits best in the view that agreement is referential and therefore controls coreference. However, notice that this cannot be the whole picture. First, the patterns shown above are only possible with 3rd person, plural DPs, as seen in (26), where the 3rd person singular DP cannot be coreferential with a 1st person singular verb ending (contrasting with (23a)).

- (26) a. El estudiante tien-e mala memoria. (Spanish)
 the student has-3.SG bad memory
 ‘The student has bad memory.’
 b. *El estudiante teng-o mala memoria.
 the students have-1.SG bad memory
 ‘The student have bad memory.’

Second, only DPs, not pronouns, or even a pronoun+DP combination show the pattern in (23). Thus, in (27), the overt pronoun *nosotros* ‘we’ blocks 3rd person agreement on the verb, and in (28), *ellos* ‘they’ blocks 1st person agreement on the verb.

- (27) a. *Nosotros los estudiantes tien-en mala memoria. (Spanish)
 we the students have-3.PL bad memory
 ‘We the students have bad memory.’
 b. Nosotros los estudiantes tene-mos mala memoria.
 we (the students) have-1.PL bad memory
 ‘We the students have bad memory.’
- (28) a. *Ellos (los estudiantes) ten-emos mala memoria. (Spanish)
 they (the students) have-1.PL bad memory
 ‘We the students have bad memory.’
 b. Ellos (los estudiantes) tien-en mala memoria.
 they (the students) have-3.PL bad memory
 ‘We the students have bad memory.’

The mismatching possibilities observed in (23) reappear in infinitival contexts even with pronoun+DP combinations, as shown in (29), which is admittedly a marked construction.

- (29) Sugirieron [hablar nosotros los estudiantes] (Spanish)
 suggested.3.pl speak us the students
 ‘They suggested that we the students should speak.’

These additional facts in (26)-(28) suggest that in addition to the role of inflection, one needs to take into account the ϕ -feature settings of DPs vs. pronouns, and perhaps the hierarchy of ϕ -features (cf. Béjar, 2003 and Shlonsky, 1997), to account for the number differences.

4.3.2 Pronominal agreement and ellipsis

One potential argument against the pronominal analysis of AGR/INFL comes from a generalization on ellipsis presented in Saab (2009, 2010). He points out that interpretable features never trigger partial identity effects under ellipsis. When a verb is gapped, that verb's interpretable tense must be identical to its overt counterpart's. We can see an example of this in (30a), where the gapped verb is interpreted as having the same tense as the overt one (*han aprendido* 'have learned'). In (30b), on the other hand, a mismatching tense interpretation renders the example ungrammatical.

- (30) a. Los estudiantes han aprendido mucho chino y Antonio también ha aprendido
 the students have learned much Chinese and Antonio also has learned
 mucho chino. (Spanish)
 much Chinese
 'The students have learned a lot of Chinese and the Antonio too.'
- b. *Los estudiantes han aprendido mucho chino y Antonio en el futuro también
 the students have learned much Chinese and Antonio in the future also
 aprenderá mucho chino.
 will learn much Chinese

Uninterpretable gender, on the other hand, is not subject to an identity requirement, as seen in (31), where the overt participle is masculine and the elided one is feminine.

- (31) Juan fue localizad-o en el restaurante y Marta también fue localizad-a en el
 Juan was located-MASC in the restaurant and Marta too was located.FEM in the
 restaurante. (Spanish)
 restaurant
 'Juan was located at the the restaurant and Marta too.'

If inflectional ϕ -features (person and number) are interpretable like Tense, then they should preserve strict identity under deletion, the way Tense was shown to do in (30). However, this is not the case, as seen in (32). In this example, person and number on the overt verb are 3rd person singular, whereas the elided verb is interpreted as 1st person plural. To the extent that the generalization on identity of interpretable features holds, this is a strong argument against considering ϕ -features interpretable on the verb.⁸

- (32) Juan fue al cine y nosotros también fui-mos al cine. (Spanish)
 Juan went.3.SG to-the movies and we too went-1.PL to-the movies
 'Juan went to the movies and we did too.'

⁸As expected, strict ϕ -feature identity is not enforced either in Jelinek's examples above: (i) can be interpreted either as *... y hoy también los asistentes fuimos a la huelga* 'and today (we) assistants also went on strike' or as *... y hoy también los asistentes fueron a la huelga* 'and today also assistants went on strike'.

- (i.) Ayer, los estudiantes fuimos a la huelga, y hoy también los asistentes. (Spanish)
 yesterday the students went.1.PL to the strike and today also the assistants.
 'Yesterday students went on strike and today assistants did too.'

An interesting fact about Saab's generalization is that it doesn't immediately extend to *pro*. In the example above, both conjuncts have an overt subject (which is contrastive in the second conjunct). However, in (33a), with a NS in both conjuncts, it is not possible to have distinct ϕ -features. Thus, this example cannot be interpreted as (33b). Intuitively, what prevents the non-identical feature interpretation in (33a) is that the second conjunct's person/number features are unrecoverable. But even in examples like (34a), where the anaphor should provide enough information, the non-identical interpretation fails. The counterpart in which the subject is overt is marginal but clearly more acceptable than the one with *pro* (cf. (34b)).

- (33) a. #El sábado salimos y el domingo también salí. (Spanish)
 the Saturday went.1.PL out and the Sunday too went.1.SG out
 'Saturday we went out and Sunday too (I went to the movie)'
- b. El sábado salimos y el domingo también salí.
 the Saturday went.1.PL out and the Sunday too went.1.SG out
 'Saturday we went out and Sunday I went out too.'
- (34) a. *El sábado, Andrés se molestó consigo mismo, y el domingo, también me molesté conmigo mismo. (Spanish)
 the Saturday, Andres CL got upset with himself, and the Sunday, too CL
 got upset with myself
- b. ?El sábado, Andrés se molestó consigo mismo, y el domingo, yo también me molesté conmigo mismo.
 the Saturday, Andres CL got upset with himself, and the Sunday, I too
 CL got upset with myself
 Saturday, Andres got upset with himself and Sunday, I too, with myself'

If we take the identity-of-interpretable-features-in-ellipsis hypothesis seriously, the result that follows is that inflectional ϕ -features are interpretable only with a NS, but not with an overt one.

Notice that these paradigms are very reminiscent of the Ancona facts presented in sect. 4.2.1.4. In that case, a 3rd person sg. verb could be interpreted as default, that is, compatible with a 3rd person plural overt DP, but not with a NS. Suppose that in (33a), *pro* must obligatorily gap with the verb because it is necessarily attached to it (in Cardinaletti's (1997) analysis, because it is a weak pronoun), then what makes that example ungrammatical is the fact that *pro*'s interpretable features are not identical with those of its antecedent.

4.3.3 Arguments against the pronominal agreement hypothesis from Irish

Irish has two distinct types of unaccusative verbs, which McCloskey (1997) calls salient and putative unaccusatives illustrated in (35a-b) respectively. The most notable difference between them is that salient unaccusatives have a prepositional argument, which, according to McCloskey remains in a VP internal position. Putative unaccusatives, on the other hand, lack the preposition and the argument acts as an external argument.

- (35) a. Neartaigh ar a ghlór. (Irish)
 strengthened on his voice
 ‘His voice strengthened.’
- b. Neartaigh a ghlór.
 strengthened his voice
 ‘His voice strengthened.’ (from McCloskey, 1996, 251, 23)

The account that rich inflection satisfies the EPP by verb-raising does not extend well to these cases. On the one hand, the fact that the internal argument remains in situ in (35a) suggests that it does not satisfy the EPP. On the other, the verb is assumed to move to I, so it could satisfy the EPP. However, in Irish there are two types of verbal morphology, the so-called analytic forms and the synthetic forms (cf. McCloskey and Hale, 1984). The first type lacks ϕ -feature information, whereas the second has person and number content. NSs can only appear with synthetic forms, not with analytic ones. Even if one does not take into account morphological richness, the fact that synthetic forms allow for NSs and analytic ones do not would classify the first type as rich and the second one as poor in A & A’s terms. In that context, one would expect that only synthetic forms would satisfy the EPP, however, all salient unaccusatives must appear in the analytic form (cf. McCloskey, 1996, 272).

4.3.4 Arguments from Finnish expletives against the pronominal agreement hypothesis

Holmberg (2005, 537) points out that if AGR receives the subject theta-role, it is interpretable, hence referential, then it may plausibly also absorb nominative Case. If *pro* exists as an independent entity, it would have to be expletive (because no theta role is available), however, in Finnish, overt expletives are barred from NS constructions such as (36).⁹

- (36) *Sitä puhun englantia. (Finnish)
 exp speak.1.SG English (from Holmberg (2005, 543, ex. 21))

Holmberg (2005, 545) argues that the ungrammaticality of this example cannot be due to the fact that the expletive is redundant because AGR checks the EPP, since this would preclude sentences like (37a), where AGR should also check the EPP, and still the expletive is possible.

- (37) a. Sitä olen minäkin käynyt Pariisissa. (Finnish)
 EXP be.1.SG I-too visited Paris.INE
 ‘I have been to Paris, too (actually).’ (from Holmberg, 2005, 543, ex. 19)
- b. Olen minäkin käynyt Pariisissa.
 be.1.SG I-too visited Paris.INE
 ‘I have been to Paris, too (actually).’ (from Holmberg, 2005, 543, ex. 20)

⁹Finnish has productive NS in 1st and 2nd person, cf. Vainikka and Levy (1999).

He concludes that the contrast between (36) and (37) follows if AGR never checks the EPP. In (36), *pro* does, hence the expletive is redundant and in (37a), the expletive checks it (because the pronoun is too low to do so). By contrast, in (37b), the pronoun is in Spec, IP, it checks the EPP and the auxiliary subsequently raises to C.

4.3.5 Word-order asymmetries in Agreement

All other things being equal, if a language has alternative SV and VS word orders, the pronominal agreement hypothesis predicts that there should be no agreement asymmetries, since agreement remains constant in both configurations. Thus, *a priori*, the data presented earlier from the Ancona variety of Italian (cf. (15)-(17)) is surprising. These data are replicated in several other Northern Italian dialects (cf. Manzini and Savoia, 2002). Recall that in the order VS, default singular agreement is possible, but in the order SV, full plural agreement is obligatory. Furthermore, the referential properties seem to be located on the overt DP, not on inflection. Thus, in both (38a and b), subject reference is 3rd person plural, even though in (38a), inflection is 3rd person singular.¹⁰

- (38) a. Questo, lo fa sempre i bambini. (Ancona variety)
 this.ACC CL do.3.SG always the.PL children.PL
 'The children always do this.'
- b. *Questo, i bambini lo fa sempre.
 this.ACC the.PL children.PL CL do.3.SG always
- c. Questo, i bambini lo fanno sempre.
 this.ACC the.PL children.PL CL do.3.PL always
 'The children always do this.' (from Cardinaletti, 1997, 39, ex. 21)

If referential properties are located on inflection, one would expect variation in reference between (38a-c), given the difference in ϕ -features. However, this does not seem to be the case.

To summarize this section, I have presented the pronominal agreement hypothesis through the implementation proposed by A & A. This hypothesis assumes that the referential and/or thematic properties of an argument can be carried by the inflectional morphology of a verb, and certain distributional patterns of coreference taken from Warlpiri and Spanish fit well within this hypothesis. I have also presented several cases that seem to challenge this account, namely the generalization that only uninterpretable features can have different values in ellipsis (but inflection can). Additionally, the Irish analytic/synthetic distinction, in conjunction with the distribution of salient unaccusatives argues against the notion that rich-agreement verbs can satisfy the EPP by raising to I. Finally, I have pointed out that the pronominal agreement hypothesis has difficulties explaining differences in verb agreement triggered by changing word orders (Ancona dialect), even though referential possibilities remain constant across examples.

¹⁰In a sense, these data are the mirror image of the Spanish examples in (23) above, from Jelinek, in which the referential properties seem located on I.

4.4 Deletion

In this section, I turn to the third answer to the question of what satisfies the EPP. This answer assumes that *pro* as not an independent category, but a pronoun without overt phonological content. Within this general idea, we find at least two separate analyses. The first one is Holmberg's (2005) account of NSs in Finnish. The second one is Saab (2009, 2010), an account that extends otherwise well-attested mechanisms of ellipsis to the case of NSs. Finally, we will briefly recapitulate Tomioka's proposal regarding discourse NSLs, which also involves NP ellipsis.

4.4.1 Finnish expletives and NSs

Holmberg's analysis of the distribution of overt expletives in Finnish leads him to propose that NSs are the null counterparts of overt pronouns. Specifically, he advances a typology of null subjects with distinct levels of internal structure, as represented in (39). This typology corresponds roughly to Déchaine and Wiltschko's (2002) proposal for pronouns. In Holmberg's account, only null DPs are fully referential, since referentiality is linked to D.

- (39)
- a. D-less ϕ Ps (Null bound pronouns and generic pronouns in partial NSLs like Finnish and also Null subjects in Spanish, Greek or Italian)
 - b. Null DPS (1st and 2nd person subjects in Finnish)
 - c. Null NPs (NS in Chinese/Japanese)

Holmberg considers and rejects the pronominal agreement alternative, based on the following facts about Finnish. In this language, expletives have a limited distribution, for example, they can optionally appear in impersonal sentences like (40a), usually when nothing else can appear in first position. Holmberg assumes this is due to an EPP requirement. If the sentence only has a verb, or a verb and some other category that cannot check the EPP, then the expletive is optional, as in (40b).

- (40)
- a. Sitä meni hullusti. (Finnish)
 EXP went wrong
 'Things went wrong.' (from Holmberg, 2005, 543, ex. 20)
 - b. Meni hullusti.
 went wrong
 'Things went wrong.' (from Holmberg, 2005, 542, ex. 17)

However, if another constituent that can check the EPP appears, either that constituent moves to first position, or the expletive is obligatory, as seen in (41).¹¹

¹¹In Finnish, arguments and certain adverbs (temporal, locative or instrumental) can check the EPP (i.e. move to preverbal position), whereas other adverbs (reason or manner) cannot do so. Holmberg (2005, 542) calls the second type "potential topics". Compare the grammaticality of (40b) with the ungrammaticality of (i), without adverbial movement and the grammaticality of (ii) with adverbial movement.

- (41) a. *Meni nyt hullusti. (Finnish)
 went now wrong
 b. Nyt meni hullusti.
 now went wrong
 ‘Now things went wrong.’
 c. Sitä meni nyt hullusti.
 EXP went now wrong
 ‘Now things went wrong.’ (from Holmberg, 2005, 541, ex. 13)

These examples show that Finnish has an overt expletive that surfaces in the expected position in order to satisfy the EPP, which is conceived as a requirement that the Spec of IP be filled. With this background in mind, suppose that we assume the pronominal agreement hypothesis for 1st and 2nd person, since Finnish only has productive NSs in these two persons. Furthermore, if we assume that pronominal agreement checks the EPP, this predicts that expletives should not be possible in those persons, and this prediction is true for certain contexts (cf. (42a)) but not for others (cf. (42b)). Given these facts, the pronominal agreement would be forced to claim either that inflection doesn’t always satisfy the EPP, or that the EPP does not apply in all cases.

- (42) a. *Sitä puhun englantia. (Finnish)
 EXP speak-1.SG English (from Holmberg, 2005, 543, ex 21a)
 b. Sitä olen minäkin käynyt Pariisissa.
 EXP be.1.SG I-too visited Paris-INE
 ‘I have been to Paris, too (actually).’ (from Holmberg, 2005, 545, ex 19b)

As an alternative, Holmberg (2005) proposes that Finnish has a null pronominal that checks the EPP, so the expletive is unnecessary in (42a). In (42b), on the other hand, the overt pronoun does not check the EPP, presumably because it is lower than Spec, IP. The counterpart of (42b) without an expletive is also grammatical (cf. (43)), and in this case, the subject merges with IP, checks the EPP and the auxiliary incorporates to C. In other words, the difference between (42b) and (43) is structural: In the first example, the verb and the subject are lower than in the second one.

- (43) Olen minäkin käynyt Pariisissa.
 be.1.SG I-too visited Paris-INE
 ‘I have been to Paris, too.’ (from Holmberg, 2005, 545, ex. 24a)

As mentioned earlier, Holmberg assumes that finite 1st and 2nd person *pros* have the representation in (39b) above, that is, they are fully referential DPs. By contrast, generic NSs like (44) below, which cannot satisfy the EPP, have the representation in (39a).

-
- (i.) *Leikkii lapsia kadulla. (Finnish)
 play children in-street
 (ii.) Kadulla leikkii lapsia.
 in-street play children
 ‘Children are playing in the street.’ (from Holmberg and Nikkane, 2002, exs. 1c, 2b)

- (44) Tässä istuu mukavasti. (Finnish)
 here sit.3.SG comfortably
 ‘One can sit comfortably here.’

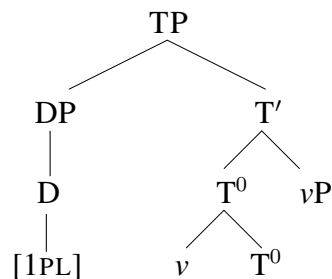
NSs in consistent NSLs have the same representation as generics in Finnish. In this case, Holmberg assumes that these languages, unlike Finnish, have a D feature in I, which is independent of AGR. When I and a ϕ P *pro* agree in a language like Greek or Spanish, the resulting category is [D ϕ P], which makes it referential.

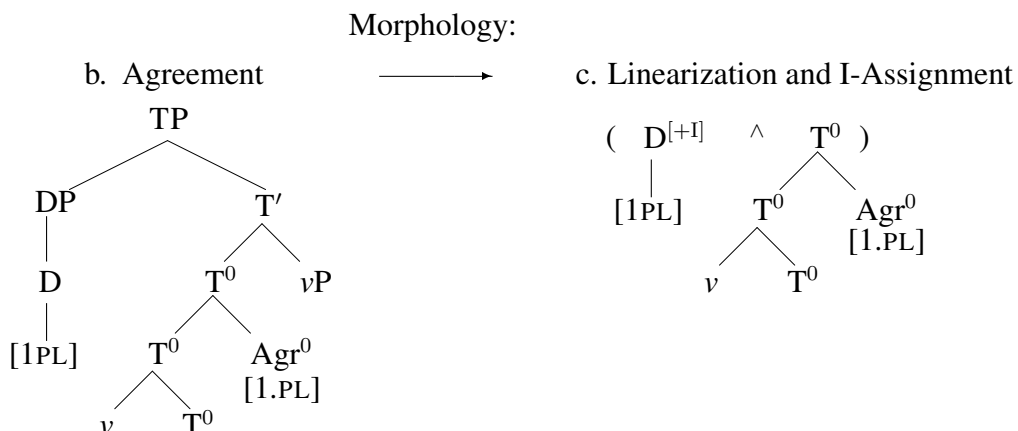
While Holmberg’s account attempts to subsume *pro* into the class of overt pronouns assuming that they are the deleted counterpart of those pronouns, it is not clear to what extent this is explicitly done. On the one hand, the precise mechanisms for deletion of *pro*’s phonological content are not made clear, on the other hand, one of the predictions this theory makes is that the types of *pro* we find should be at most the same types of overt pronouns, and at least a subset of those types. In practice, both Holmberg’s and Saab’s account (see below) restrict *pro* to the subset of weak pronouns, following Cardinaletti and Starke (1999). One lingering question is why this should be the case (i.e. why don’t we have the null counterpart of overt, non-emphatic pronouns).

4.4.2 *Pro* as ellipsis

Saab (2009), Tomioka (2003) and Barbosa (2010) all develop different versions of the idea that NSs are subjects under ellipsis. Saab (2009, 2010) develops the idea of NSs as deleted versions of overt pronouns within a much larger theory of ellipsis. Working in the framework of Distributed Morphology, he takes NSs to be deleted as a PF phenomenon only when they are adjacent to T. Formally, a (simplified) syntactic structure like (45a) is linearized as (45b), and D can be assigned a feature [+I]. This feature is added to elliptical heads under formal identity, and blocks lexical insertion rules, resulting in a null subject. Thus, the only difference between an overt pronoun and a null one is that the former lacks [+I] and gets content through lexical insertion rules, whereas the latter has [+I], which blocks the application of those rules.

- (45) a. Syntax of *compramos* ‘bought.1.PL’





Since [+I] is assigned to subjects adjacent to T, for all intents and purposes, NS are also the null equivalent of Cardinaletti and Starke's (1999) weak pronouns.¹²

4.4.3 Null/overt asymmetries

Both Holmberg's and Saab's deletion analysis of *pro* rely on Cardinaletti & Starke's (1999) assertion that *pro* is a weak pronoun. This raises the two following questions regarding the null/overt distinction. First, why is *pro* necessarily weak (at least in Spanish and Italian)? Second, why must it necessarily occupy a dedicated (Spec, IP) position?

To understand the relevance of the first question, consider that all languages that have *pro* also have overt pronouns, so it is puzzling that only weak pronouns can be null. There is a trivial answer to this question: Phonological nullness is the weakest possible manifestation, so in some sense, null pronouns simply deplete already deficient content. However, the notion of weak pronoun also encodes semantic correlates, and in this sense, it is not clear why both things should go together.¹³

On the other hand, deletion is generally not concerned with the size of the deleted element, which can be as large and structurally complex as a whole IP in ellipsis contexts (cf. (46)), so the intuitive notion that NSs are null because they are syntactically simpler therefore phonologically smaller does not find a clear parallel in other areas where ellipsis is well defined.

- (46) Doris compró una tartaleta de limón en la pastelería de San Antonio a las 4 ayer
 Doris bought a tart of lemon in the pastry-shop of San Antonio at the 4 yesterday
 y Samuel también compró una tartaleta de limón en la pastelería de San Antonio a
 and Samuel also bought a tart of lemon in the pastry-shop of San Antonio at

¹²Cardinaletti & Starke's distinction between weak and strong pronouns relies on a set of interpretive and prosodic properties. For example, weak pronouns cannot be modified by adverbs like *only*, or conjoined. Cf. sect. 7.1.

¹³In Cardinaletti and Starke's analysis, pronominal deficiency stems from having fewer syntactic/functional projections, and this correlates with having less phonological content. As they put it: [T]he more a pronoun is deficient, the less it has syntactic structure." (fn. 50). However, it should be clear that fewer syntactic/functional heads does not always mean less prosodic content, and viceversa, less phonological content does not necessarily mean less syntactic complexity as Bayer (1999, 237) points out.

las 4 ayer (Spanish)
 the 4 yesterday
 Doris bought a lemon tart in the San Antonio pastry shop at 4 yesterday and Samuel too.'

A third point of contention is the following: If most of the semantic properties associated with strong pronouns have as a prerequisite some kind of prosodic prominence, then the reason why NSs do not have those properties is precisely because they cannot be prosodically prominent. In fact, Frascarelli (2007) explicitly argues for deriving the weak/strong distinction from the prosody of pronouns.

Cardinaletti and Starke (1999, 164) make a point of rejecting this alternative, noting that certain weak pronouns can bear stress (cf. for example clitics in Argentinean Spanish, Huidobro, 2005), and that not all strong pronouns must be prosodically prominent (cf. French *Jean a vu seulement lui* 'Jean has seen only him', in which the pronoun is prosodically flat, according to Cardinaletti and Starke, 1999, 162). Additionally, they argue, if one were to derive prosodic weakness/strength separately from the semantic properties associated with those properties, one would miss the generalization that they tend to appear together. However, Cardinaletti and Starke (1999) themselves point to instances in which changes in semantic properties are not accompanied by changes in prosodic ones. For example, they suggest that weak pronouns can usually not be used to point at something in the discourse, but if there is a salient antecedent in the discourse, then ostension becomes possible with weak pronouns. In other words, a weak pronoun can remain prosodically weak but show one of the properties of strong pronouns.

On the other hand, at least one of the properties they ascribe to strong pronouns (conjoinability) has no obvious correlate in the semantic side. Furthermore, they explain the restriction on coordination by stating that it requires CP/DP and since weak pronouns do not project CP, they cannot be coordinated. By contrast, strong pronouns project a CP and can be conjoined. This leaves open the question of why overt verbal IPs are typically coordinated in many languages. Additionally, in languages like Irish, *pro* can be conjoined (cf. (47)).¹⁴

(47) da mbeinn -se agus tusa ann. (Irish)
 if be.COND.1.SG -CONTR and you there
 'If you and I were there.' (from McCloskey and Hale, 1984, 501, ex. 31a)

The data from Irish challenge the analysis of *pro* as a weak pronoun in general, because Irish *pro* lacks at least two of the properties of weak pronouns (it can be conjoined and it can appear with contrastive modifiers), but it is still null. To the extent that *pro* in Irish is necessarily linked to rich

¹⁴In Irish, conjoined pronouns whether overt or null, have an added emphatic particle, as shown in (i)-(ii).

- (i.) *mé agus tú (Irish)
 I and you (from McCloskey and Hale, 1984, 503, 35c)
- (ii.) mise agus tusa
 I.CONTR and you.CONTR
 'You and I.' (from McCloskey and Hale, 1984, 503, 36a)

agreement (cf. McCloskey and Hale, 1984), one would assume that it occupies the same position as *pro* in other languages, but without the corresponding properties.

For languages like Spanish that disallow conjunction of *pro*, one can independently show that prosodic status determines coordination possibilities, as in the case of adverbs in *-mente* in Spanish (cf. (48)). In this example, two bound morphemes can be conjoined, and they happen to behave as two separate prosodic words (for example, they have independent stress, cf. Bosque (1987) and Camacho (2003, 67)). In other instances of bound morphemes, lack of conjoinability is associated with lack of prosodic independence. Extending the CP/IP analysis to these cases seems rather difficult.

- (48) a. La boda más [ceremoniosa- y solemne-]mente celebrada. (Spanish)
 the wedding most ceremonious- and solemnly celebrated
 ‘The most ceremoniously and solemnly celebrated wedding.’
- b. La boda más [ceremoniosamente y solemnemente celebrada.
 the wedding most ceremoniously and solemnly celebrated
 ‘The most ceremoniously and solemnly celebrated wedding.’

(from Camacho, 2003, 67)

To summarize, the deletion analysis faces the challenge of explaining why only weak pronouns can be deleted in NSLs. As an alternative account, it is possible that part of the distribution of *pro* comes from the fact that it is null, not from the contention that it is weak. I have also reviewed some evidence from Irish that not all instances of *pro* can be treated as weak.

4.5 Is the EPP universal?

In sect. 4.5.2 and in the past sections of this chapter, I have assumed that the EPP is universal. The question we will address in this section is whether the EPP can be maintained as a unified requirement that some nominal category appear in Spec, IP. Part of the answer to this question hinges crucially on the status of null expletives. After all, if null thematic subjects are independently required by the θ -Criterion, the sole remaining reason for postulating expletives would be that Spec, IP needs to be filled.¹⁵ But if the EPP is not universal, then there is no need to postulate the existence of null expletives in NSLs. If that is correct, then the presence of overt expletives in languages that have them would have a different explanation not directly related to argument structure.

4.5.1 Evidence for expletives in NSLs

For obvious reasons, empirical evidence for or against pure null expletives is very hard to find. Despite that difficulty, McCloskey (1996), has argued that Irish grammar shows no evidence of

¹⁵Manzini and Savoia (2002, 162) point out that if one treats theta roles as independent features that be merged with the corresponding argument, the need for thematic *pro* as bearer of a theta role disappears. However, proposing independent feature theta roles is not enough, one still needs to have something to associate that feature with, be it inflection or something else.

null expletives, and that the EPP does not hold in that language. First, Irish is consistently VSO in finite clauses, but McCloskey (1991, 1997) argue that the subject raises from its vP position to a position lower than IP. To reach this conclusion, he points out that subjects of finite clauses consistently place to the left of adverbs (cf. (49)). At the same time, VP-ellipsis requires the subject to obligatorily elide (cf. (50)), suggesting that the subject and other VP-material form a constituent that excludes the verb (cf. (51)).

- (49) Chuala Róise go minic roimhe an t-amhrán sin. (Irish)
 heard Róise often before-it that-song
 ‘Róise had often heard that song before.’ (from McCloskey, 1997, 219, ex. 50b)
- (50) Ní tháinig muid ’na bhaile anuraidh ach tiocfaidh [muid ’na bhaile]
 NEG came we home last-year but come.FUT we home
 i mbliana. (Irish)
 this-year
 ‘We didn’t come home last year but we will this year.’ (from McCloskey, 1997, 211, ex. 23)
- (51) [IP V+I [XP Subj [vP Subj [VP V O . . .]]]]

This suggests that subjects do not overtly occupy the canonical Spec, IP subject position. Is this position occupied by an expletive? McCloskey believes that it is not. As part of his evidence, he presents the asymmetric behavior of unaccusative verbs in Irish. For a subclass of unaccusative verbs in this language (his salient unaccusatives), the sole argument systematically remains inside VP, whereas for another subclass (putative unaccusatives), the argument is external by all the relevant tests. The most notable difference between both classes is that the internal argument of salient unaccusatives has a preposition, as seen in (52a) vs. b).

- (52) a. Neartaigh ar a ghlór. (Irish)
 strengthened on his voice
 ‘His voice strengthened.’
- b. Neartaigh a ghlór.
 strengthened his voice
 ‘His voice strengthened.’ (from McCloskey, 1996, 251, 23)

If Irish has a null expletive to comply with the EPP, and expletive-associate chains must match in syntactic category, the existence of salient unaccusatives, with a PP argument becomes problematic. As McCloskey (1997) points out, the same argument can be recreated for perfective passives. On the other hand, if Irish lacks expletives altogether, the existence of salient unaccusatives (and the perfective passive patterns) no longer becomes an issue.

For languages like Spanish or Italian, empirical evidence of null expletives is also very hard to come by. On the one hand, unlike thematic *pro*, expletives never alternate between overt and null forms, so it is hard to analyze the distribution of null expletives based on the distribution of overt ones. Even in varieties that seem to have overt expletives (cf. sects. 3.2.1.1 and 3.2.1.2),

the evidence suggests that they are not uniquely related to the subject position, but potentially to a left-peripheral CP position.

Furthermore, I think one can find indirect evidence against null expletives. Consider, for example, extraposed clauses with verbs such as *resultar* ‘turn out’, shown in (53a). These clauses cannot be preposed, as seen in (53b). If the EPP holds in Spanish, one interpretation of this fact suggests that these clauses can never satisfy the EPP.

- (53) a. Resultó [que Blanca va a venir] (Spanish)
 turned out that Blanca is going to come
 ‘It turned out that Blanca is going to come.’
 b. ?*[Que Blanca va a venir] resultó.
 that Blanca is going to turned.out
 ‘That Blanca is going to come turned out.’

Other verbs that take extraposed clauses do allow for the clause to raise to a preverbal position, as seen in (54), although this is not their preferred position. In this position, the extraposed clause tends to be interpreted as focused. The same pattern holds in English, but with the overt expletive *it*.

- (54) a. Es cierto [que Blanca va a venir] (Spanish)
 is true that Blanca is going to come
 ‘It is true that Blanca is going to come.’
 b. [Que Blanca va a venir] es cierto.
 that Blanca is going to come is true
 ‘That Blanca is going to come is true.’

As in the case of Irish unaccusative verbs, it is not clear why one type of predicate (*ser cierto* ‘be true’) allows for the extraposed clause to satisfy the EPP whereas another type (*resultar* ‘result’) does not.¹⁶ Nevertheless, assume, for the sake of argument, that a null expletive appears obligatorily with *resultar* ‘turn out’ (hence raising is impossible in (53b)), but optionally with *ser cierto* ‘be true’, hence the optional raising of the clause in (54). However, consider two more cases in the paradigm presented in (55)-(56): When an adjective or adverb is added as a predicate of the main verb, then preposing the clause becomes possible with *resultar* ‘result’.¹⁷

¹⁶If one assumes that inflection can satisfy the EPP (cf. 4.3 below), then clausal raising must be unrelated to the EPP in (53)-(54), since in both cases inflection satisfies the EPP.

¹⁷Without a predicate, *ser* shows the same pattern as in (53). For example, if I ask ¿Por qué estás tan contento? ‘why are you so happy?’ someone could reply (i) but not (ii).

- (i.) Es [que Blanca va a venir]
 is that Blanca is going to come
 ‘It’s that Blanca is going to come.’
 (ii.) *[Que Blanca va a venir] es.
 is that Blanca is going to come

- (55) a. Resultó excelente [que Luis viniera] (Spanish)
 turned.out excellent that Luis come.SUBJ
 ‘It turned out great that Luis came.’
 b. [Que Luis viniera] resultó excelente.
 that Luis come.SUBJ turned.out excellent
 ‘That Luis came turned out (to be) great.’
- (56) a. Es bueno [que Luis venga] (Spanish)
 is good that Luis come.SUBJ
 ‘It is good that Luis came.’
 b. [Que Luis venga] es bueno.
 that Luis come.SUBJ is good
 ‘That Luis comes is good.’

Given that the only difference between (53)-(54) and (55)-(56) is the presence of the adverb or adjective, it would seem that the possibility of preposing the extraposed clause is not related to the EPP. In fact, if anything, this contrast suggests a VP-internal EPP effect: If no adverb is present, the clause must remain in a VP-internal position, but if the adverb/adjective is present, the clause may move or remain in situ.

Existential constructions with *haber* ‘be’ also disallow the overt DP from appearing in pre-verbal position both in Spanish and in English, as seen in (57)-(58). Even if an overt locative appears, the associate-V order is ungrammatical in Spanish and in English it lacks an existential, non-specific reading.¹⁸

- (57) a. Habían tres personas. (Spanish)
 were.3.PL three people
 ‘There were three people.’
 b. *Tres personas habían.
 three people were.3.PL
- (58) a. There were three people. (English)
 b. *Three people were.
- (59) a. Tres personas habían en la conferencia. (Spanish)
 were.3.PL three people in the conference
 ‘There were three people at the conference.’
 b. #Three people were at the conference. (English)

Needless to say, there are multiple analyses of existential constructions. The main point here is that these contrasts raise issues about the universality of the EPP as a requirement to have a category in

¹⁸There are questions about whether the associate is an external argument. Rodríguez-Mondoñedo (2006) argues that it is an object, although it agrees with the verb.

Spec, IP at least for languages like Spanish/Irish, where they do not interact with overt expletives and the syntactic distribution does not provide any independent evidence.

To summarize, the proposal that a null expletive fulfills the EPP requirement (as a requirement of Spec, IP) does not provide much explanation for the relative position of extraposed clauses, hence it questions whether the EPP is a universal requirement, in line with McCloskey (1997).¹⁹ However, there are plenty of “EPP-effects” described in the literature. In most cases, these remain a poorly understood cover term for situations in which some constituent must move to a clause-initial position, and perhaps these cases are best seen as phonological alignment constraints, in other cases, they may be the result of interaction between assertion structure and prosody.

4.5.2 Positional freezing

In this section I present a slight detour on an alternative perspective for the correlation between VS word order and *that*-trace effects. This alternative is based on Rizzi and Shlonsky (2007) (henceforth R & S). Although this is not directly relevant to the NSP, since I have shown that NSs are not directly connected with *that*-trace violations, this detour allows me to explain some of the facts presented in section 3.2.4.

R & S present a different perspective on the EPP. Their proposal rests on the idea that the SubjP requires a subject-like category in its specifier in order to satisfy the Subject Criterion. Whatever category satisfies the criterion is frozen in that position, preventing it from further movement. The most important consequence of this freezing is that subjects cannot be extracted (if frozen in SubjP), giving rise to subject/object extraction asymmetries. In order to extract the subject, languages resort to alternative strategies, for example inflected complementizers. Another possibility is to satisfy the Subject Criterion with a category different than the subject (an expletive, among other possibilities, see below).

Specifically, one possible by-passing strategy involves merging a nominal category with ϕ -features above the canonical IP subject position. The resulting configuration satisfies the Subject Criterion, according to R & S. Two consequences follow: the complementizer shows some kind of agreement properties and subject extraction becomes possible.

For French, R&S propose a representation along the lines of (60). An agreement-bearing complementizer (*que*) merges with Subj', producing a structure in which SubjP lacks an overt complementizer position, and where the Subject Criterion is satisfied by the agreeing complementizer (Fin + ϕ in (60)). This, in turn, allows for the subject to raise. Similar results obtain in English.²⁰

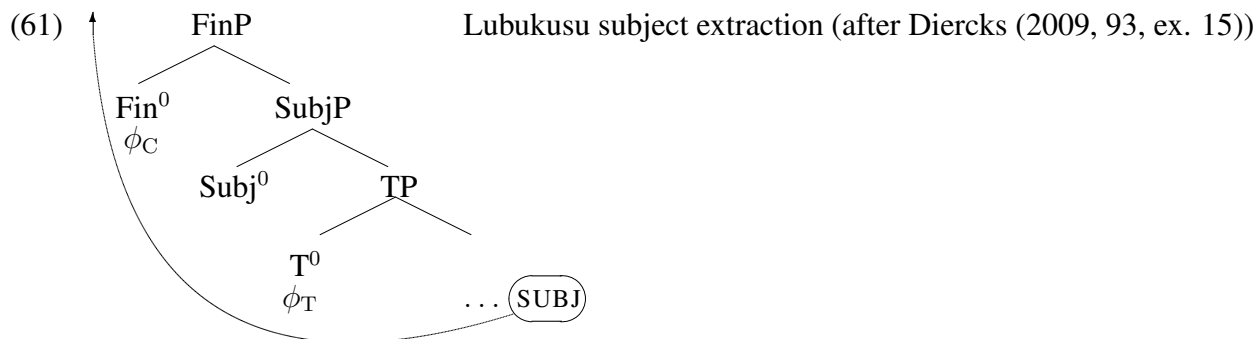
- (60) Subj. extraction in French:
- $$[\text{FocP wh-} [\text{Foc} [\text{FinP wh-} [\text{Fin}+\phi [\text{Subj} \dots [\text{vP wh-} \dots]]]]] \quad \text{Fin}+\phi = \textit{qui}$$
- ↑-----↓-----

Following these ideas, Diercks (2009) proposes the following account for the Lubukusu paradigms presented in sect. 3.2.4.5: Subject extraction in this language requires merging of an agreement

¹⁹Since much of the distribution just discussed can be replicated in English, where an overt *it* does appear, the logical conclusion is that the EPP does not explain much of the English distribution either.

²⁰Both (60) and (61) below assume Rizzi's expanded structure for the left periphery, which includes, among others, a FocP, FinP, cf. Rizzi (1997).

morpheme above Subj to satisfy the Subject Criterion, as in French and English. As a result, complementizer agreement becomes obligatory whenever subject extraction takes place, as in (61).



The extraction facts from BP in sect. 3.2.4.4 poses a slightly more challenging problem for extending the freezing analysis. On the one hand, the data from inflected infinitives (cf. (56)) seem to follow a similar pattern as the one in Lubukusu: Whenever agreement is merged above Subj, extraction is licensed. On the other, if we adopt Menuzzi's (2000) proposal as is, extraction must take place from the preverbal subject position. Since this position would freeze the subject, according to R&S's analysis, one would not expect extraction to be possible. However, I think there are reasons to argue for a slight reanalysis of the data. Menuzzi (2000) assumes that *nada* 'nothing' is adjoined to I' because it precedes the inflected verb but follows the subject (cf. (62)), and the subject is in Spec, IP.

- (62) a. João nada tinha feito ajudar Maria (BP)
 João nothing has done help Maria
 'João has done nothing to help Maria.'
 b. *João tinha nada feito ajudar Maria
 João has nothing done help Maria
 'João has done nothing to help Maria.' (from Menuzzi, 2000, 34, ex. 37)

However, a different analysis is possible: If *nada* is dislocated, then arguably the subject is not in its SubjP position. In an articulated left periphery (cf. Rizzi, 1997 and much subsequent work), this would involve the structure in (63).²¹

- (63) [_{TopP} João [_{TopP} nada [_{SubjP} [_{TP} tinha [feito ajudar Maria]]]]]

In the case of the inflected infinitive clauses, the verb must usually precede the subject (cf. (64)), a fact that Menuzzi (2000) takes to show V-to-C movement. However, whenever the object *nada* is dislocated, the verb can precede or follow the subject-object sequence, as in seen in (65)).

- (64) a. *O Manuel recorda os rapazes terem dado um presente pr'a
 the Manuel remembers the boys to-have.3.PL given a present to-the
 Maria. (BP)
 Maria
 'Manuel remembers that the boys gave a present to Maria.'

²¹ Alternatively, *nada* could be in Spec, SubjP, but I will not pursue that option.

- b. O Manuel recorda [terem os rapazes dado um presente pr'a Maria.
the Manuel remembers to-have.3.PL the children given a present to-the Maria
'Manuel remembers that the boys gave a present to Maria.' (from Menuzzi, 2000, 32,
ex. 34)
- (65) a. João lamentou os rapazes nada terem feito para ajudar Maria (BP)
João regretted the boys nothing to-have.3PL to done help Maria
'João regretted that the boys did nothing to help Maria.'
- b. *João lamentou os rapazes terem nada feito, para ajudar Maria
João regretted the boys to-have.3PL nothing done to help Maria
'João regretted that the boys did nothing to help Maria.'
- c. ?João lamentou terem os rapazes nada, feito para ajudar Maria
João regretted to-have.3PL the boys nothing done to help Maria
'João regretted that the boys did nothing to help Maria.' (from Menuzzi, 2000, 34,
ex. 38)

One possible account of these facts would have the verb in a constant position in all of these examples (for example, Rizzi's Fin^0 position), and the two arguments in different, dislocated positions above or below $FinP$, as in (66).²²

- (66) a. ... [$FinP$ terem [$TopP$ João [$TopP$ nada [$SubjP$ [TP feito ajudar Maria]]]]]
b. ... [$TopP$ João [$TopP$ nada [$FinP$ terem [$SubjP$ [TP feito ajudar Maria]]]]]

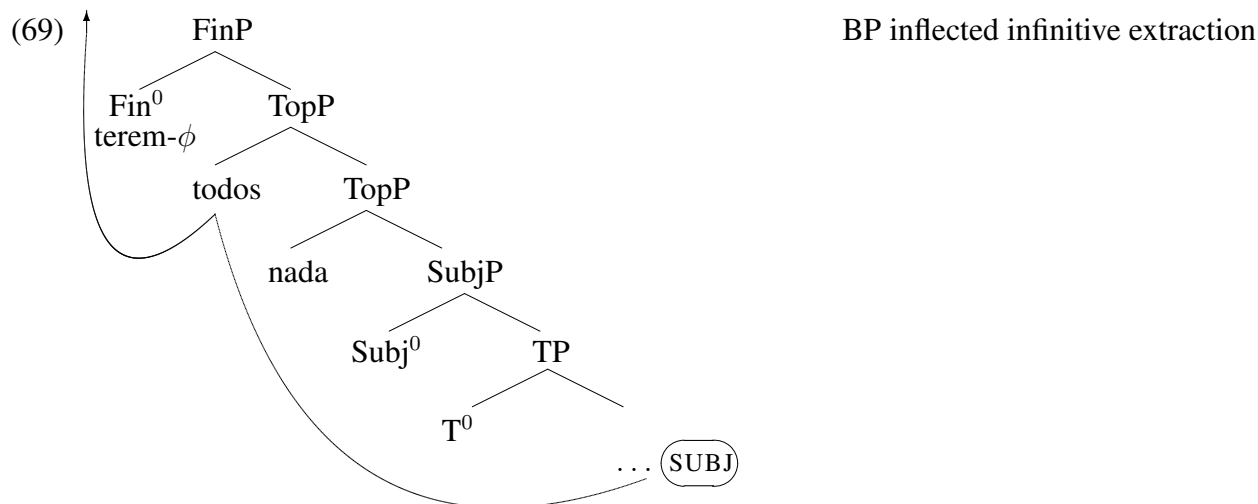
Notice, though, that this means that the presence of the floated quantifier *tudo* 'all' in (56) above, repeated below, no longer indicates that the extraction site is Spec, IP (or, in more precise terms, Spec, $SubjP$), but another topicalized position.

- (67) Que rapazes, o Manuel afirma terem todos nada oferecido de presente pr'a
which boys the Manuel claims to-have.3.PL all nothing offered of present to-the
Maria no aniversário dela? (BP)
Maria in her birthday
'Which boys does Manuel claim to all have offered nothing as a gift to Maria for her
birthday?' (from Menuzzi, 2000, 35, ex. 40a)

This, in turn, raises the question of what satisfies the Subject Criterion, and two possible answers come to mind: One, it is still the verb+ ϕ in Fin^0 , the other one, it is a null *pro*. Given the ungrammaticality of (68), where presumably *pro* should be able to appear in $SubjP$, I adopt the first option. Crucially, the difference between (67) and (68) is that in the first case there is an inflected infinitive but in the second one there is *que*. The proposed structure for the first example is presented in (69).

²²Merging Rizzi's (1997) left-periphery proposal with R&S $SubjP$ proposal would result in the clausal template $TopP > FocP > TopP > FinP > SubjP > TP$.

- (68) Que rapazes, o Paulo desconfia [que gostem **todos** de Maria]? (BP)
 which boys the Paulo suspects that like all of Maria
 ‘Which boys does Paulo suspect suspect all like Maria?’



4.5.3 The Subject Criterion revisited

R&S’s core proposal raises the following question: If SubjP is the position for the Subject Criterion, why would merging ϕ -features above that position satisfy the criterion?

The alternative I will propose is that the Subject Criterion may target a number projections in the left periphery, which can vary depending on the language. Its default position will be SubjP, but it can also target positions higher in the structure. Thus, in Icelandic and in DSEC, expletives appear in Spec, TP (what corresponds to R&S’s SubjP), whereas in CCS and Northwestern Iberian languages, the expletive has the feature Top, merging directly in that position. In BP, English/French, Övdalian and Lubukusu, merging of an agreeing head over SubjP, will satisfy the Subject Criterion and allow for subject movement.

Conceptually, formalizing the EPP as a requirement that targets a set of functional projections in the left periphery makes sense of the two independent properties associated with subjects. According to R&S (p. 118), the DP in SubjP “receives an interpretation paraphrasable as ‘About DP, I’m reporting event XP’. Subjects thus share an interpretive property of topics, the ‘aboutness’ relation linking subjects and predicates as well as topics and comments.” Thus, subjects are frequently aboutness topics in addition to subjects of a predication. These are independent properties so one expects them to be able to surface independently, and many researchers have argued that in certain languages one of them has more prominence than the other (cf. for example Li and Thompson, 1976). Thus, separating each of them in different heads is conceptually attractive.

Specifically, I propose the constraints in (70)-(71). These requirements are alignment requirements and can be ranked with respect to each other.²³ Languages for which (70) is more prominent

²³(70) and (71) differ from R & S’s proposal in at two respects: First, they are seen as alignment constraints,

Objects are never marked for Case or topicality, as seen in (74). Given the preceding discourse, the object is a perfect candidate for topic, but it is not marked as such.

- (74) [After *Intala takka-á magaalaa dhuf-e*. ‘A girl came to the market’.]
 Nama tokko-ó intala arke. (Oromo)
 person one.MASC-SUBJ girl saw
 ‘A man saw the girl.’ (from Clamons et al., 1999, 62, ex. 3)

Clamons et al. (1999, 62) point out that topic marking in (72) is distinct from topicalization, a process that fronts any NP, and marks it with a different morphology than the one illustrated in (72). This is seen in (75).

- (75) a. Yo Salmaa-ti, (isi-f-n) dhuf-t-e. (Oromo)
 as for-S (she-SUBJ-TOP) COMe-FEM-PAST
 ‘As for Salma, she came.’
 b. Yo Salmaa-ti, gurbaa-n isi ark-e.
 as for-S boy:SUBJ-TOP her see-PAST
 ‘As for Salma, the boy saw her.’ (from Clamons et al., 1999, 62, ex. 4)

These examples suggest that only subjects marked as topics agree with the verb. If we distribute the Subject Criterion requirement over several heads, we have a principled way of accounting for this pattern: Only subjects that are marked as topics satisfy this requirement in Oromo, and hence trigger agreement. If topic and subjecthood are conflated as part of the Subject Criterion, then the facts from Oromo cannot be properly explained.

In sum, I assume that two interacting constraints underly the Subject Criterion: Subject alignment and topic alignment. In languages like English, the requirement applies to SubjP, and in languages like Chinese, plausibly to TopicP. In languages like Oromo, we see overt evidence of targeting TopicP in that only subjects marked as topics can agree with the verb.

This requirement is usually satisfied by the subject, and when it is, the subject is frozen in place. However, there are other ways to satisfy the requirement for example by merging an expletive with SubjP, which frees the subject for extraction. In other cases, a complementizer that has agreement properties targets a left-peripheral position, also allowing movement of the subject.

As we will see in chapter 5 (particularly in sect. 5.5) the notion of the extended Subject Criterion will play an important role in accounting for why apparently very similar varieties have differential outcomes when licensing NSs.

4.6 The revised NSP

So, what is the NSP? Based on the evidence presented in previous sections, a successful analysis of NSLs and the NSP should at least take into account the following properties.

1. **Typology of NSLs.** Certain languages have productive null subjects, others have restricted null subjects, others only have overt subjects.

2. NSs and agreement

- a) The inflectional properties of a verb frequently correlates with the availability of null subjects, both language internally (i.e. Irish, Hebrew) and across languages (English vs. Spanish).
 - b. There is no universally valid notion of agreement-richness that serves as a sufficient condition for NS.
 - c. In some languages, NSs are identified in discourse.
3. **Null expletives.** Null subjects do not imply obligatorily null expletives (cf. Finnish), and null expletives do not imply available null subjects (cf. German).
 4. ***Pro* vs. pronominal agreement.** Overall, the evidence argues in favor of *pro* as a separate syntactic category (Finnish expletives, agreement asymmetries in Northern Italian dialects).
 5. There is no convincing evidence for the universality of the EPP as a requirement to fill the Spec, IP position. Hence null expletives cannot be obligatory in NSLs.

Orthogonally to these properties, I have also assumed a revised version of subject extraction asymmetries based on R & S, in which extraction asymmetries are the result of strategies for by-passing criterial freezing (induced by subject movement to SubjP).

We are now in a position to state the properties of the NSP as follows:

- (76)
- a. *Pro* is an lexical entry available as part of Universal Grammar.
 - b. *Pro* contains unvalued ϕ -features.
 - c. The features of *pro* are filled in as part of AGREE.

The first three properties in (76) are part of what Rizzi (1986a) termed identification, although, as we will see below, identification entails not only feature-valuation of *pro* but also discourse identification. What remains is to develop a specific theory of ϕ -feature valuation, and to see how this interacts with the discourse contexts.

4.7 Summary

In this section, I have reviewed the three lines of analysis of NS that have been proposed in the literature: NSs as an independently defined category *pro*, NSs as deleted versions of overt pronouns, and NSs as pronominal affixes attached to the verb. I have argued that the evidence for *pro* is, overall, slightly more compelling than for the deletion analysis or the pronominal agreement proposal. As Roberts (2010) points out evidence for pronominal AGR/INFL comes mostly from interpretive facts (see more below), whereas the evidence for a separate null subject come from syntactic paradigms. In order to distinguish the *pro* theory from the pronominal agreement theory, three general sets of facts appear particularly relevant: First, the distribution of overt expletives in Finnish. These expletives directly argue against the pronominal agreement hypothesis and in favor

of an independent *pro*. Second, ellipsis seems to separate overt subjects (where person, number and gender appear not to be interpretable on INFL) from true NSs, where person, number and gender appear to behave like Tense, i.e. as an interpretable feature. Third, the facts from Northern Italian dialects (represented by the Ancona dialect) suggest that in the general case reference does not track inflection but rather a separate category.

I have also argued that the EPP (or subject criterion) should be seen as two separate alignment constraints that interact through constraint ranking.

Part II
On Identification

Chapter 5

Identification and morphology

One of the earliest intuitions regarding the NSP was the notion that consistent NSLs rely on the morphological information available on the verb to recover the reference of the NS. On the one hand, looking at languages like Italian, Spanish or Swahili, one can see that the verb productively and distinctively encodes person and number, whereas in languages like French or English this is not the case. On the other hand, we find that in many languages, null subjects are possible only when they are recoverable from verbal information. For example, NSs are possible in Irish in the synthetic form, which has person/number information (cf. McCloskey and Hale, 1984), but not in analytic form, with no person/number information.

Hebrew presents a different aspect of the same phenomenon. As is well-known, this language only allows for NSs in contexts where person/number inflection is available (either on the verb or on some other functional category, cf. Shlonsky, 2009), which is typically in 1st/2nd person in the past and future tense, as well as with certain types of negation in the present.

A third type of example is illustrated by languages like Pashto (cf. Huang, 1984, 535). This is a split ergative language in which for present events, the verb follows a nominative-accusative pattern, it has a rich agreement system, and it agrees with the subject in transitive and intransitive clauses, as illustrated in (1). For past events, the pattern is ergative, and the verb agrees with intransitive subjects but with transitive objects, as shown in (2).

- (1) a. Jān ra-z-i. (Pashto)
John DIR-come-3.MASC.SG
'John comes.'
- b. zə maṇa xwr-əm.
I apple eat-1.MASC.SG
'I eat the apple.' (from Huang, 1984, 535, ex. 12)
- (2) a. Jān ra-ǧ-ay. (Pashto)
John ASP-come-3.MASC.SG
'John came.'
- b. ma maṇa wə-xwar-a.
I apple PRF-eat-3.FEM.SG
'I ate the apple.' (from Huang, 1984, 535, ex. 13)

When a null category appears, it tracks agreement morphology, thus the subject can drop in both (1a-b) and (2a) (cf. (3) and 4a), but not in (2b) (cf. (4b)). Conversely, the object can drop in (2b) but not in (1b) (cf. (5)).

- (3) a. *pro* ra-z-i. (Pashto)
 DIR-come-3.MASC.SG
 'He comes.'
- b. *pro* maṇa xwr-əm.
 apple eat-1.MASC.SG
 'I eat the apple.' (from Huang, 1984, 536, ex. 13)
- (4) a. *pro* ra-ǧ-ay. (Pashto)
 ASP-come-3.MASC.SG
 'He came.' (from Huang, 1984, 536, ex. 14a)
- b. **pro* maṇa wə-xwar-a.
 apple PRF-eat-3.FEM.SG
 'Ate the apple.' (from Huang, 1984, 535, ex. 16)
- (5) a. ma *pro* wə-xwar-a. (Pashto)
 I PRF-eat-3.FEM.SG
 'I ate it.' (from Huang, 1984, 535, ex. 14b)
- b. *zə *pro* xwr-əm.
 I eat-1.MASC.SG
 'I eat (it).' (from Huang, 1984, 535, ex. 15)

Irish, Hebrew and Pashto illustrate that overt morphological agreement correlates fairly closely with the availability of NSs within single languages. Other languages where this correlation between agreement and availability of NSs holds are Standard, Moroccan, Lebanese and Beni Hassan Arabic, Älvdalsmålet and Angami (cf. Cole, 2009, 574). Even in consistent NSLs like Spanish and Italian, in tenses where the morphological paradigm is not as distinct as in general, NSs become impossible (cf. Lozano (2002); Sheehan (2007) and other references. This is illustrated in (6)). In the first example, the preterite tense in the second clause displays clearly distinct person endings (*abrí* 'opened.1.SG', *abri-ste* 'opened-2.SG', *abri-ó* 'opened-3.SG', *abri-mos* 'opened-1.PL', *abri-eron* 'opened-2/3.PL'), hence the subject can be dropped. In the second example, on the other hand, a NS becomes ungrammatical, because the imperfect singular paradigm has homophonous endings for the 1st and 3rd persons (*tenía* 'had') vs. the 2nd person (*tenías* 'had').¹

¹The unacceptability of (6b) with *pro* substantially improves if the second clause is negated, as in (i). I have no explanation for why this is.

- i. María y yo llegamos a casa. *pro* no tenía las llaves. (Spanish)
 Maria and I arrived to home. *pro* not had.1/3.SG the keys
 'Maria and I arrived home. I/she didn't have the keys.'

- (6) a. María y yo llegamos a casa. Yo/*pro* abrí la puerta. (Spanish)
 Maria and I arrived to home. I/*pro* opened1.SG the door.
 ‘Maria and I arrived home. I opened the door.’
- b. María y yo llegamos a casa. Yo/ella/**pro* tenía las llaves.
 Maria and I arrived to home. I/she/*pro* had.1/3.SG the keys
 ‘Maria and I arrived home. I found the keys.’ (from Sheehan, 2007, 84, ex. 158)

Although the correlation between NSs and agreement is clear, it is also well-known that rich inflection is neither a necessary nor a sufficient condition for NSs, since many languages with rich agreement lack NSs (Icelandic is arguably one of them), and discourse-related NSLs with no agreement have productive NSs. Given this paradox, many have given up on the possibility of establishing a well-delimited connection between overt morphological inflection and the availability of NSs (cf. for example Speas, 1995 and Alexiadou and Anagnostopoulou, 1998, 522-3, fn. 38). Rather, agreement richness is stipulated as an abstract syntactic property of certain heads (cf. Speas, 1995) and the availability of NSs follows from that syntactic property. Cole (2009), on the other hand, proposes a theory that relativizes the notion of morphological richness to individual grammars that allows him to capture general tendencies as well as cross- and intralinguistic variability in availability of NSs.

In the following sections, I will review and adopt Cole’s (2009) proposal for morphological identification, which suggests that grammars set the minimal level of morphological specification at which a NS can be recovered, and this minimal level ranges from full ϕ -feature specification (person, number and gender) as in Tarifit to no morphological specification whatsoever as in Chinese. I will combine this proposal with a hierarchical approach to ϕ -features along the lines of Harley and Ritter (2002) and Béjar (2003), and a particular implementation of AGREE along the lines of Pesetsky and Torrego (2001, 2007, 2004, 2006); Camacho (2011).

5.1 The minimal morphological threshold (MMT)

Cole’s (2009, 569) point of departure is the following observation: In many languages, NSs are possible if uniquely identified by agreement morphology. If that fails, they are identified by reference to an antecedent in context, and if that is not possible, overt pronouns are used. This cascading strategy can be observed in the following Bengali examples. In this language, the subject and the verb agree for person and formality if they are nominative, but not for number; overt (redundant) pronouns are used for focus and to change topics. With this background, consider (7)-(9). (7a) introduces an overt subject *Iqbal*. The NS in (7b) is uniquely identified by verbal morphology (3.PAST) and naturally interpreted as coreferent with *Iqbal*, illustrating identification through verbal morphology.

- (7) a. Iqbal Calcutta-e budhbar-e gœlo. . . (Bengali)
 Iqbal to-Calcutta on-Wednesday go.3.PAST
 ‘Iqbal went to Calcutta on Wednesday. . .’

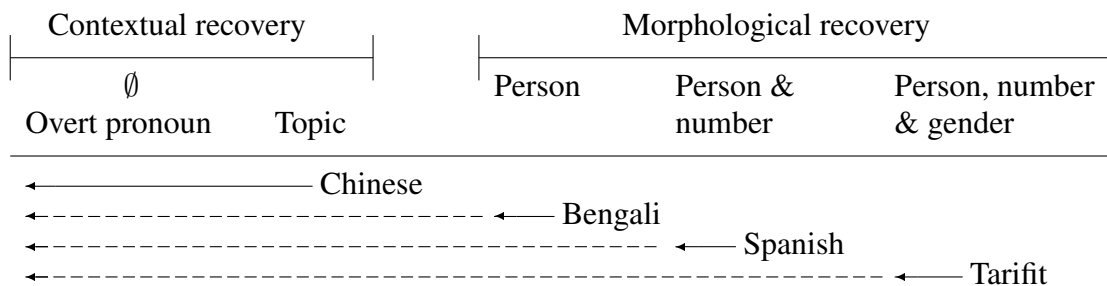
- b. Brihoshpothibar-e *pro* gari kinlo.
 on-Thursday car buy.3.PAST
 ‘On Thursday he bought a car.’ (from Cole, 2009, 576, ex. 38)

By contrast, in (8), the same discourse structure is set up, but the verbs in (8b) are both ambiguous between 2nd/3rd person, so they can only be interpreted through their antecedent in discourse. In this case, morphology is not enough to identify the NS. The same situation arises with (9b), so that if the intended interpretation of the third subject is 2nd/3rd person, there must be an overt pronoun (*tumi* ‘you’) because the verb is ambiguous between 2nd/3rd person, so morphological identification fails and the only overt antecedent is 3rd person.

- (8) a. Shonibar-e Iqbal Calcutta-e jabe... (Bengali)
 on-Saturday Iqbal to-Calcutta go-2/3.FUT
 ‘On Saturday, Iqbal will go to Calcutta.’
 b. *pro* gari kinbe *pro* Restaurant-e jabe.
 car buy-2/3.FUT to-restaurant go-2/3.FUT
 ‘He will buy a car. He will go to a restaurant.’ (from Cole, 2009, 576, ex. 39)
- (9) a. Shonibar-e Iqbal Calcutta-e jabe ... (Bengali)
 on-Saturday Iqbal to-Calcutta go-2/3.FUT
 ‘On Saturday, Iqbal will go to Calcutta.’
 b. *pro* Gari kinbe. Tumi restaurant-e jabe.
 car buy-2/3.FUT you-S to-restaurant go-2/3.FUT
 ‘He will buy a car. You will go to a restaurant...’ (from Cole, 2009, 576, ex. 40)

This basic idea that recoverability first resorts to morphological identification, then to identification by antecedent and finally by inserting an overt pronoun, is combined with a variable scale of what constitutes successful morphological identification depending on the language. The general scale is presented in (10), and languages define at which point in the scale subjects can be recovered from the morphology (Cole’s so-called “morphological maximality”). Thus, NSs are morphologically recoverable by a bundle of ϕ -features that include person, number & gender in Tarifit, otherwise by Topic, otherwise an overt pronoun will be required. For Spanish, the morphological break-point is Person & number, etc.

(10) Recoverability scale



Under this perspective, what constitutes morphological richness to identify a NS is not uniformly defined for all languages, and a given language may define it as including person, while a paradigm that encodes the same distinctions in an other language may not allow recoverability. I will call the point at which a language allows for morphological identification of a NS **Minimal morphological threshold** (MMT) for NSs, as in (11). The MMT for Tarifit will be overt morphological encoding of person, number and gender, whereas for Bengali it will just be person.

- (11) The **Minimal morphological threshold** (MMT) defines the minimal set of values overtly encoded in the morphology that a language requires to identify a null subject.

One might consider whether an MMT at given point in the scale allows NSs to be identified by a overt *phi*-feature settings to the right of that point in the scale. Consider the case of gender in Arabic. As Cole (2009, 579) points out, gender is restricted to a 2nd and 3rd person singular and plural in Arabic, but does not appear in 1st, as illustrated by comparing (12a) with (12b-c). However, since NSs are possible in all persons, the MMT for Arabic cannot contain gender, otherwise NSs should not be possible in (12c).

- (12) a. (ʔanā) raʔay-tu zayd-an. (Arabic)
 (I.NOM) see-PAST.1.SG.NOM Zayd-ACC
 ‘I saw Zayd.’
- b. (ʔanti) takallam-ti
 (you.SG.FEM.NOM) speak-PAST.3.SG.FEM.NOM
 ‘You (fem.) spoke.’ (from Johns, 2007, 129, ex. 14, 15)
- c. (huwa) ɕāʔ-a
 (he.NOM) came-3SG.MAS.NOM
 ‘He came.’ (from Fassi Fehri, 1993, 115, ex. 48b)

Cole’s solution involves assuming that “morphological maximality [in Arabic] is for person and number in the 1st person, but for person, number and gender in the 2nd and 3rd persons.” (p. 579).

This issues raises the question of whether ϕ -features are simply listed, or whether they have internal structure that would derive which features are accessible to identify a NS in any given paradigm. I will explore the issue in section 5.2.

Before doing so, it is worth discussing a consequence of this point of view: In languages with obligatorily overt subjects, identification by morphology fails, although it may do so in different ways. For example, in English, one might speculate that the MMT is set to have a person value of at least 1st or 2nd, so that even distinctive 3rd persons in the present do not identify a NS. In other languages, it may be specified as any person, so that even if the verb marks 1st, 2nd and 3rd persons distinctively, NSs will still not be identified.

However, if a NS is not identified morphologically in a language like English or German, why can’t it be identified by a discourse antecedent, the way it is in Chinese? So isn’t (13) possible?

- (13) Mary left last night and Bill left this morning. *As for John, is leaving tomorrow. (English)

In chapter 6 we will propose a solution to this question based on two ideas: The nature of NSs is different in Chinese-type and English-type languages and the locality requirements for identifying the antecedent are also different.

5.2 The structure of ϕ -features

Cole's (2009) proposal does not specifically assume an internal hierarchy for ϕ -features, but as I pointed out in the preceding section, there may be reasons to think that ϕ -features are hierarchically organized. In addition to the issue of languages like Arabic, where gender is expressed in some persons but not others, the scale in (10) seems to have an underlying structure: Points of maximality display a certain implicational relation, such that number implies person and gender implies both person and number. It is thus worth considering whether research on the internal structure of ϕ -features can shed light on this issue.

5.2.1 Φ -feature hierarchies

Following Harley and Ritter (2002), Béjar (2003) proposes that ϕ features are organized hierarchically as in the partial representation in (14). One of the primary purposes of the hierarchy is to capture entailment relations among features. Thus, speaker entails participant, and inanimate entails class. Notice that these representations are underspecified, and features are privative.

(14)

REFERRING EXPRESSION			
π	INDIVIDUATION		
PARTICIPANT	GROUP	MIN.	CLASS
SPEAKER	ADDRESSEE	ANIMATE	INANIMATE/ NEUTER
	FEMININE	MASCULINE	

Regarding person, the hierarchy doesn't explicitly label any node as "1st or 2nd person, rather, those are represented as in (15). The feature π (person) formalizes the fact that "3rd persons subcategorize in more ways than can be represented by the binary contrast between participants and nonparticipants" (Béjar, 2003, 49).

(15)

Person:	3rd person	2nd person	1st person
	R	R	R
	π	π	π
	PART	PART	PART
		ADDR	SPKR

This feature hierarchy is intended to be a general template that applies to elements bearing ϕ -features. However, verbal inflection and pronouns have two crucial differences: Pronouns have potential referential capacity and categorial specification, whereas verbal inflection need not. In fact, if the arguments for *pro* given in sect. 4.2.1 are correct, inflection does not have either property. This distinction is not reflected in the structures in (14) and (15) in any obvious way.

However, Béjar (2003, 49) suggests the additional entailment hierarchy in (16) which includes the features DEICTIC, DEFINITE, SPECIFIC and sc d. π , on the other hand, can range over any of the features in this proposal.

(16) SPEAKER > PARTICIPANT > DEICTIC > DEFINITE > SPECIFIC > D > R(EFERENT)

The second problem, i.e. the categorial nature of pronouns comes to light when considering languages like Chinese or Japanese, which seem to have argumental NPs (as opposed to DPs). This possibility clusters with four others, according to Chierchia (1998, 354, ex. 20): a) generalized bare arguments (i.e. absence of determiners), b) the extension of all nouns is mass, c) absence of plural morphology and d) generalized classifier system. Tomioka (2003), in turn, has argued that null pronouns simply null counterparts of overt categories (cf. Hoji, 1998 and Tomioka, 1999).

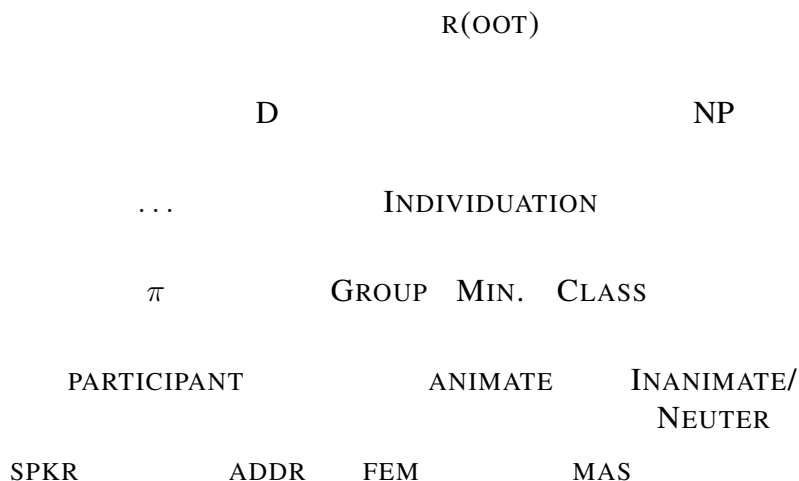
One possible way to capture these two properties (where referentiality is located and the NP/DP category of pronouns) would be to assume first that referentiality is not located in REFERENT. One reason for this is that expletives can lack referential properties but have the same agreement properties as fully referential pronouns, as seen in (17).² There must be some way to signal that *it* in the first case lacks reference whereas it refers to *the tree branch* in the second case, and if this property is encoded in R, then we have no way of representing (17a). Based on this evidence, let us assume that R simply stands for ROOT, and that referentiality can be encoded in any of the features dominated by R up to DEICTIC.

- (17) a. It is snowing. (English)
 b. I was walking under a tree branch when it fell and hit me.

Second, following Abney's (1987) original intuition that pronouns are Ds with a null NP, we can assume that the feature geometry in (14) (including necessary amendments from (16) corresponds to D, and N/NP is a sister to D, as in (18).

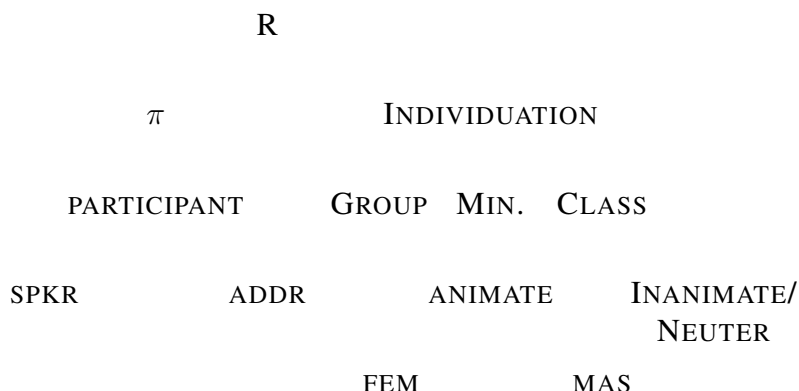
(18) ϕ -feature geometry for pronouns

²As mentioned in sect. 2.1.1.1, several authors have proposed that weather-verb expletives have referential properties, I will disregard that possibility.



In the case of inflectional ϕ -features, the assumptions made above suggest that they do not have referential properties. This raises the question whether the nodes related to referential properties (D, SPECIFIC, DEFINITE and DEICTIC) should be present and represented as in (18) or absent as in (19).

(19) Impoverished ϕ -feature geometry for Infl

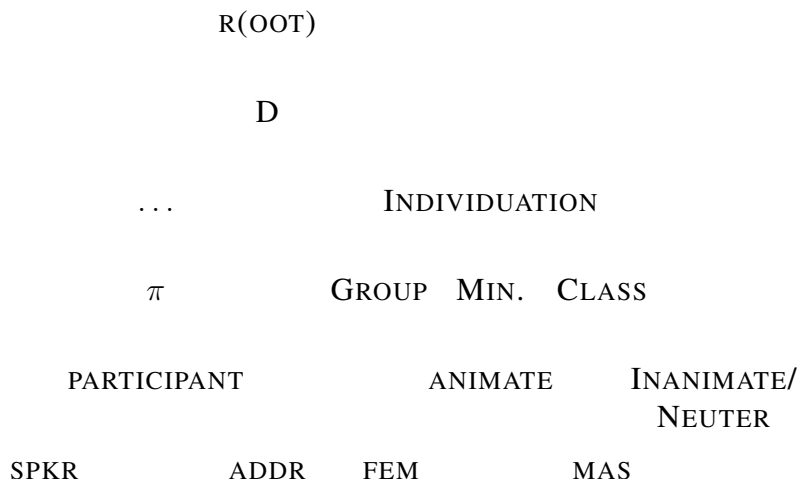


Several morphological systems are sensitive to specificity, particularly with objects. For example, Swahili definite objects trigger overt agreement but indefinites do not, as seen in (20). Since we want the feature geometry to be as general as possible, I tentatively conclude that inflectional agreement should include the nodes in (18), although those nodes that have referential properties are uninterpretable.

- (20) a. U- me- leta kitabu? (Swahili)
 2.SG PERF bring book
 ‘Have you brought a book?’
- b. U- me- ki- leta kitabu?
 2.SG PERF 3.SG bring book
 ‘Have you brought the book?’ (from Perrott, 1972, 38, quoted in Croft, 1988, 161)

However, what is different between a pronominal ϕ -feature geometry and Infl's geometry is the absence of the NP branch for Infl. While it may make sense to match definiteness or deixis, I can't think of a reasonable counterpart in the inflectional morphology to the content NP. Rather, as we will see, I will propose that languages like Chinese or Japanese have pronouns that lack D but have NP, and in those languages we expect no agreement. Hence, the proposed geometry for Infl would be the one in (21), where . . . represents SPECIFIC, DEFINITE and DEICTIC.

(21) ϕ -feature geometry for Infl



5.3 AGREE

Béjar's (2003) analysis is built on the operation AGREE (cf. Chomsky, 2000, 2001), which involves at least two crucial steps: First, the probe and the goal must MATCH, then the goal VALUE the probe's features. The first operation is only possible if the features of the probe are a subset of the features of the goal as (22) (cf. Béjar (2003, 53)).

(22) a. MATCH is valued at the root.

b. Probe (F) and Goal (F') match if Goal (F') entails root Probe (F).

(from Béjar, 2003, 53, ex 37)

Given these principles, MATCH would have the outcomes in (23). Only the combinations in the first two lines MATCH, whereas the third one does not, because the feature specification of the goal does not entail that of the probe.

(23) Probe-goal match outcomes depending on feature specification.

(from Béjar, 2003, 55, ex. 41)

[F] Probe	[F'] Goal	MATCH
$[\pi]$	$[\pi]$	Yes
$[\pi]$	$[\pi[\text{PART}]]$	Yes
$[\pi]$	$[\text{INDIV}]$	No

Finally, Bejar assumes an asymmetry between interpretable and uninterpretable features. A feature [F] on an uninterpretable probe does not automatically entail features that dominate [F], whereas it does if the feature is interpretable. Thus, for example, (24) does not entail the features in parenthesis, otherwise the Probe and the Goal would match at the root ([R]).

(24)	Probe		Goal
	(R)		R
	π		

The second part of AGREE involves copying the features of a goal to a probe, and subsequently deleting them, as in (25).

(25) Agreement involves copying the features of a goal to a probe and deleting them (after Béjar, 2003, 60, ex. 47)

Probe(F)			Goal(F')
π	←		π
PART			PART
SPKR			SPKR

I adopt a the definition of AGREE presented in (26), which departs from the conception adopted in Chomsky (2000, 2001); Béjar (2003) in important ways and builds on Pesetsky and Torrego (2001, 2007, 2004, 2006); Baker (2008); Camacho (2010b) and Camacho (2011). Specifically, the three most important differences with the Chomsky-based AGREE theory are the distinction between probes and goals, the distinction between interpretability and valuation and the elimination of c-command from the definition.

(26) A and B AGREE iff:

- a. At least one of the categories is functional and acts as a probe
- b. Both categories' ϕ -features MATCH
- c. A and B's ϕ -features are VALUED
- d. AGREE is locally constrained to the phase that contains the matching categories

I assume, together with Béjar (2003), that the probe must be a subset of goal for MATCH, so that a probe specified as [π] would match a goal specified as [π [PART]], but not a goal specified as [INDIV]. VALUATION is defined as in (27).

- (27) a. VALUATION: When two categories have matching ϕ -features, the values of one can be copied onto the other.
- b. VALUATION can only specify a node, not delete or change its value

In the conception of AGREE in (26), only the probe has to be functional, the goal can be functional or lexical. Thus, a Determiner (D) head probes a lexical Noun (N), but I (or *v*, depending on one's theory) probes a functional DP. Second, I assume that at the end of AGREE, both matching categories must have with valued features, but the direction of copying can be from the probe to the goal or the goal to the probe. What determines directionality is which of the ϕ -feature structure is more specified, so copying cannot delete or replace existing features, but it can append features to existing nodes. This is illustrated in (28).

- (28) Outcomes of VALUATION

Before Valuation		After Valuation		
A	B	A	B	Outcome
$[\pi]$	$[\pi[\text{PART}]]$	$[\pi[\text{PART}]]$	$[\pi[\text{PART}]]$	✓
$[\pi[\text{PART}]]$	$[\pi]$	$[\pi[\text{PART}]]$	$[\pi[\text{PART}]]$	✓
$[\pi]$	$[\pi[\text{PART}]]$	$[\pi]$	$[\pi]$	*
$[\pi[\text{PART}]]$	$[\pi]$	$[\pi]$	$[\pi]$	*

Finally, no c-command condition is built into the definitions. In Baker (2008), the c-command condition is loosened to allow for cases in which either the Probe c-commands the Goal or the Goal c-commands the probe. In most cases, c-command constrains agreement patterns that would also be ruled out by cyclicity. Consider, for example, cases in which an argument in base-position blocks agreement with a lower argument, but movement of the higher argument unblocks agreement, as schematically represented in (29). Cases like these are well-documented in Icelandic (cf. Sigurðsson and Holmberg, 2008), and many other languages (cf. Béjar, 2003). These contexts are readily explained if one assumes that in (29b), DP_{DAT} no longer intervenes because the probe (V) does not c-command it.

- (29) a. EXP V DP_{DAT} DP_{NOM}
*
- b. DP_{DAT} V t DP_{NOM}
✓

However, the contrast in (29) can also be explained if in (29b), DP_{DAT} is outside of the agreement domain. To that effect, one of the dialects of Icelandic described by Sigurðsson and Holmberg shows the pattern in (30), where displacing the dative does not allow for full agreement with the nominative. If all that is at stake in (29) is the c-command configuration, then the facts in (30) come as a surprise.

	Singular:	3rd person	2nd person	1st person
		R	R	R
		D	D	D
	...	INDIV	...	INDIV
	π	CLASS	π	CLASS
		FEM MAS	PART FEM MAS	PART FEM MAS
			ADDR	SPKR
	Plural:	3rd person	2nd person	1st person
		R	R	R
		D	D	D
	...	INDIV	...	INDIV
	π	CLASS GROUP	π	CLASS GROUP
		FEM MAS	PART FEM MAS	PART FEM MAS
			ADDR	SPKR

Assume that *pro* is generally specified as R, or perhaps R[D]. As a result, the AGREE operation would proceed as in (33) for 2nd person.³

(33) AGREE between 2nd p. sg. Infl and *pro* in Tarifit

³In this representations, I have omitted some intermediate nodes, signaling them with ..., but the full geometry corresponding to 2nd person singular in (32) would be copied.

MATCH: Probe (Infl) \longleftrightarrow Goal (*pro*)

R R

ldots INDIV

PART CLASS

ADDR FEM MAS

VALUATION: Probe (Infl) \longleftrightarrow Goal (*pro*)

R R

... INDIV *ldots* INDIV

PART CLASS PART CLASS

ADDR FEM MAS ADDR FEM MAS

If the probe is underspecified, *pro* would be unvalued for some relevant node. As a result of how Tarifit defines the MMT, *pro* would not be sufficiently identified.

5.4.1 MMT set to person and number: Italian and Spanish

In the case of Italian and Spanish, NS require at least person and number, as illustrated in (34). This means that MMT is set to person and number, and the ϕ -feature geometry is the one proposed in (35) ignoring the intermediate nodes between R and π .

(34) Comi-mos almendras. (Spanish)

eat-1.PL almonds
'We ate almonds.'

(35) Finite ϕ -feature probes in Spanish

Singular:	3rd person	2nd person	1st person
	R	R	R
	D	D	D
	... π	... π	... π
		PART	PART
		ADDR	SPKR
Plural:	3rd person	2nd person	1st person
	R	R	R
	D	D	D
	... π INDIV	... π INDIV	... π INDIV
	GROUP	PART GROUP	PART GROUP
		ADDR	SPKR

The *pro* identification process proceeds as in (36).

(36) AGREE between 2nd p. sg. Infl and *pro* in Spanish

MATCH: Probe (Infl) \longleftrightarrow Goal (*pro*)

R	R
D	D
... π	
PART	
ADDR	

VALUATION: Probe (Infl) \longrightarrow Goal (*pro*)

R	R
D	D
... π	... π
PART	PART
ADDR	ADDR

Notice that the remaining *pro* is not specified for CLASS, because the probe is unspecified for that feature. However, there is some reason to believe that *pro* can be potentially specified for gender, as in the case of secondary predicates or small clause subjects, illustrated in (37). In this example, *pro* is necessarily identified as feminine because the secondary predicate is specified as feminine. For these cases, I take the feminine specification of *pro* to come from the overt morphology on the secondary predicate. In the case of secondary predicates, gender is not enough to value/identify *pro*, as seen in (38).

- (37) a. *pro*_i salió *pro*_i content-a. (Spanish)
 left happy-FEM
 ‘She left happy.’
- b. La_i consider-o [*pro*_i atrevid-a]
 CL.FEM consider-1.SG daring-FEM
 ‘I consider her daring.’

- (38) *Notamos [*pro* content-a] (Spanish)
 saw happy-FEM
 ‘We saw Ana happy.’

Assuming that the functional head associated with the adjective is specified as in (39), *pro*’s resulting specification will fail the MMT for Spanish.⁴

(39) Representation for Adjectival ϕ -features in Spanish

⁴Baker (2008) argues that adjectival agreement fundamentally differs from verbal agreement in that the first type lacks 1st and 2nd person agreement. In his theory, this follows from two principles, first, structurally, verbs project a Specifier, whereas adjectives do not, and second, 1st and 2nd person agreement takes place in a specifier configuration (his SCOPA, p. 52). Both theories might converge in the following sense. Some of the agreement mismatches Béjar (2003) analyzes are triggered when a target for agreement moves to the specifier of the probe, expanding the agreement search space. If adjectives lack a Specifier by design, these types of effects would be systematically absent with adjectives. I do not know if this prediction is correct.

(42) Finite ϕ -structure probes for Bengali

Sg/pl:	3 rd person	2 nd person	1 st person
	R	R	R
	D	D	
	... π	... π	... π
	PART	PART	PART
		ADDR	SPKR

MATCHING and VALUATION in Bengali are represented in (43). Like for gender in Spanish, I do not assume that number is specified for *pro* in Bengali, although the corresponding overt pronouns do distinguish between singular and plural (cf. Sengupta, 1999, 278).

(43) MATCH: Probe (Infl) \longleftrightarrow Goal (*pro*)

R	R
D	
... π	
PART	
ADDR	

VALUATION: Probe (Infl) \longrightarrow Goal (*pro*)

R	R
\emptyset	D
... π	... π
PART	PART
ADDR	ADDR

5.4.3 Gender in Arabic

Recall that Arabic 1st/2nd persons are specified for gender, whereas 3rd is not. In the current approach, this would be captured by distinct ϕ -feature hierarchies for each type of probe, as shown in (44) for the singular. MMT in Arabic includes at least π , so all of those representations would value NSs that could be identified.

(44) ϕ -structure for Arabic probes:

Singular:	3rd person	2nd person	1st person
	R	R	R
	D	D	D
	... π INDIV	... π INDIV	... π
	CLASS	PART CLASS	PART
	MAS FEM	ADDR MAS FEM	SPKR

5.4.4 Mixed paradigm in Hebrew

In sect. 2.3.3, we observed two types of partial NSLs, ones in which NSs are available depending on person and/or tense, and those in which NSs are restricted to main/embedded clauses. In this section, I consider the first type, in particular Hebrew. This language shows three separate asymmetries that regulate the distribution of NSs (cf. Shlonsky, 2009, 133): the referentiality asymmetry, the person asymmetry and the tense asymmetry:

- (45)
- a. The referentiality asymmetry: Non-referential argumental NSs are possible in every tensed environment but one.
 - b. The person asymmetry: Referential NSs are permitted with 1st and 2nd person inflection; 3rd person NSs are only possible in contexts of (non-standard) binding and/or Control.
 - c. The tense asymmetry: Referential NSs are only possible in past and future tense clauses, not in present tense clauses (from Shlonsky, 2009, 133)

For present purposes, I will focus on the person and tense asymmetries, illustrated in (46)-(47) and (48) respectively.⁶

⁶The person asymmetry reappears in the present with negative *eyn*, which carries inflection. In these contexts, only 1st/2nd are possible, cf. Shlonsky (2009, 137).

- (46) a. *Lamad albanit. (Hebrew)
 study.PAST.3.MAS.SG Albanian
 ‘He studied Albanian.’
- b. *Lamd-da albanit.
 study.PAST.3.FEM.SG Albanian
 ‘She studied Albanian.’
- c. *Lamd-du albanit.
 study.PAST.3.PL Albanian
 ‘They studied Albanian.’
- (47) a. Lamad-ti albanit. (Hebrew)
 study.PAST.1.SG Albanian
 ‘I studied Albanian.’
- b. Lamad-ta albanit.
 study.PAST.2.MAS.SG Albanian
 ‘You (mas) studied Albanian.’
- c. Lamad-t albanit.
 study.PAST.2.FEM.SG Albanian
 ‘You (fem) studied Albanian.’
- d. Lamad-nu albanit.
 study.PAST.1.PL Albanian
 ‘We studied Albanian.’
- e. Lamad-tem albanit.
 study.PAST.2.PL Albanian
 ‘You (pl) studied Albanian.’ (from Shlonsky, 2009, 135, ex. 5)
- (48) a. *Lomed albanit. (Hebrew)
 study.PRES.MAS.SG Albanian
- b. *Lomed-et albanit.
 study.PRES.FEM Albanian
- c. *Lomd-im albanit.
 study.PRES.MAS.PL Albanian
- d. *Lomd-ot albanit. (from Shlonsky, 2009, 136, ex. 7)
 study.PRES.FEM.PL Albanian

Shlonsky (2009) analyzes these two asymmetries in the following way: The ϕ -feature matrix of the present and future Tense lacks the slot for [person]. 1st and 2nd persons, on the other hand, are taken to head a separate syntactic head, a Speech Act Participant Phrase located above TP, in the left periphery (cf. Sigurðsson, 2010 for a similar proposal for Finnish). “Sap⁰ is filled by a pronominal clitic, moved from the position of the thematic subject” (Shlonsky, 2009, 142). The

clitic is moved to Sap^0 , and the verbal complex in T adjoins to that head. In essence, Shlonsky suggests that Hebrew lacks covert 1st and 2nd pronouns, but it has 1st and 2nd clitics.

In the approach I have proposed here, Hebrew would set the MMT to [PART]. Regarding the tense asymmetry, Shlonsky's suggestion that ϕ -features lack [person] means that it lacks the relevant [PART] node, as illustrated in (49). In this representation, *pro* does not have a value that satisfies Hebrew's MMT, and additionally, D is not deleted on Infl.

(49) AGREE between 2nd p. sg. present Infl and *pro* in Hebrew

VALUATION: Probe (Infl) \longrightarrow Goal (*pro*)

R	R
D	

The account correctly predicts that *pro* should not be possible in the present. In the past and future Infl values *pro* so that it satisfies the MMT only in 1st and 2nd person, because only those persons are specified for [PART], whereas 3rd persons lacks the relevant feature, as seen in (50).

(50) Hebrew 2nd p. sg. past Infl Hebrew 3rd p. sg. past Infl
 VALUATION: Probe (Infl) \longrightarrow Goal (*pro*) Probe (Infl) \longrightarrow Goal (*pro*)

R	R	R	R
\emptyset	D	\emptyset	D
... π	... π	... π	... π
PART	PART \leftarrow MMT		
ADDR	ADDR		

Shlonsky (2009) argues that 3rd person null subjects are possible when controlled by another argument, an issue we will return to below. For now, these cases illustrate that in the relevant agreement domain, 3rd *pro* cannot be identified, as predicted by the theory.

5.4.5 Mixed paradigms in German Vernaculars

A similar account applies to Zurich German, Schwabian and Övdalian, illustrated in (51)-(53). The inflectional paradigm of these varieties is presented in (54), where the persons that allow for NSSs are in bold.

(51) a. ... daß scho des Buch kauft hasch. (Schwabian, only 1st/2nd sg.)
 that already the book bought have.2SG
 ‘... that you already have bought the book.’

- b. Geschtern han-mr en Bobbel Eis kauft.
yesterday have-1.SG.me.CL a ball ice cream bought
'Yesterday I bought myself a ball of ice cream' (from Rosenkvist, 2009, 164, ex. 13)
- (52) a. Ha der das nöd scho verzellt? (Zurich German, only 1st/2nd
have.1.SG to-you it not already told
sg.)

'Haven't I told you that already?'
- b. Wänn nach Züri chunnsch, muesch mi bsueche.
when to Zurich come.2.SG must.2.SG me visit
'When you come to Zurich, you must visit me.' (from Rosenkvist, 2009, 164, ex. 15)
- (53) a. ... dar wilum glãmå min wennanan. (Övdalian, only 1st/2nd pl.)
when want-to.1.PL chat with each-other
'when we want to chat with each other.'
- b. Nū irið iema.
now are-2.PL. home
'Now you are home.' (from Rosenkvist, 2009, 169, ex. 22)
- (54) Verbal inflection paradigm for *to come* in Schwabian, Zurich German and Övdalian (from Rosenkvist, 2009, 171, table 3)

	singular			plural			Inf.
	1	2	3	1	2	3	
Zurich German	chume	chunnsch	chunnt	chömed			chu
Swabian	komm	kommsch	kommt	kommet			komma
Övdalian	kumb			kumum	kumið	kumå	

The distribution of NSs in Zurich German, Schwabian, and Övdalian fit well in the proposed analysis. All three varieties set the MMT to [PART]. This predicts that none of the varieties will allow for NSs with 3rd person probes. For Zurich German and Schwabian, we can assume that all persons in the plural lack a [PART] node, given that they are identical, as represented in (55). In Övdalian, on the other hand, the reverse situation holds: All persons in the singular are identical, so they are underspecified for [PART]. Finally, Zurich German and Schwabian 1st and 2nd person singular geometries specify [PART] as does Övdalian 1st and 2nd person plural.

(55) ϕ -structure for Zurich German, Schwabian and övdalian:

Singular: **ZG and Schwab 1/2 p.**

Ovd 1/2/3 p.

R

R

D

D

... π

... π

PART

SPKR ADDR

Plural: **ZG and Schwab 1/2/3 p.**

Ovd 1/2 p.

R

R

D

D

... π INDIV

... π INDIV

GROUP

PART GROUP

ADDR SPKR

5.4.6 Summary

In the preceding sections, I have adopted the idea that that NS identification is constrained by Cole's recoverability scale in (10) that ranges from morphological recovery to contextual recovery. Each language sets a minimal morphological threshold (MMT) to recover a NS. I have applied the notion of MMT, combined with the feature geometry presented earlier to Tarifit (MMT set to person, number and gender), Italian/Spanish (MMT set to person, number), Bengali (MMT set to person) and to mixed paradigms such as Hebrew and German Vernaculars.

I have also argued that ϕ -feature bundles are best viewed as structured hierarchies, rather than lists. Additionally, I have adopted the idea that agreement between two categories involves copying of the probe's feature root-node to the the goal's root node. These assumptions have as a consequence that agreement will not target intermediate nodes in the hierarchy.

5.5 Identification and locality

All of the instances of identification by morphology presented in the preceding section involve AGREE between a probe and a goal within the same clause. In other words, subjects agree with the verb in their own clause, not with another verb. So, for example, (56) cannot mean ‘I think that I left’, with a structure in which the embedded *pro* is identified by the matrix probe.

- (56) Cre-o que sali-ó. (Spanish)
 think-1.SG that left-3.SG
 ‘I think that s/he left’ not ‘I think that I left.’

In general, then, AGREE is constrained to IP/CP. In this section, I analyze the contrast between Standard German and Germanic vernaculars as a difference in the location of the agreeing probe. This distinction derives the fact that in German, despite strong agreement, NSs are not licensed, whereas in the vernaculars, it is.

5.5.1 On the locality of head-agreement: Standard Germanic vs. C-agreeing German Vernaculars

Standard German is frequently quoted as a problem for accounts that link rich agreement with the ability to have a NS. Specifically, the German inflectional paradigm in (57) shows that endings for all persons are distinctive, but this variety lacks NSs, as shown in (58).

- (57) Verbal inflection in Standard German for *arbeiten* ‘to work’ (from Jaeggli and Safir, 1989, 28, ex. 38)

	singular			plural		
	1	2	3	1	2	3
Standard German	arbeit-e	arbeit-est	arbeit-et	arbeit-en	arbeit-t	arbeit-en

- (58) Sie kenne *(ich) nicht. (Standard German)
 her know I not
 ‘Her I do not know.’ (from Rosenkvist, 2009, 151, ex. 3)

At the same time, several regional varieties of Germanic do have NSs, as already shown in (51)-(53) above. Another set of German varieties, including Bavarian, Lower Bavarian and Frisian, also have NSs, as illustrated in (59)-(61).

- (59) a. ... obst noch Minga kummst. (Bavarian, 2nd sg. and pl. only)
 if.2.SG to Munich come.2.SG
 ‘Whether you come to Munich’

- b. Hobbds khoa geld nimma.
 have-2.PL no money not-anymore
 ‘You have no money anymore.’

(from Rosenkvist, 2010a, ex. 9)

- (60) Fahrma noch Minga? (Lower Bavarian, only 2nd sg./pl. and 1st pl.)
 travel.1.PL to Munich
 ‘Are we going to Munich?’

- (61) a. Miskien moatst my helpe. (Frisian, only 2nd sg.)
 perhaps must.2.SG me help
 ‘Perhaps you must help me.’

- b. Ik tink datst my helpe moatst.
 I think that.2.SG me help must.2.SG
 ‘I think that you must help me.’

(from Rosenkvist, 2009, 167, ex. 19)

However, NSs in these varieties are only possible if C is inflected, as the contrast between (59) and (62) shows. In the second set of examples, the complementizer lacks inflection, although the verb still has it.⁷

- (62) a. *Ob *pro* noch Minga kumm-st. (Bavarian)
 if to Munich com.2.SG
 b. *Ob *pro* noch Minga kumm-ts.
 if to Munich com.2.PL (from Bayer, 1984, 240)

In sum, abstracting away for the moment from the person restrictions, we seem to have three patterns: Those that have NSs in the I-field (Zurich German, Schwabian and Ovdalian), those that NSs in the C-field (Lower Bavarian, Bavarian and Frisian) and those that lack NSs (Standard German). One possible way to account for these differences would be to relativize licensing of *pro* to the two distinct areas, IP and CP. Suppose, then, that the I-field languages identify *pro* in IP, hence MMT is defined with respect to probes in Infl. Thus, since Zurich German, Schwabian and Övdalian have probes specified for [PART] in Infl, 1st/2nd person subjects can be null.

In Bavarian and Lower Bavarian, on the other hand, the probe is located in C, so the only NSs that will be licensed are 2nd sg./pl., 2nd sg./pl. and 1st plural respectively. This are the only persons that are specified at the C-level, as seen in (63).

- (63) C-related inflectional paradigm in Bavarian and Lower Bavarian

	Singular			Plural		
	1	2	3	1	2	3
Bavarian		-st			-ts	
Lower Bavarian		-st		-ma	-ts	

⁷Bayer (1984); Fuß (2005) give convincing evidence that the C-related morphemes of Bavarian and Lower Bavarian are inflectional morphemes, not independent clitics.

This means that recoverability in these two dialects is set as in (64), whereas π -feature geometries for probes are the ones in (65). Consequently, only 2nd person sg./pl. in both varieties and 1st person pl. in Lower Bavarian will satisfy morphological maximality.

(64) MMT in Bavarian and Lower Bavarian: [R [D ... [π [PART]]]]

(65) Probe ϕ -structure for Bavarian and Lower Bavarian:

Singular:	Bav./L. Bav.		Plural:	Bav./L. Bav.		Bav.	L. Bav.
	2nd p.	1/3rd p.		2nd p.		1st p.	1st p.
	R	R		R		R	R
	D	D		D		D	D
	... π	... π		... π	INDIV	... π	INDIV ... π
	PART			PART	GROUP	PART	GROUP
	ADDR			ADDR		SPKR	

Consistent with this account, Bavarian displays *that*-trace effect violations, as illustrated in (66). Under the Rizzi-Shlonsky freezing approach, this is expected, since one of the ways to by-pass *that*-trace effect violations is to have an agreeing complementizer.

- (66) a. Wer moanst du [t mog d'Emma]? (Bavarian)
 who think.2.SG you loves Emma
 'Who do you think that loves Emma?'
 b. Wer moanst du [daß t d'Emma mog]?
 who think.2.SG you that Emma loves
 'Who do you think that loves Emma?'
 c. Weam moanst du [daß d'Emma t mog]?
 who.ACC think.2.SG you that Emma loves
 'Who do you think that Emma loves?' (from Bayer, 1984, 210, ex. 2)

Turning now to Standard German, we can assume that like Bavarian and Lower Bavarian, the probe that agrees with the subject *pro* is in C, but unlike in those varieties, in Standard German, C-agreement does not have inflection, hence the probes are unspecified for [PART]. If the MMT is the same for all three varieties, it follows that referential NSs are not identified in Standard German.

To summarize, the proposed analysis locates the crucial difference between German/Bavarian and Zurich German/Schwabian (and Spanish/Italian, for that matter) in where the probe for subject agreement is located: In German/Bavarian, in C, in Zurich German/Schwabian (or Italian/Spanish)

in I/T. German and Bavarian, in turn, vary because of how specified the probe is in each language. It follows, then, that while German shows a rich inflectional paradigm on the verb, it does not have referential NSs, whereas another one with an essentially identical inflectional paradigm to Standard German does.⁸

5.5.2 Local agreement vs. distant reference: First and second person shifted readings

The so-called shifted readings of Slave and Amharic show instances in which an embedded 1st/2nd person is fully valued within its clause, but it is interpreted as if it had the features of a matrix antecedent (cf. Rice, 1989; Schlenker, 2003; Baker, 2008). This suggests that in addition to valuation, there is an independent mechanism that determines actual discourse anchoring.

Person shift is illustrated in (67). As Baker (2008, 124-150) points out, the embedded 1st or 2nd person subject is interpreted as if coreferential with a 3rd person antecedent in these examples. In addition to that reading, there is a non-shifted one, where the 1st person refers to the speaker and not to the matrix subject.⁹

- (67) Tony łue ghəshéohťí enídhe. (Slave)
 Tony fish 1.SG.SUB.OPT.eat 3.SG.SUB.wants
 ‘Tony wants to eat fish.’
 (Lit. ‘Tony wants I eat fish’) (from Rice, 1989, 1273, quoted in Baker, 2008, 129)

In Slave, we find three types of main verbs with respect to shifted readings: Those that allow for shifting of 1st persons (as in (67)), those that also allow for shifting of 2nd persons, and those that do not allow for shifting at all. Whether a given verb is of one type or another is a lexical property. For example, the equivalent of ‘know’ does not induce shifting, whereas the equivalent of ‘think’ does. One important syntactic property that distinguishes shift-inducing verbs from those that do not allow person shifting is that the first type lacks a complementizer, whereas the second type must have one.

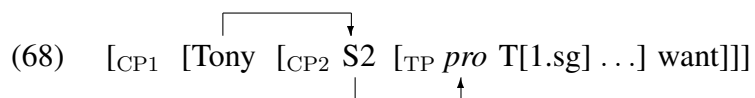
At first sight, it might look like shifted readings involve two instances of AGREE, one local (between *pro* and the embedded verb/probe) and one non-local (between the antecedent in the main clause and *pro*). As a consequence of these two instances, the ϕ -feature values of the second iteration of AGREE are superimposed on the first iteration. However, I don’t think this is the right way to view these cases. From the point of view of the theory I have proposed, such an analysis would involve erasing of an already specified ϕ -feature node, something that would predict overgeneration in other cases. Empirically, this solution would not explain why the morphology is the way it is: The embedded verb and subject (if overt) show signs of 1st person, not of 3rd person. Finally, if erasure of ϕ -features is involved, it is not immediately obvious why it is restricted to

⁸I must assume that V-to-C movement of the verb in German is not enough to identify a NS in that variety, although the equivalent configuration still does in Bavarian.

⁹Although the shifted reading has a similar interpretation to that of a direct quote in English (‘Tony said ‘I eat fish’’), its syntactic properties distinguish it both in Slave and in Amharic, another language where it has been observed, cf. Schlenker (2003); Baker (2008).

1st/2nd persons, so on this view, we would expect that an embedded 3rd person could be interpreted as correferential with a matrix 1st, but this does not seem to happen.¹⁰

For these reasons, I will suggest that the mechanisms underlying person shift in Slave are fundamentally distinct from AGREE. Baker's (2008) analysis is consistent with this conclusion (although his overall proposal is different in other respects). In his account, shifted readings arise indirectly through control. The embedded clause has a speech-act participant operator S (for speaker) or A (for addressee) that bind the embedded subject of the embedded clause, yielding a 1st and 2nd person interpretation respectively.¹¹ In shifted readings, the embedded S also binds the embedded subject, yielding a 1st person, but S is controlled by the subject of matrix clause, resulting in coreference, as shown in (68).



Thus, shifted readings are a consequence of control across clause boundaries, whereas as agreement is a locally bound type of AGREE.

¹⁰Note that while all the research on person shift readings goes to great lengths to point out that they are not cases of direct quotation, the interpretive patterns are very similar to direct quotation cases, thus a 1st person inside a quotation can be interpreted as referring to a 3rd person matrix subject, as in (i), but a 3rd person inside a quotation cannot be interpreted as referring to a 1/2nd person matrix subject (cf. (ii)). These common patterns remain unexplained if person shifting is completely independent from quotations.

(i.) Peter said “I am going away” (where I = Peter)

(ii.) #I said “He is going away” (where he = I)

¹¹In Baker's analysis, absence of binding by S or A results in 3rd person interpretation.

Chapter 6

Discourse identification

In this chapter, we turn to discourse identification of *pro*. Many researchers have pointed out that *pro* is topic-oriented (cf. Givón, 1983; Huang, 1984, Samek-Lodovici, 1996; Grimshaw and Samek-Lodovici, 1998; Frascarelli, 2007; Camacho, 2011; Sigurðsson, 2011). We will study two separate cases: Instances in which *pro* is identified by inflection and also by a topic, and cases in which *pro* is identified only by a topic. In order to understand the discourse conditions of *pro*, we will review Frascarelli's (2007) proposal for Italian. This will lead me to reassert Rizzi's double requirement of licensing and identification recast in terms of valuation and discourse linking. I will then present cases within NSLs where morphological valuation fails, and only topic linking applies. Then I will turn to languages with obligatorily overt subjects with limited cases of NSs (Germanic topic drop), and after that I review discourse NSLs like Chinese. Finally, I will adopt the analysis that one of the relevant differences between agreement-based NSLs and discourse-based NSLs relates to the specification of *pro*.

6.1 Topics and *pro*

A number of researchers have noted that *pro* is topic-oriented among many others). Samek-Lodovici (1996, 29) explicitly states that NSs must be licensed by topic antecedents and as evidence for this, he notes that a complement of a *by*-phrase cannot be an antecedent to a NS in Italian, Chinese, Hebrew or Greek, as seen in (1)-(2) for the first two languages. *By*-phrase complements are independently known not to be topics.

- (1) a. Questa mattina, la mostra è stata visitata da Gianni. (Italian)
this morning the exhibition has been visited by John
'This morning, the exhibition was visited by John'
- b. Più tardi, *pro*_{*i}/ egli_i / lui_i ha visitato l' università.
more late he / he has visited the university
'Later on, he visited the university.' (from Samek-Lodovici, 1996, 31, ex. 3)
- (2) a. Zuotian na yizhi beizi bei Lisi_i dapo le. (Chinese)
yesterday that one cup by Lisi break ASP
'Yesterday, that cup was broken by Lisi.'

- b. Jintan *pro*_i/ ta_i dapo le linwai yizhi.
 today / he break ASP another one.
 ‘Today he broke another one.’ (from Samek-Lodovici, 1996, 31, ex. 6)

Conversely, an overt, preverbal subject in a declarative sentence can serve as antecedent to a subsequent NS in all four languages (cf. (3) for Italian). In Hebrew and Chinese, an overt pronoun can also appear in that context (cf. (4)-(5)).

- (3) a. Questa mattina, Gianni_i ha visitato la mostra. (Italian)
 this morning, John has visited the exhibition.
 ‘This morning, John visited the exhibition.’
 b. Più tardi, *pro*_i/ ?egli_i / ??lui_i ha visitato l’università.
 More late, / he / he has visited the university
 Later, he visited the university.’
- (4) a. Ba-slos a be Juli ha-nasi_i xatam al ha-xoze ha-ze. (Hebrew)
 in-three in July the-president signed on the-contract the-this.
 ‘The third of July, the president’s brother signed this contract.’
 b. Lemoxorat *pro*_i / hu_i xatam al xoze xadas.
 the next day / he signed on contract new
 ‘The next day he signed a new contract.’ (from Samek-Lodovici, 1996, 33, ex. 9)
- (5) a. Zuotian, Lisi_i dapo le yizhi beizi. (Chinese)
 yesterday, Lisi break ASP one cup
 ‘Yesterday, Lisi broke a cup.’
 b. Jintian *pro*_i/ ta_i dapo le linwai yizhi.
 today he break ASP another one
 ‘Today he broke another one.’ (from Samek-Lodovici, 1996, 33, ex. 10)

We can reach the same conclusion from another angle by looking at cases where *pro* cannot be uniquely identified by morphology in agreement-rich languages. In such cases, the topic becomes essential for determining *pro*’s reference. For example, Cole (2000) (quoted in Sheehan, 2007, 158) notes that in (6) the subject can be null, because *abrí* ‘opened.1.SG’ identifies it as a 1st person. However, in the imperfect, *tenía* ‘had.1/3.SG’ is ambiguous between 1st and 3rd person. In that context, the antecedent becomes crucial for identifying the NS. When this antecedent is underdetermined, as it is in (6b) because it includes two conjoined DPs, the NS cannot be identified, hence an overt subject is required.

- (6) a. María y yo llegamos a casa. Yo/*pro* abrí la puerta. (Spanish)
 Maria and I arrived to home. I/*pro* opened1.SG the door.
 ‘Maria and I arrived home. I opened the door.’
 b. María y yo llegamos a casa. Yo/ella/**pro* tenía las llaves.
 Maria and I arrived to home. I/she/*pro* had.1/3.SG the keys
 ‘Maria and I arrived home. I found the keys.’ (from Sheehan, 2007, 84, ex. 158)

The same kind of situation is pervasive in the subjunctive in several languages, since agreement paradigms are impoverished. For example, Cardinaletti (1997) notes that the present subjunctive verb form in Italian does not distinguish between 1st/2nd/3rd person. In that context, 2nd persons cannot be null, as seen in (7). We will return to these cases below.

- (7) a. Che possa riuscir-ci non è chiaro. (Italian)
 that can.SUBJ.SG manage-CL not is clear
 ‘It isn’t clear that I/*you/he can manage it.’
 b. Che tu possa riuscir-ci non è chiaro.
 that you can.SUBJ manage-CL not is clear
 ‘It isn’t clear that you can manage it.’ (from Sheehan, 2007, 83 ex. 154-5)

These examples suggest that *pro* is crucially identified by a topic. Is topic identification a sufficient condition for NSs in all languages? Frascarelli (2007) argues that it is, but this is too radical a conclusion, since reducing NS identification to the availability of a topic cannot account for the contrast between (6a) and b), where the only difference is the inflectional paradigm of the imperfect vs. the preterit. Thus, I suggest two conditions on NSs, as in (8).

- (8) Conditions on referential categories
 a. Referential categories must be contentful.
 b. Referential categories must be discourse-linked.

Overt nominal categories comply with (8a) automatically, whereas their corresponding null counterparts do so through valuation as described in the previous chapters. We can think of this condition as a prerequisite for a referential category to be discourse-linked. The second condition is intrinsically connected to topic linking, as suggested earlier. We will now turn to the specific details of topic linking in Frascarelli’s proposal.

6.2 Typology of Topics

Frascarelli (2007) builds on the observation that NSs are usually topic-oriented, and proposes that *pro* is always bound by a topic. Her analysis is partly based on the discourse distribution of NSs and partly on the prosodic properties associated with them. Based on these, she proposes that NSs are bound by an Aboutness Shift Topic, which has an intonational contour L*+H. These elements introduce a new topic or shift from a previous one. For example, in (9), the phrase *l’ultima unità* ‘the last unit’ is a new topic introduced after a general description of the materials of a self-learning course, and it is CLLD.

- (9) Il materiale era tantissimo quindi all’ inizio l’ ho fatto tutto di corsa cercando di impiegarci il tempo che dicevate voi magari facendolo un po’ superficialmente pur di prendere tutto - l’ultima unità la sto facendo l’avevo lasciata un po’ da parte . . . (Italian)
 ‘The material was quite a lot, so at the beginning I did it all in a rush, trying to do it in the time that you had fixed, perhaps a little superficially, so as to do everything - I’m doing the last unit now, I had put it aside before . . .’ (from Frascarelli, 2007, 698, ex. 8)

In addition to Aboutness Shift Topics, Frascarelli (2007) distinguishes two other kinds: Contrastive topics and Familiar topics. The first type has a H pitch, and provides an alternative to an existing topic, as seen in (10), where *a lei* ‘to her’ presupposes that alternative possibilities for the speaker position were available in discourse.

- (10) Invece a lei non l’ ha presa come speaker. (Italian)
 instead to her not CL.3.SG have.3.SG taken.FEM as speaker
 ‘On the contrary he didn’t choose her to be the speaker.’
 (from Frascarelli and Hinterhölzl, 2007, 96, ex. ix)

Finally, a Familiar topic “refer[s] to given information in the discourse” (Frascarelli, 2007, 699), it is D-linked (i.e. it relates to referential sets pre-established in discourse), and it is “used to refer to background information, for topic continuity (Givón 1983) or in the right periphery, as an ‘afterthought’ (p. 699). Thus, in (11), *la conferma* ‘the check’ is first introduced by speaker A, then resumed by speaker B and repeated as a right-hand topic by student A.

- (11) A: Io dovevo studiare le regole qui e lì fare solo esercizio, invece mi aspettavo di trovare dei punti a cui far riferimento ogni volta per vedere la regola, questo mi è mancato praticamente: **la conferma** di ricordare tutto insomma. (Italian)
 ‘I was supposed to study the rules here and do the exercises at home, while I expected to find some outlines I could refer to, at any point, to check the relevant rule, this is what I missed: the check that I could remember everything.’
 B: Comunque quelle domande ti davano **la conferma** che avevi capito.
 however those questions gave you **a check** for your understanding.’
 A: Ma . . . magari non me la- non riesco a darmela da sola **la conferma**.
 ‘A well, maybe I cannot do **this check** on my own.’
 (from Frascarelli, 2007, 699, ex. 9)

Frascarelli (2007) proposes that these three topics are projected in an expanded C-domain, following Rizzi (1997) and subsequent work. Each of these topics is a based-generated CLLD phrase, associated with a clitic (cf. Cinque, 1990).

- (12) [_{ForceP} [**ShiftTopP** [_{GroundP} [**ContrastTopP** [_{FocP} [**FamTopP** [_{FinP} . . .]]]]]]]]]

As mentioned, *pro* is assumed to be bound by a local Aboutness Shift topic or a silent local copy of it. As an example, the antecedent of the first *pro_i* in (13) is *il mio capo* ‘my boss’, the Aboutness Shift Topic. If this topic is maintained from clause to clause, a silent copy of it appears in FamP and locally binds *pro*, as in (14) (cf. Frascarelli, 2007, 709).

- (13) [il mio capo]_i come diceva Carlo . . . *pro_i* è un exreporter . . . *pro_i* è stato in giro per il mondo . . . (Italian)
 ‘My boss as Carlo used to say . . . *pro_i* is a former reporter . . . *pro_i* has been all over the world.’

- (14) [_{AboutShift}TopP [il mio capo]_i . . . [_{AgrSP} [_{vP} *pro*_i è un exreporter]] [_{FamP} Top_i [_{AgrSP} [_{vP} *pro*_i]]]]

Let us now turn to instances in which morphological valuation fails in an otherwise consistent NS language. In such cases, as predicted, discourse identification is the only available option.

6.3 Topic-identification in an inflection-rich NSL

Although Spanish and Hebrew typically have NSs valued by inflection, in both languages certain adjunct clauses have NSs not valued by inflection.

6.3.1 Absolute clauses in Spanish

Absolute clauses in Spanish (cf. Pérez Jiménez, 2007; Camacho, 2011 among others) illustrate a case where NSs are not valued by verbal inflection. In the current framework, the probe lacks the minimal ϕ -feature specification (π).

Absolute clauses are adjunct clauses that consist of a predicate of any type (a participial, a PP, an AdvP or a DP) and a subject, and they obligatorily appear in the order Pred-S. As pointed out in Camacho (2011), since the subject can be null (cf. (15)), the question arises as to how it can be identified, since the predicate lacks person and potentially number values, as in the case of (15b), where the adverb *lejos* lacks person, number and gender features.

- (15) a. Incómodo por el incidente, Jaime evitó encontrarse con su jefe. (Spanish)
 uncomfortable by the incident, Jaime avoided run into with his boss
 ‘Uncomfortable about the incident, Jaime avoided running into his boss.’
 (from Camacho, 2011, 990,
 ex. 3c)
- b. Lejos ya en el tiempo, la guerra perdió importancia en la vida diaria
 far already in the time, the war lost importance in the life daily
 ‘Once far away in time, the war lost relevance in day-to-day life.’

As Camacho (2011) points out, absolute-clause NSs show common properties with *pro* and *PRO*. Specifically, like *pro*, they alternate with overt subjects, they are Case-marked in the way overt subjects are and they allow for split antecedents. Like obligatorily controlled *PRO*, they must have an antecedent, and it must be local, as seen in (16): In (16a), the feminine gender of the adjunct-clause predicate renders the main-clause subject unavailable as an antecedent, so the NS is not recoverable. The fact that the clitic is also unavailable as an antecedent illustrates another property: The antecedent must be a topic, and the clitic cannot serve as one. This contrasts with a doubled dative DP appears, which can be a topic, as seen in (15b). Finally, like controlled *PRO*, absolute-clause NSs have obligatory *de se* readings.

- (16) a. *Cansada por la caminata, Miguel le sirvió un café. (Spanish)
 tired.FEM by the walk, Miguel CL served a coffee.
 ‘Tired by the walk, Miguel served her coffee.’

- b. Cansada por la caminata, Miguel le sirvió a Ana un café.
 tired.FEM by the walk, Miguel CL served to Ana a coffee.
 ‘Tired by the walk, Miguel served Ana coffee.’

Since the MMT for NSs in Spanish is set to [PART] and probes for absolute-clause predicates lack that specification, this means that *pro* will not be valued by the probe. However, if *pro* is linked to the topic antecedent, then that process establishes a chain between the topic, *pro* and the probe, yielding agreement. The topic can, in turn, be null, but since the topic is located at the edge of the CP-phase, it can be discourse-linked to an antecedent, as in (17).

$$(17) \left[\text{CP}_{\text{main}} \dots \phi \dots \left[\text{CP}_{\text{abs}} \left[\text{TopP} \text{Top}_i \left[\text{TP} \text{T}_i \text{pro}_i \dots \right] \right] \right] \right]$$

\uparrow ----- \uparrow ----- \uparrow ----- \uparrow

If Camacho’s (2010) analysis of absolute-clause NSs is correct, it yields three interesting conclusions: First, even in an agreement-rich language like Spanish, there are instances of NSs that are not identified by rich agreement, but by a topic antecedent. Second, this NS-antecedent relationship, which seems to be long-distance, can be reduced to shorter, phase-bound links, all subsumed under the valuation version of AGREE. Third, it confirms Cole’s (2009) intuition that agreement recoverability and topic recoverability are parts of a continuum.

6.3.2 Adjunct clauses in Hebrew

Shlonsky (2009, 146-151) discusses a fairly similar situation in Hebrew (cf. also Borer, 1989; Gutman, 2004. In (18), the embedded NS “is obligatorily dependent on a nominal in the matrix clause” (p. 146).

- (18) Dani₁ kibel mi-Dafna ve-Rina matana yafa axarei še pro_{1/*2}
 Dani received-3.MAS.SG from-Dafna and-Rina present pretty.FEM.SG after that
 siyem et ha doktorat. (Hebrew)
 finished.3.MAS.SG ACC the doctorate
 ‘Dani received a fine present from Dafna and Rina after he had finished his doctorate.’
 (from Shlonsky, 2009, 146, ex. 18)

Like in Spanish, the embedded NS in these clauses can have several possible constituents as antecedents, and according to Shlonsky (2009, 148) “Gutman [Gutman, 2004, J.C.] views the accessibility of the antecedent is a necessary condition for the grammaticality of sentences with covert subjects. One of the factors that determine accessibility is the discourse salience of the antecedent.” Thus, like in Spanish absolute clauses, a main-clause passive agent is not a good antecedent for the embedded NS. Shlonsky (2009, 149), following Borer (1989), assumes that these cases of non-standard binding/control entail assigning a [person] feature to the NS. This feature, which makes the NS referential, is copied to T’s ϕ matrix through a mechanism he calls feature synchronization. Synchronization is possible because matrix T and the embedded NS are in a probe-goal relation.

Regardless of the specific mechanics of each proposal, the configurations of Spanish absolute clauses and Hebrew non-standard binding control are very similar. In both languages, control by

a discourse-salient antecedent is obligatory and in both cases valuation in the embedded domain is not possible because inflection lacks the relevant specification (in Hebrew, 3rd persons do not license NSs in general).

6.3.3 Subjunctive clause NSs

As we saw earlier, several tenses across the Spanish or Italian conjugation do not distinguish between two or more persons. Consider, for example, the present subjunctive or the imperfect indicative in Spanish, where 1st and 3rd person verb endings are identical, and 2nd person is distinct in the singular, as seen in (19).

- (19) a. Spanish present subjunctive:
 consiga ‘manage.1/3.SG.SUBJ’ *consigas* ‘manage.2.SG.SUBJ’
 b. Spanish imperfect indicative:
 conseguía ‘manage.1/3.SG.IMP’ *conseguías* ‘manage.2.SG.IMP’

Assuming that 1st/3rd persons are represented identically, and they should lack [PART], as in (20). If so, *pro* will remain unvalued for [PART] in its clause. Nevertheless, 1st/3rd NSs are possible even in out-of-the-blue contexts, such as as (21a)-(22).

- (20) Spanish singular present subjunctive ϕ -feature geometry

1 st /3 rd p.	2 nd p.
R	R
D	D
	... π
	PART
	ADDR

- (21) a. Que lo consiga no es claro. (Spanish)
 that CL manage.1/3.SG.SUBJ not is clear
 ‘It is not clear that I/he/she will manage it’
 b. Que lo consigas no es claro.
 that CL manage.2.SG.SUBJ not is clear
 ‘It is not clear that you will manage it’
- (22) Durante ese verano, cada día salía a buscar el periódico. (Spanish)
 during that summer, each day went.out.IMP to look.for the newspaper
 During that summer, I/he/she went out to look for the newspaper every day.’

The question is how is *pro* discourse-linked in (21a) and (22)? The proposal is that both 1st and 3rd persons are derived by default. In the case of 3rd person, it is frequently the default morphological setting, for example in cases of agreement asymmetries. Regarding 1st person, a number of researchers, 1st and 2nd persons are syntactically represented as syntactic heads in the left-periphery (cf. Poletto, 2000; Speas, 2004; Sigurdsson, 2004; Sigurðsson, 2010, 2011; Sigurðsson and Holmberg, 2008; Baker, 2008), which bind the corresponding pronouns. Safir (2004), on the other hand, argues against this view, and suggests that 1st and 2nd persons pick out an individual in the context through a constant function, rather than through binding. I will assume that 1st persons are subject to a constant function, contrary to Safir (2004, 138), I will not extend this proposal to 2nd persons. Whereas the 1st person can always be contextually recovered (i.e. the speaker), the 2nd person, on the other hand, cannot always be, because the reference of the 2nd person addressee requires some pointing gesture. Consider, for example, the situation described by Safir (2004, 138) in which “a squad leader is assigning tasks to his or her assembled group, s/he might point to each member saying *you do this, you do that.*” In that context, if the squad leader walks into the room looking at the ceiling and says, without any gesture, *you do that*, the addressee is not uniquely fixed, but if s/he says *I’ll do that*, there is no question as to who the speaker is. Turning to the examples above, the 1st person interpretation comes from the constant function, which can be seen as a last-resort operation that allows the underspecified [D] output of (20) to be interpreted as a 1st person.¹

Extending this analysis to cases where underspecification affects all three persons, such as the Italian present subjunctive, reveals some interesting properties. Cardinaletti (1997, 51) notes that in out-of-the-blue contexts, a NS can be interpreted either as 1st or 3rd person, but not as 2nd person, as seen in (23). This is because in that particular tense all three singular person endings are identical.

- (23) a. Che possa riuscir-ci non è chiaro. (Italian)
 that can.SUBJ.SG manage-CL not is clear
 ‘It isn’t clear that I/*you/he can manage it.’
 b. Che tu possa riuscir-ci non è chiaro.
 that you can.SUBJ manage-CL not is clear
 ‘It isn’t clear that you can manage it.’
 (from Cardinaletti, 1997, 51, ex. 66, quoted in Sheehan, 2007, 83)

The proposed analysis for the Spanish cases above carries over fairly well to these cases: Inflection does not value *pro*, hence either the constant function adopted from Safir (2004) assigns a 1st person interpretation, or the default 3rd person interpretation comes into play, but nothing yields the 2nd person interpretation.

¹As we will see below, topic drop in Swedish (cf. (i)) tends to be identified as 1st person by default, cf. Sigurðsson (2011, 279).

- (i.) Kommer strax. (Swedish)
 come right.away
 ‘(I’ll) be there in a minute. Sigurðsson, 2011, 279, ex. 23b

In this context, it is worth pointing out that even contextual saliency does not allow for 2nd person interpretation in Italian present subjunctive examples, as noted by Rizzi (p.c., quoted by Sheehan 2007, fn. 40). Thus, in (24) the second clause introduces a NS that is clearly identified by the verb *hai* ‘have.2.SG.’ as 2nd person. However, this is not enough to interpret the subjunctive clause’s subject as referring to a 2nd person as well.

- (24) #So [che hai provato] ma non è facile [che possa riuscir-ci] (Italian)
 know that have.2.SG tried but not is easy that can.SUBJ succeed-CL
 ‘I know that you’ve tried but it’s not going to be easy for you to succeed.’
 (from Rizzi, p.c., quoted in Sheehan, 2007, 83, fn. 40)

This is a fairly robust effect, as the following examples show (thanks to Vieri Samek-Lodovici for judgements, glosses and translations are my own). Thus, regardless of whether the potential antecedent is an overt or null subject pronoun (cf. (25)-(26), preverbal or postverbal (cf. (25a) vs. b)) or a CLLD object or a topic ((cf. (27)), a 2nd person interpretation is not possible.

- (25) a. *So che tu hai provato ma non è facile che possa riuscir-ci. (Italian)
 know that you have2.SG tried but not is easy that can.SUBJ.SG succeed-CL
 ‘I know that you have tried but it’s not going to be easy for you to succeed.’
- b. *So che (ci) hai provato tu, ma non è facile che possa riuscir-ci
 know that CL have2.SG tried you but not is easy that can.SUBJ succeed-CL
 ‘I know that you have tried but it’s not going to be easy for you to succeed.’
- (26) a. *So che (tu) vuoi andare ma non credo che possa far-lo. (Italian)
 know that you want go but not think.1.SG that can.SUBJ.SG do-CL
 ‘I know that you want to go but I don’t think that you will be able to.’
- b. *So che (tu) provi spesso a chiamar-mi, ma non sembra che riesca a trovar-mi
 know that you try frequently to call-CL but not seems that succeed.SUBJ find-CL
 ‘I know that you frequently try to call me but it doesn’t seem that you succeed to find me.’
- (27) a. *A te, non te l’ ho mai detto, ma non è facile che possa
 to you, not CL.2 CL.3 have ever told, but not is easy that can.SUBJ
 riuscir-ci. (Italian)
 succeed-CL
 ‘You, I have never told you this, but it’s not easy for you to succeed.’

- b. *Tu, fra tutti noi, so che avresti desiderato molto aiutar-lo, ma non
 you, among all us, know.1.SG that would-have wanted much help-CL, but not
 credo che possa riuscir-ci.
 believe that can.SUBJ succeed-CL
 ‘You, among all of us, I know that you would have very much wished to help him, but
 I don’t think that you can succeed.’

Spanish is partially similar, in the sense that (21a) above (with a morphologically ambiguous 1st/3rd verb ending) cannot be interpreted as if it were the polite form of talking to an addressee (corresponding to the pronoun *usted* ‘you-formal’). Thus, (21a) cannot mean the same as (28).²

- (28) Que usted lo consiga no es claro. (Spanish)
 that you-formal CL manage.1/3.SG not is clear
 ‘It is not clear that you will manage it.’

This confirms that there may be something special about the way an addressee is picked in discourse, whether it is morphologically 2nd or 3rd person. However, Spanish and Italian differ because Spanish is much more liberal when it comes to identifying an addressee by a salient antecedent, so that (29) is acceptable, by contrast to the Italian examples in (25)-(27) above.

- (29) Se que usted lo ha intentado, pero no creo que lo
 know.1.SG that you-formal CL have.3.SG tried but not think.CL that CL
 consiga. (Spanish)
 manage.1/3.SG
 ‘I know that you have tried it, but I don’t think that you will manage it.’

6.3.4 Topic drop

Topic drop is a phenomenon in which a subject or an object can be dropped provided that the Spec, CP is null (cf. Sigurðsson, 2011). The phenomenon, which occurs in a range of Germanic languages, is not generally contingent on verb agreement, although “[e]ven within Germanic, however, agreement *constrains* identification” (Sigurðsson, 2011, 279). Thus, although both Swedish and Icelandic have topic drop with similar properties, in Swedish (without agreement), null topics can be construed as different persons depending on the context, whereas in Icelandic (with agreement), they are constrained by agreement. Dropped topics in Swedish tend to be interpreted as 1st person, so example (30) would typically be interpreted as ‘I am just lying on the beach’, but in the context of a question like *where is Anna?*, the interpretation would be ‘she is just lying on the beach’.

- (30) Ligger bara på stranden. (Swedish)
 lie just on beach.the
 ‘Just lying on the beach.’ (from Sigurðsson, 2011, 280, ex. 23a)

²Thanks to Liliana Sánchez (p.c.) for pointing this out.

Compare the Swedish example with the corresponding translation in Icelandic. First, there are four different renderings of (30) in that language, as shown in (31). Needless to say, each of those examples can only be interpreted in a way consistent with the agreement on the verb.

- (31) a. Ligg bara á ströndinni. (Icelandic)
 lie.1.SG just on beach.the
 ‘I am lying on the beach.’
- b. Ligg-ur bara á ströndinni.
 lie-3.SG just on beach.the
 ‘S/he is lying on the beach.’
- c. Liggj-um bara á ströndinni.
 lie-1PL just on beach.the
 ‘We are lying on the beach.’
- d. Liggj-a bara á ströndinni.
 lie-3PL just on beach.the
 ‘They are lying on the beach.’ (from Sigurðsson, 2011, 280, ex. 25)
- (32) a. (Ich) kenne das nicht. (German)
 (I) recognize that not
 ‘(I) do not recognize that.’
- b. *Jetzt kenne das nicht.
 now recognize that not
 ‘Now (I) do not recognize that.’ (from Sigurðsson, 2011, 271, ex. 8a, 9a)

In Sigurðsson’s analysis topic drop is possible because the null category raises to the CP-area, and it is “C/Edge linked” by either by a Topic operator or by a 1st/2nd person logophoric operator, or by both.³

In the current analysis, topic drop is also a case of discourse-identification, although the contrast between Icelandic and Swedish, for example, follows from the fact that inflection values the NS in Icelandic but not in Swedish. Nevertheless, the MMT in both languages implies that valuation is not enough to license NSs. As a result, both NSs must be discourse-linked, but only Icelandic constrains the null topic depending on verbal agreement.⁴

6.3.5 Constraining the system

One immediate concern that arises with the proposed system is how to constrain the availability of NSs. On the one hand, if discourse linking can supersede lack of agreement valuation, then why isn’t this option generally available in all languages? Second, what are the locality restrictions that regulate discourse-linking? The answer to the second question will have to wait until I

³Although Sigurðsson assumes that agreement constraints topic drop, as noted above, it is not clear how this is formally done, given that weight of the 1st/2nd person interpretation is carried by the logophoric operator. Presumably, a NS linked to a 1st-person operator must also match the verb’s morphology (if it has one).

⁴The fact that the Spec, CP must be null, can follow if the null category must be moved to that position.

present the account of consistent discourse-based NSLs. I do not have a clear answer to the first question, although I have several observations to make. First, agreement valuation is always local, i.e. it is constrained to the clause containing the subject and Infl. By contrast, discourse-linking is typically extra-clausal. Even within extra-clausal discourse-linking, two subtypes, depending on locality. Absolute clauses in Spanish, Hebrew non-standard control have mid-range locality, i.e. the antecedent is typically in the adjacent clause, whereas in other cases, a NS can skip a clause, as seen in (33), where the antecedent can be *Maria*, *Pedro* or even *I*.

- (33) A María le molesta que a Pedro le guste que *pro* escriba en su pared del
 to Maria CL bothers that to Pedro CL pleases that write on POSS wall of-the
 Facebook. (Spanish)
 Facebook
 ‘It bothers Maria that Pedro likes that s/he/I write on his/her Facebook wall.

There are possible explanations for the mid-range cases exist. One is that the set of constructions that require a mid-range antecedent share some structural property that allows the domain to be extended. The notion of an extended domain can be due to a number of possible factors. The two most natural candidates are a defective C-system, or a defective T-system. As noted, the adjunct clauses in many of these cases lack inflectional information, and are temporally dependent on the main clause. However, there does not seem to be a one-to-one correlation between temporal defectivity and other symptoms of phase-extension. Rather, it would seem that the embedded clauses that show this mid-range control property form a unit with the main clause. Temporal dependencies are a manifestation of this fact, and binding-domain extensions are another. Suppose, then, that they share a paragraph-level discourse layer of functional projections where the topic is uniquely identified. Thus, in addition to clause-level phases, we can define the notion of paragraph-level phases, which include a main clause, its adjuncts and any embedded clauses.⁵

Consider, for example, the fact that successive-cyclic movement can range over several embedded clauses but cannot cross a paragraph-level phase in the sense just defined. Thus, whereas the *wh*-word can be an argument of *buy* in (34a), crossing several successively embedded clausal boundaries, the same is not possible in (34b), when one attempts to move a *wh*-word from one main clause to another main clause.

- (34) a. What did Mary hear [that say Bill said [that Susan bought]]? (English)
 b. *What did Mary go to the mall. Did Peter buy?

A similar kind of argument can be made when considering exceptions to the coordinate structure constraint, noted by Lakoff (1986). These tend to be possible whenever the two conjuncts are interpreted as a close unit (causally or temporally linked). Thus, whereas in (35a) the two conjuncts are interpreted as temporally linked and the second one can be seen as a result of the first one, in (35b), this is not the case, and extraction is not possible.

- (35) a. Which shirt did John go to the store and buy? (English)

⁵Shipibo NSs have a distribution that is sensitive to this notion of a paragraph-level phase, cf. sect. 3.4.2.

- b. *Which shirt did John go to his mother's house and also buy?

Assuming the notion of a paragraph level phase that has a single topic, then we can delimit the antecedent for NSs to the closest topic, which will be the paragraph-level one.

6.4 Subject identification in discourse-NSLs

As mentioned in sect. 2.3.2, a number of languages allow for NSs and do not have verbal morphology to identify them. For Chinese, the traditional analysis dating back to Huang (1984, 1989) assumes that NSs are identified in discourse. However, as the following example shows, topic-binding does not seem to be the right notion. In this sentence, the embedded *pro* is interpreted as coreferential with the matrix subject, not the topic.

- (36) Lisi_i, Zhangsan_j ku de [*pro*_{*i/j} hen shangxin] (Chinese)
 Lisi Zhangsan cry till very sad
 'Lisi, zhangsan cried until he got very sad.' (from Huang, 1989, 198, ex. 41b)

This type of example led Huang (1989) to propose an alternative account based on the notion of control, which he extends to Italian-type NSs. Within this alternative, a null pronominal is controlled in its domain, and if it lacks one, then it may either be identified by the closest c-commanding antecedent in a preceding clause, or be interpreted arbitrarily, or be pragmatically controlled. Technically, a control domain is the S/NP containing the null category and an accessible SUBJECT. So, for example, in (37), the control domain is the minimal sentence that contains the NS, because AGR acts as an accessible SUBJECT. In (38), on the other hand, there is no AGR, hence no accessible SUBJECT, so the minimal clause does not constitute a control domain. In (39), on the other hand, the matrix clause does contain an accessible SUBJECT *Zhangsan*, which becomes the antecedent for *pro*.

- (37) [_S *pro* verra] (Italian)
 will come
 'S/he will come.' (from Huang, 1989, 195, ex. 29b)
- (38) [_S *pro* lai le] (Chinese)
 came ASP
 '[he] came.' (from Huang, 1989, 195, ex. 30)
- (39) [Zhangsan qi ma qi de [*pro* hen le] (Chinese)
 Zhangsan ride horse ride till very tired
 'Zhangsan rode a horse until he got very tired.' (from Huang, 1989, 197, ex. 40a)

In preverbal adjunct clauses, the embedded subject need not be controlled by a matrix antecedent. In Huang's analysis, this follows because the preverbal adjunct clause lacks an accessible SUBJECT, presumably because the clause is not c-commanded by the subject position in the matrix clause.

- (40) [Ruguo *pro*_{i/j} bu lai], ta_i keneng hui shengqi. (Chinese)
 if not come he possibly will angry
 ‘If we/you/he . . . doesn’t come, he will probably be angry.’
 (from Huang, 1989, 198, ex. 42b)

Chinese displays a subject-object asymmetry with respect to topicalization from within an island, as noted by Huang (1982) (highlighted by Li (2007)). Topicalization is generally constrained by syntactic islands, but examples like (41a) seem to be an exception, since topicalization from the relative- clause subject-position is possible. By contrast, topicalization from the object position is not possible, as seen in (41b).

- (41) a. Zhangsan_i, [[\emptyset _i xihuan de shu] hen duo]. (Chinese)
 Zhangsan like DE book very plenty
 ‘Zhangsan_i, the books [he_i] likes are many.’
 b. *Zhangsan_i, wo kan-guo [\emptyset _i xihuan de shu].
 Zhangsan I see-GUO like DE book
 ‘Zhangsan_i, I see the books [he_i] likes.’ (from Li, 2007, exs. 5, 7)

Huang’s analysis traces this asymmetry to his Generalized Control Rule (GCR). In the first example, the NS is *pro*, identified within the matrix clause by the closest c-commanding antecedent (the topic). In the second example, on the other hand, the closest c-commanding antecedent is the embedded subject, so it cannot be a resumptive pronoun for the topic. In other words, (41a) is not an extraction across an island boundary, whereas (41b) is.

Li (2007) points out that this analysis correctly predicts the possible range of interpretations of null subjects inside islands. In particular, if the NS is not co-indexed with the subject of the higher clause (the closest c-commanding NP), the sentence is ungrammatical, as seen in (42).⁶

- (42) Wo_i yinwei [\emptyset _{i/*j} bu xihuan Zhangsan] you diar shiwang/bu-hao-yisi. (Chinese)
 I because not like Zhangsan have slight disappointment/embarrassment
 ‘I am somewhat disappointed/embarrassed because I/*he/*she does not like Zhangsan.’
 (from Li, 2007, ex. 23)

In terms of the analysis advanced here, both Chinese and Japanese are languages where a NS is not recovered morphologically. As mentioned, in Chinese, it is controlled by the closest c-commanding antecedent in the control domain. How does the notion of control domain translate

⁶Huang’s analysis has been questioned on the basis that control is not always obligatory, as shown in (i). In this example, the control domain is the main clause, but the embedded subject can have an alternative antecedent. As Audrey Li (p.c.) points out, *say*-type verbs (which can take quotations), can have the NS refer to a discourse topic if the context is very clear and strongly favors the interpretation. Cf. also Huang (1992) for other arguments against J-T. Huang’s analysis.

- (i.) Xiaohong de meimei shuo \emptyset xihuan tan gangpin. (Chinese)
 Xiaohong GEN younger-sister say like play piano
 ‘Xiaohong_i’s younger sister_j says that (I/you/he/she_{i/j/k}/we/they) like(s) to play the piano.’
 (from Y. Huang 2000, quoted in Cole, 2009, 560, ex. 1)

into current terminology? In the preceding section, I argued for a paragraph-level unit to account for cases of mid-range control/binding. This notion can translate the concept of a control domain in a straightforward manner. The only case where results are unexpected is (40), where the paragraph-level phase should include the adjunct clause and the main clause, so topic control should be obligatory, but it is optional. Assume, however, that the topic does not c-command the adjunct clause. If so, we expect the reference of the NS to be fixed through pragmatic control.

The proposed analysis also accounts in a straight-forward way for an apparent counterexample to the GCR (cf. Huang, 1989). In (43), the GCR predicts that the null object should seek the subject of the embedded clause as an antecedent, since it is the closest c-commanding NP. However, it is interpreted as coindexed with the topic. In order to account for this apparent by-passing of the first c-commanding antecedent, Huang proposes that the null object topicalizes within the embedded clause, as in (44). We can assume the same type of proposal, although I assume that the position to which *pro* moves is not a topic position.⁷

- (43) Zhe-ben shu_i, [[Lisi kan Ø_i] zui heshi]. (Chinese)
 this-CL book Lisi read most appropriate
 ‘This book, for Lisi to read [it] is most appropriate.’ (from Li, 2007, ex. 10)

- (44) Topic_i [CP_{main} [I_{sland} *pro*_i Lisi_j . . . t_i] . . .]
 ↑-----↑

An alternative compatible with the proposed analysis is Zushi’s (2003), who suggests that Japanese and Chinese allow for a base-generated zero topic. Let us assume, then, that a null pronominal in Chinese must be identified in the paragraph-phase that contains it, by the first c-commanding NP in that phase (typically the subject or topic).

6.4.1 Two types of nominals

A question that comes up at this point is whether crosslinguistic variation is simply a matter of how a NS is identified (agreement vs. discourse-only), or whether the structure of NSs in each language also plays a role. From the analysis developed so far, nothing seems to distinguish them in a principled way. In particular, Chinese or Japanese are the way they are because they lack verbal morphology, hence NS identification can only be done through the discourse. However, as we pointed out earlier, Tomioka (2003) has advanced the idea that perhaps there is a principled distinction between null categories in both types of language, which relates to the structure we observe in overt nominals. Specifically, following Chierchia (1998), languages like Japanese have NPs, but languages like Spanish or Italian have DPs. Correspondingly, according to Tomioka, null categories in Japanese are NPs.

While this proposal is interesting, extending it to other discourse-related languages presents a challenge. Li (2007) points out that Japanese and Chinese have different behaviors with respect

⁷The GCR applies not only to subjects, but also to null objects. In the case of objects, it predicts that they should not be possible as pronominals, since the first c-commanding antecedent in the control domain would be the subject of their clause, yielding a Principle B violation.

to the distribution and interpretation of null categories. Specifically, she argues, overt nominals in Chinese show properties typical of DPs that are not present in Japanese. For example, nominals exhibit a fixed word order of [Demonstrative+Classifier+N] without the marker *de*, whereas all other modifiers have quite free word order and are followed by *de*. In contrast, Japanese demonstratives and modifiers have a similar distribution as other modifiers in the sense that the Japanese counterpart of *de* must occur with demonstratives and classifier phrases:

- (45) a. ko-no/ so-no/ a-no (Japanese)
 this-NO/ that-NO/ that-NO
 b. *zhe-de/ na-de (Chinese)
 this-DE/ that-DE (from Li, 2007, ex. 43)
- (46) a. san-satu-no hon/ san-nin-no gakusei (Japanese)
 three-CL-NO book/ three-CL-NO student
 b. *san-ben-de shu/ san-ge-de xuesheng (Chinese)
 three-CL-DE book/ three-CL-DE student (from Li, 2007, ex. 43)

For Li, this suggests that nominals have a DP-like structure in Chinese but not in Japanese (cf. Tomioka, 1999, 2003). Assuming that the typology of null categories is the same as that of overt ones in an given language, this means that the corresponding null categories would presumably be D and N respectively (cf. Fukui, 1988; Hoji, 1998; Noguchi, 1997; Tomioka, 2003).

The different settings for *pro* in Chinese and Japanese suggests an explanation for another difference between both languages, as Li (2007) points out. She observes that Japanese lacks true sloppy interpretations in ellipsis (cf. Hoji, 1998), whereas Chinese has them: If the null category in ellipsis is pronominal in both of these languages (as argued for by Hoji, 1998 for Japanese), then the sloppy reading asymmetry can follow from *pro*'s different category in each language.

Assuming that this analysis is correct, the representation of pronouns would be the ones in (47). However, this means that Japanese pronouns (and nominals in general) are not referential in the sense of (8), repeated in (48), rather their reference is achieved through a different mechanism (along the lines of Chierchia, 1998).⁸

- | | | |
|------|-----------------------|------------------------|
| (47) | Chinese null pronouns | Japanese null pronouns |
| | R | R |
| | D | N |

- (48) Conditions on referential categories
- a. Referential categories must be contentful.
 - b. Referential categories must be discourse-linked.

⁸Chierchia actually treats Chinese and Japanese as belonging to the NP language-type.

One final consequence of this parametric difference, according to Li (2007), is that while Chinese null categories are regulated by the GCR, Japanese ones are not. For example, NSs need not be bound by the first c-commanding antecedent, as illustrated in (49a). By contrast, the NS in Chinese in (49b) cannot skip the subject to seek the topic.

- (49) a. Bush-wa, maikeru muua-ga senkyo-ni katta node
 Bush-TOP Michael Moore-NOM election-DAT won because
 gakkari siteita. (Japanese)
 disappointed
 ‘Bush, Michael Moore was disappointed because (he) won the election.’
 (from Li, 2007, ex. 40)
- b. *(Lisi₁,) wo yinwei [e₁ hai bu renshi naxie ren] hen danxin. (Chinese)
 Lisi I because still not know those people very concerned
 ‘(Lisi,) I am very concerned because Lisi still does not know those people.’
 (from Li, 2007, ex. 41b)

These proposals entail that Japanese *pro_N* is not subject to valuation or agreement, since it lacks the node that triggers such mechanisms, whereas Chinese *pro_D* is. However, Chinese differs from Spanish or Italian in one crucial respect, that is the availability of another type of null category in Chinese. One of the consequences of Huang’s analysis is that Chinese null objects cannot be pronominal. Since they would be subject to the GCR, the first c-commanding antecedent in the control domain would be the subject, so they would have to be interpreted as coreferential with the subject, i.e. as anaphors. Given that they need not be interpreted as anaphors, then they must either be some other category or be interpreted by some other mechanism. Li (2007) argues for the former option, suggesting that the null objects in Chinese are a type of True Empty Position (TEP), which is a last-resort empty element necessary for subcategorization purposes, but devoid of any content and whose interpretation is done by purely pragmatic means. TEPs can be interpreted as pronouns but also as other categories, for example in the context of ellipsis. This is possible because they lack content. In languages like Spanish, these categories do not seem to be available. In terms of our analysis, TEPs would correspond to a feature geometry that only has an [R] node.

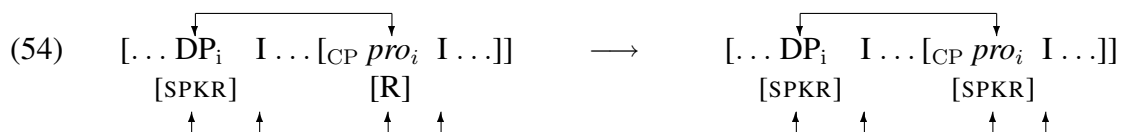
6.4.2 Shipibo and Capanawa

Shipibo is a particularly interesting example because it displays person asymmetries but lacks morphological agreement, as we already described in section 3.4. To summarize again, 1st and 2nd person subjects are obligatorily overt but 3rd person ones are optional in main clauses. In embedded, switch-reference clauses, both are optional. This is illustrated in the paradigm in (50), where the SR clause appears in first position, indicated by brackets. As these examples show, a 1st person subject can be overt in both clauses (cf. (50a)), or null in one of them (cf. (50b-c)), but not in both of them. Thus, (50d) cannot be interpreted as the NS counterpart of (50a-c). However, this example can be interpreted as a third person.

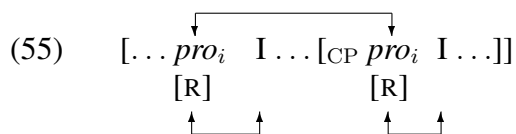
Third, the language distinguishes between mass and count nouns. Thus, according to Valenzuela (2003b, 204), “[n]on-count nouns do not combine with numerals and cannot take the plural *-bo*.” Finally, Shipibo lacks a classifier system. The first property would group it with NP languages, but the other three would suggest it is a DP-like language. Additionally, demonstratives seem to be able to function as definite article in certain cases and the numeral *westiora* ‘one’ can function as an indefinite determiner (cf. Lorient et al., 1993, 43).

- (53) a. Ja oi beain-bira, ea kai. (Shipibo)
 this rain come-despite, I leave.INC
 ‘Even if the rain comes, I am leaving.’
- b. Moatian-ronki ipaoni-ke westiora joni kikin-bires koshi.
 long ago-HRSAY was-INC one man extremely-INTEN strong
 ‘In the old days (they say) there was a very strong man. (from Lorient et al., 1993, 43, my glosses)

Within the SR clause, a NS will remain unvalued, because inflection lacks the relevant specification required, as in Chinese (cf. (54)). Following the assumptions we made earlier, Shipibo has a paragraph-level strong phase (cf. Camacho, 2010c, and at that level (the main clause), an overt 1st and 2nd person must serve as an antecedent, just as in Chinese.



At the main-clause level, the NS requires an overt antecedent if interpreted as 1st or 2nd, as in (54). If no antecedent appears, then I assume the last resort default 3rd person interpretation applies, as in (55).



6.5 Summary

In this chapter, I have explored how the theory of NS identification extends discourse identification. First, I have shown that discourse identification is required even in languages with strong agreement, involving absolute clauses in Spanish, adjunct clauses in Hebrew and subjunctive in Spanish/Italian. I have then explored discourse identification in languages where it is the only available mechanism, including Germanic topic drop, Japanese, Chinese and Shipibo. One of the resulting proposals is a paragraph-level strong phase where a single discourse-level topic is identified. This projection delimits what I have called mid-range binding/control of NSs, a set of phenomena that include adjuncts in Spanish and Hebrew, as well as control in Chinese.

Chapter 7

Null/overt subject contrasts

The proposed analysis of NSLs in terms of valuation of *pro* by a morphologically specified agreement and by a discourse antecedent raises the issue of what happens in agreement-strong NSLs when subjects are overt. Consider the Quechua examples from 2.1.1, repeated below. In (1a), *pro* will be valued as 3.SG. by the probe, however in (1b), the overt subject is already valued, VALUATION would involve erasing the existing values of one of the elements.

- (1) a. Papa-ta mikhu-n-mi. (Quechua)
potato-ACC eat-3.SG-FOC/EVID
'S/he eats potatoes.'
- b. Huwan-mi papa-ta mikhu-n.
Huwan-FOC/EVID potato-ACC eat-3.SG
'Huwan eats potatoes (attested).'

Consider more closely the logical possibilities VALUATION entails within the model I have proposed. When two categories agree, either only one of them has values (as in (2a)), or both do. In this second case, the values can be identical (as in (2b)) or different (as in (2c)). If valuation involves sharing of a terminal node by the root, then the outcomes (presented at the righthand side of the arrows in (2)) would only yield a consistent result in (2a, b) but not in (2c), where each node must be interpreted as simultaneously having the values A and B. We can formulate VALUATION in these terms (terminal node sharing), with the proviso that the resulting structure must be interpretable and hence cannot contain contradictory information.

- (2) a. F F → F F b. F F → F F
A A A A A A
- c. F F → F F
A B A B

Thus, although agreement between an overt subject and inflection can be subsumed under the proposed agreement mechanism, the question remains as to whether an overt subject always has the representation in (3a), or whether the structure in (1b) is also possible. In this last structure, T values *pro*, and the overt subject is dislocated.

- (3) a. [DP T]
 [3SG] [3SG]
 ↑ ↑
 └──────────┘
- b. [DP_i [*pro*_i T]]
 [3SG] [3SG] ← [3SG]
 ↑ ↑
 └──────────┘

In this chapter, I review existing arguments in the literature for a principled distinction between NSs and overt pronominals, and in chapter 8 I turn to the nature and position of overt DPs.

7.1 *Pro* as a weak pronoun

Cardinaletti (1997) and many others assume that *pro* is a weak pronoun. In sect. 4.4.3 I already introduced the properties ascribed to weak pronouns, which I briefly review here: According to Cardinaletti and Starke (1999), weak pronouns must occupy a dedicated, non-basic position, although this property is not directly observable in the case of *pro*. They generally cannot be coordinated (cf. (4)), although they can in Irish (cf. (5)).

- (4) a. *Pedro/ella y salimos a la calle. (Spanish)
 Pedro/she and went-out to the street.
- b. Pedro/ella y yo salimos a la calle.
 Pedro/she and I went-out to the street.
 ‘Pedro and I went out on the street.’

- (5) da mbeinn -se agus tusa ann. (Irish)
 if be.COND.1.SG -CONTR and you there
 ‘If you and I were there.’ (from McCloskey and Hale, 1984, 501, ex. 31a)

Additionally, weak pronouns cannot acquire their reference through ostention (pointing) (cf. (6)) and they can refer to a [-human] referent. Thus, for example, (7a) is ambiguous between a [+human] and [-human] interpretation, whereas (7b) is not.

- (6) a. #*Pro* è veramente bello. (Italian)
 is truly beautiful
 ‘He (over there) is truly beautiful.’
- b. Lui è veramente bello.
 he is truly beautiful
 ‘He (over there) is truly beautiful.’ (from Roberts, 2010a, 71, ex. 17d)
- (7) a. *Pro* se cayó. (Spanish)
 CL fell
 ‘It/he/she fell.’

- b. Él/ella se cayó.
 he/she CL fell
 ‘S/he/*it fell.’

Weak pronouns cannot be modified, that is (8a) cannot mean something similar to (8b), although Irish is also an exception to this (cf. (5)). Finally, they can appear in impersonal clauses (cf. the contrast between the possible impersonal reading in (9a) vs. the referential reading in (9b)).

- (8) a. #Solo *pro* saben la respuesta. (Spanish)
 only know the answer
 b. Solo ellas saben la respuesta.
 only they.FEM know the answer
 ‘Only they know the answer.’
- (9) a. *Pro* me vendieron un aguacate dañado. (Spanish)
 CL sold.3.PL an avocado damaged
 ‘I was sold a damaged avocado.’
 b. Ellos me vendieron un aguacate dañado.
 they sold mCL sold.3.PL an avocado damaged
 ‘They sold me a damaged avocado.’

Cardinaletti and Starke (1999) derive the difference between strong/weak pronouns from distinct syntactic structures. In particular, strong elements have a CP, but weak ones do not. This category, in turn, has a referential index, which derives some of the referential asymmetries observed above.

7.2 On the difference between a null and overt pronouns

We can translate Cardinaletti and Starke’s assumption that strong pronouns have a referential index in the following way: By virtue of being unspecified, *pro* lacks any of the referential nodes that an overt category has by default (D, SPECIFIC, DEFINITE, DEICTIC). Even after VALUATION, *pro* will not have reference, inflection does not have interpretable features. Thus, the representations for an overt 3rd person pronominal, a null pronominal and a 3rd person inflection (for a language without gender distinctions) are given in (10), where the crucial difference between an overt pronoun and inflection is that inflection has uninterpretable D/*pi* (signaled by *u*).

- (10) Overt pronominals, null pronominals and inflection

Overt pron. (3 rd p. sg.)	<i>pro</i>	Infl. (3 rd p. sg.)
R	R	R
D		<i>u</i> D
... π INDIV		... <i>u</i> π INDIV
CLASS		CLASS
FEM MAS		FEM MAS

After VALUATION, *pro* will have an uninterpretable D/ π feature, so if nothing else happens, it will not receive a referential interpretation. Where does the value referential content for D/*pro* come from? One natural source is discourse linking, as discussed in chapter 6. Thus, the referentially-defective nature of null pronouns provides us with a unified theme for the two major strands of analysis of NSs. On the one hand, Rizzi (1986a) argues that *pro* needs to be identified by copying rich agreement. On the other hand, Frascarelli (2007) argues that *pro* is always identified by a topic, even in a language like Italian. Assuming Rizzi's perspective, one might wonder why topic identification is needed, given that morphology already values *pro*. Conversely, if topic identification is always obligatory, as Frascarelli proposes, what is the purpose of having verbal person/number/gender morphology? Throughout this book, we have seen examples that both types of identification are attested across languages and even within a single language. Within the perspective adopted in this chapter, morphological identification does value *pro*, but not sufficiently for it to have referential properties. Thus, it is possible to have an expletive *pro* without discourse identification, but in order to have a referential interpretation, discourse linking becomes crucial.¹

7.2.1 Binding effects

Recall from sect. 2.2.3 that for some speakers of NSLs, overt pronouns cannot be bound by a quantifier (cf. Montalbetti, 1984). This effect happens both in languages with rich morphological agreement and no morphological agreement, as illustrated in (11)-(13), for Spanish, Chinese and Japanese respectively. In both cases, the overt pronoun cannot be bound by a quantifier.

- (11) a. Todo estudiante_i cree que *pro*_i es inteligente.
 every student thinks that *pro* is intelligent
 'Every student_i thinks that he_i is intelligent'

¹In the proposed analysis, I do not assume a unified concept of weak pronouns, but rather derive the properties described by Cardinaletti and Starke (1999) from separate principles: Lack of overt content and possibly reduced structure. The semantic effects observed for weak pronouns stem from reduced structure (lack of D), but lack of conjoinability derives from prosodic conditions on coordination.

- b. Todo estudiante_i cree que él_{*i/j} es inteligente.
 every student thinks that he is intelligent
 ‘Every student_i thinks that he_j is intelligent.’
- (12) a. Meige ren_i xihuan Ø_i neng xingfu. (Chinese)
 every man wish can happy
 ‘Everybody wishes that s/he can be happy.’
 b. *Meige ren_i xihuan ta_i neng xingfu.
 every man wish he can happy
 ‘Everybody wishes that s/he can be happy.’
 (from Montalbetti, 1984, 187-188, exs. 13b, 14b)
- (13) a. Dare_i-ga [Ø_i sore-o mita to] itta no. (Japanese)
 who-NOM that-ACC saw that said Q
 ‘Who_i said that (he_i) saw that?’
 b. *Dare_i-ga [kare_i-ga sore-o mita to] itta no.
 who-NOM he-NOM that-ACC saw that said Q
 ‘Who_i said that (he_i) saw that?’ (from Kanno, 1997, 266, ex. 3)

For languages like Japanese, experimental data from Kanno (1997) suggest that native speakers unanimously disallow bound readings with overt pronouns like *kare*. Lozano (2002) found the same results for Spanish and Greek natives, but for some speakers of Spanish, OPC effects do not hold (cf. ?).

Under the assumptions I have been making, in both Chinese and Italian/Spanish, *pro* lacks a valued D, and in examples like (11a)-(13a), the quantifier serves as the antecedent that values D.² One can see the quantifier as valuing *pro*'s referential value through quantification, as in (14).

- (14) QP_i . . . [[R[D_i]] VP]

If, on the other hand, the overt pronoun has a valued referential D as part of its lexical entry, then the quantifier does not directly value the pronoun's [D] node, so a bound interpretation requires some independent mechanism. We can speculate that this alternative mechanism is less economical if valuation is available, hence a preference for the null pronoun counterpart.

This analysis does not entail that overt pronouns can never be bound. For example, certain overt pronouns can be interpreted as generic in Spanish. Specifically, a 2nd person can be used to denote a generic reading, as in (15a), even if the pronoun is overt (cf. (15b)).

- (15) a. Cuando visitas una bodega, normalmente te dejan probar los
 when visit.2.SG a winery, normally CL.2.SG let try the
 vinos. (Spanish)
 wines
 ‘When you visit a winery, they normally let you try the wines.’

²The quantifier must do something else, given that the interpretation in all three cases is quantificationally bound. I leave that aside.

- b. Cuando visitas una bodega, tú no sabes si te van a cobrar por
 when visit.2.SG a winery, you not know if CL.2.SG are-going to charge for
 probar.
 taste
 ‘When you visit a winery, you don’t know if they are going to charge you for tasting.’

It follows that even pronouns that have an overt D can treat it as non-referential (cf. Vergnaud and Zubizarreta, 1992). Perhaps this alternative semantic interpretation of D is also more generally available for speakers that do show OPC effects.

7.2.2 Backwards coreference in temporal adjunct clauses

The overt/null contrast surfaces again in cases of backwards anaphora in temporal clauses. Larson and Luján (1989, 1) point out that an overt pronoun cannot be coreferent with the main-clause subject in (16a), but its null counterpart can in (16b) in Spanish. Italian seems to pattern in the same way as Spanish (thanks to Vieri Samek-Lodovici, p.c. for judgements).³

- (16) a. Cuando *pro*_{i/j} trabaja, Juan_i no bebe. (Spanish)
 when work.3.SG Juan not drink.3.SG
 When he works, Juan doesn’t drink.’
 b. Cuando él_{*i/j} trabaja, Juan_i no bebe.
 when he work.3.SG Juan not drink.3.SG
 When he (someone else) works, Juan doesn’t drink.’
 (from Larson and Luján, 1989, 1, ex. 1)

In their analysis, this follows from the fact that the overt pronoun/emphatic pronoun is focused and occupies a higher position than the null one, as in (17), from which the focused pronoun can c-command the R-expression, yielding a Principle C violation of the Binding Theory.

- (17) a. [... *pro* ...]
 b. [él]_i [... t_i ...] (from Larson and Luján, 1989, 1, ex. 2)

As expected, when the emphatic pronoun does not c-command the R-expression, the contrast disappears, as seen in (18).

- (18) a. Cuando Juan_i trabaja, *pro*_{i/j} no bebe. (Spanish)
 when Juan work.3.SG not drink.3.SG
 When Juan works, he doesn’t drink.’

³Larson and Luján (1989, 4) note that the contrasts in (16) can be reproduced in English with unstressed and stressed pronouns, as in (i)-(ii).

- (i) *When HE_i works, John_i doesn’t drink. (English)
 (ii) When he_i works, John_i doesn’t drink. (from Larson and Luján, 1989, 4, ex. 14)

- b. Cuando Juan_i trabaja, él_{i/j} no bebe.
 when Juan work.3.SG he not drink.3.SG
 'When Juan works, he doesn't drink.' (from Larson and Luján, 1989, 3, ex 7)

Larson and Luján's pronominal typology includes two types of pronouns: emphatic and non-emphatic. However, Sánchez (1994, 489) suggests that Larson and Luján's contrasts disappear when the pronoun is stressed. For other speakers of Spanish, however, emphasis does not necessarily improve the coreference possibilities in (19).

- (19) Cuando ÉL_i trabaja, Juan_i no bebe. (Spanish)
 when HE works, Juan not drinks.
 'When HE works, Juan doesn't drink.' (from Sánchez, 1994, 489, ex. 41)

Italian also shows improved coreferential possibilities when the pronoun is stressed, as in (19). As Samek-Lodovici (p.c.) points out, extending the anti-c-command analysis for the pattern in (20), requires that the focused pronoun in (20) does not move, or that it moves to a position that is lower than the unstressed pronoun counterpart of (16b) above.

- (20) Quando guida LUI_i, Gianni_i non beve. (Italian)
 when drives HE Gianni not drinks.
 'When HE drives, Gianni doesn't drink.' (from Vieri Samek-Lodovici, p.c)

An alternative account more consistent with the analysis developed here would be the following: As assumed, the crucial difference between the null and overt pronoun is that the former requires a topic antecedent to provide interpretable content to the D feature, whereas the latter does not. Consider the topic-structure for the different examples.

- (21) a. Top_i [Cuando *pro*_i ...] [DP_i ...]
 └──────────┬──────────┘
 [*u*D]
 ↑
 b. Top_i [Cuando él_i ...] [DP_i ...]
 [*i*D]
 c. Top_i [Cuando ÉL_j ...] [DP_j ...]
 [*i*D]

In (21a), the topic provides an interpretable value to *pro*'s [D], and also acts as a topic for the following DP. On the other hand, if the topic is coindexed with the overt pronoun (as in (21b)), the resulting structure involves two independently valued but coreferential D features, but as suggested in the previous section, valuation of the null option is preferred (over independent values that result in coreference). On the other hand, (21b) with a disjoint reading is fine, because no alternative option based on valuation is available. Finally, in (21c), the stressed pronoun is interpreted as a contrastive topic, which makes it possible for it to be indexed differently from the preceding topic.

7.2.3 Backwards anaphora in conditional sentences

Frascarelli (2007) quotes patterns for Italian conditional clauses that are reminiscent of the ones just seen in the preceding section for Spanish. Like in the Spanish temporal adjunct clauses, the overt pronoun cannot serve as an antecedent to an overt DP. However, unlike in the Spanish temporal adjunct clauses, the overt pronominal can never be coreferential with a DP, regardless of which position each one occupies, as seen in (22a)-(24a). Conversely, a null subject can be coreferential in most instances (with one exception I will return to below), as seen in (22b)-(24b).

- (22) a. *Se lui_i si sentirà bene Jim_i andrà. (Italian)
 if he CL will feel well Jim will leave
 ‘If he feels well, Jim will go.’
 b. Se *pro*_i si sentirà bene Jim_i andrà.
 if CL will feel well Jim will leave
 ‘If he feels well, Jim will go.’ (from Frascarelli, 2007, 695, ex. 6b)
- (23) a. *Jim_i andrà se lui_i si sentirà bene. (Italian)
 Jim will go if he CL will feel well
 ‘Jim will go if he feels well.’
 b. Jim_i andrà se *pro*_i si sentirà bene.
 Jim will go if CL will feel well
 ‘Jim will go if he feels well.’ (from Frascarelli, 2007, 695, ex. 6a)
- (24) a. *Se Jim_i si sentirà bene lui_i andrà. (Italian)
 if Jim will feel well he will go
 ‘He will go if Jim feels well.’
 b. Se Jim_i si sentirà bene *pro*_i andrà.
 if Jim CL will feel well will go
 ‘If Jim feels well, he will go.’ (from Frascarelli, 2007, 695, ex. 6d)

These patterns are identical in Spanish, so conditionals seem to have a different distribution with respect to the null/overt pronoun contrast in both languages than temporal adjunct clauses. This raises a challenge to extending the Larson & Luján account to the conditional cases, because that analysis is crucially based on whether the overt pronoun c-commands the DP, and c-command obtains in certain contexts but not others (cf. (16b) vs. (18b) above). If that analysis were to apply to the conditional cases, one would have to show that the overt pronoun c-commands the DP in all of the relevant examples, regardless of whether the pronoun appears in the *if*-clause or the consequence clause, as in (25).

- (25) a. *He_i [_{CP1} si/se t_i . . .] [_{CP2} DP_i . . .]
 b. *He_i [_{CP1} t_i . . .] [_{CP2} si/se DP_i . . .]

However, proposing a unified account to all of the conditional cases runs into problems because of evidence that extraction from both clauses is not equally possible in conditional constructions. In Spanish, for example, *wh*-extraction is possible from the consequence-clause, but not from the *if*-clause (cf. (26)).

- (26) a. Si Ana organiza la fiesta, Blanca va a traer la comida. (Spanish)
 if Ana organizes the party Blanca goes to bring the food
 ‘If Ana organizes the party, Blanca is going to bring the food.’
- b. ¿Qué va a traer Blanca t [si Ana organiza la fiesta]?
 what go to bring Blanca if Ana organizes the party
 ‘What is Blanca going to bring if Ana organizes the party?’
- c. *¿Qué si Ana trae t] Blanca va a organizar la fiesta?
 what if Ana brings Blanca goes to organize the party

If pronoun raising and *wh*-extraction are subject to the same locality constraints, then one would expect coreference to be impossible when the pronoun is in the consequence clause but not in the *if*-clause, because *wh*-movement is not possible from the *if*-clause (cf. (26c)).

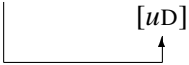
Note, furthermore, that the two clauses are not completely symmetric. Rather, it seems that the *if*-clause occupies the same position as a clitic left-dislocated phrase, that is, to the left of a *wh*-word. Consider, for example, the paradigms in (27). In the most natural order, the *wh*-word appears adjacent to its own condition clause (cf. (27a-b)), but as (27c) shows, a *wh*-word to the left of the *if*-clause is fairly marginal, and only improves with a rhetorical question flavor. This is similar to the distribution of clitic left-dislocated phrases, as seen in (28).

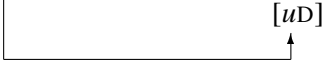
- (27) a. [Si Ana trae el pastel] ¿quién va a organizar la fiesta? (Spanish)
 if Ana brings the cake who is going to organize the party
 ‘If Ana brings the cake, who is going to organize the party?’
- b. ¿Quién va a organizar la fiesta [si Ana trae el pastel]?
 who goes to organize the party if Ana brings the cake
 ‘Who is going to organize the party if Ana brings the cake?’
- c. ??¿Quién [si Ana trae el pastel] va a organizar la fiesta?
 who if Ana brings the cake is going to organize the party
 ‘Who, if Ana brings the cake, is going to organize the party?’
- (28) a. A la hermana de Ana, ¿quién la va a invitar a la fiesta? (Spanish)
 to the sister of Ana, who CL is going to invite to the party
 ‘Ana’s sister, who is going to invite her to the party?’
- b. *?¿Quién a la hermana de Ana la va a invitar a la fiesta?
 who to the sister of Ana CL is going to invite to the party
 ‘Who, Ana’s sister, is going to invite (her) to the party?’

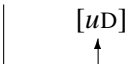
This parallelism between CLLD phrases and conditionals suggests a structure like (29) for conditionals, assuming an extended version of CP. In this structure, the *if*-clause is in a specifier of a left-peripheral position of the consequent clause. This derives why extraction from the *if*-clause is not possible (cf. (26c)).

(29) [XP₁ [CP₂ si ...] [X₁ [CP₁ quién [IP₁ va a ...]]]]

Extending the analysis proposed for adjunct temporal clauses, let us assume that in all configurations a topic is present, yielding the following structures. In (30a)-(32a), the topic can identify *pro*, which in turn can be coindexed with the DP. In (30b)-(32b), on the other hand, since the pronoun is overt, it should be used to change a topic, but this would lead to a clash in indices between the topic and the subsequent DP, hence coindexing would be impossible.^{4,5}

(30) a. Top_i [Si/se *pro*_i ...] [DP_i ...]

 b. *Top_i [Si/se él/lui_i ...] [DP_i ...]
 [iD]

(31) a. Top_i [Si/se DP_i ...] [*pro*_i ...]

 b. *Top_i [Si/se DP_i ...] [él/lui_i ...]
 [iD]

(32) a. Top_i [*pro*_i ...] [si/se DP_i ...]

 b. *Top_i [él/lui_i ...] [si/se DP_i ...]
 [iD]

To summarize, the evidence from backwards pronominalization better supports the analysis based on topic binding of *pro* than the alternative based on the anti c-command condition.

⁴In (32), I am abstracting away from the sentence-internal structure of the sequence. Since the *if*-clause follows the consequence clause, there must be some additional movement of that clause to a left-peripheral position higher than the *if*-clause.

⁵Example (24b) above minimally contrasts with (i) below, which is ungrammatical, for reasons that are not clear to me.

(i.) **Pro*_i andrà se Jim_i si sentirà bene.
 will go if Jim CL will feel well
 ‘If Jim feels well, he will go.’

(Italian)

(from Frascarelli, 2007, 695, ex. 6c)

7.3 Deriving expletives

Although in section 4.5 I suggested that NSLs may lack expletives altogether, in this section, I will show how an expletive could arise in the proposed account. One of the most notable properties of expletives is that the verb appears in 3rd person singular, and it would be an added value if the theory could derive this restriction. It seems reasonable to relate this property to the other property of expletive constructions, namely the fact that they do not assign a regular theta role.

Svenonius (2002, 6) notes that several researchers have argued that certain types of expletives have some kind of semantic content, as already noted in sect. 2.1.1.1. In support for this contention, Chomsky (1981, 323-5) points out to examples of alleged control by weather-verb expletives (cf. (33)). Additionally, Hoekstra (1983) argues that extraposition *it* is referential, because it is always coindexed with a CP (cf. (34)). However, it is not clear that coindexing here means the same as coreferential, just as agreement involves some notion of matching that does not imply coreference.

- (33) a. It often clears up here right after snowing heavily. (English)
(from Svenonius, 2002, 6, 5a)
- b. Korai szürkületet okozva befelhösödött. (Hungarian)
early twilight.ACC causing clouded
'It clouded, causing early twilight.'
(from Kiss, 2002, 119, ex. 39a)
- (34) It is obvious where you got that hickey. (English)
(from Svenonius, 2002, 5, ex. 4a)

A slightly different argument for the semantic content of expletives comes from the fact that in Hungarian, extraposition pronouns can be focused (cf. (35)), as noted by Kiss (2002).⁶

- (35) Számomra csak az volt nyilvánvaló hogy Éva megbukik. (Hungarian)
for.me only that.NOM was obvious that Eve fails
'Only that was obvious for me that Eve would fail.'
(from Kiss, 2002, 112, ex. 17b)

Assuming for the sake of argument that some expletives are not semantically vacuous, it is clear that their content is not fully referential, and that if they receive a theta role, it is not one of the usual ones. Suppose that *pro* appears with a 3rd person inflection, gets valued but has an uninterpreted D feature. By assumption, this feature must be recovered in discourse. If a topic were to value it, then the resulting structure would have a topic with no semantic connection to the clause, because no argument in the clause bears a theta role. Thus, the only possible interpretation would entail

⁶Note that the kind of expletive illustrated in (35) is only possible in a NSL like Spanish with an overt demonstrative, as shown in (i). This is expected if elements modified by *solo* must be focused and focus is typically associated with stress in these languages. But this also means that strictly speaking, one cannot test whether a null expletive can be focused.

- (i.) Solo ESO es obvio que Eva va a fracasar. (Spanish)
only that is obvious that Eva will fail
'Only that was obvious that Eve would fail.'

a non-canonical theta role (if Chomsky is right). If there is no topic, the result is the same. The same logic applies to other persons, with the caveat that we assume the default interpretation to be associated with the 3rd person, not with the others.⁷

7.3.1 Full and default agreement with associate DPs

The analysis in the preceding section suggests that the expletive agrees with the verb, but the truly default interpretation resulting from lack of topic valuation is 3rd person. These expletives contrast with other types of expletives that can appear in other persons in the context of expletive-associate constructions, discussed by Manzini and Savoia (2002). They state two important points: First, whether the postverbal subject agrees with the verb or not is entirely independent of whether the language has null subjects. Indeed, English tends to have agreement with the postverbal associate, but French does not (Manzini and Savoia (2002, 170)). Second, when associates can be definite, 1st and 2nd persons can appear as associates, and the expletive can agree in all three persons, as shown in (36)-(38). However, when the postverbal associate is 1st-2nd person, full agreement is obligatory, as seen in the contrast in (38). Thus, agreement mismatches are only possible for number and gender, but person agreement is obligatory (cf. Manzini and Savoia, 2002, 175-176).

- (36) a. vengg i (Urbino)
 come I
 ‘I come.’
 b. si nut te
 have come you
 You have come
 c. sen nuti no
 have come we
 ‘We have come. (from Manzini and Savoia, 2002, 176, ex. 28)
- (37) t vi t3 (Alfonsine Italian)
 CL.SUB come.2.SG you
 ‘You came.’ (from Manzini and Savoia, 2002, 175, ex. 27b)
- (38) a. i faɲtʃ i veɲan dɔpo (Casaccia Italian)
 the children CL-SUB come.3.PL afterward
 ‘The children come afterward

⁷The default 1st person assignment we discussed earlier will be unavailable precisely because there is no theta-role assignment: It makes little sense to assign a speaker role to a situation in which no theta roles are defined. One could assign such an interpretation to an evidential, but in such case the constant function does not apply to the subject of the clause, but to a separate layer that contains evidential heads, as illustrated by the following Quechua example provided by Liliana Sánchez (p.c.)

- (i.) Paramu-chkan-mi. (Quechua)
 rain-PROG-EVID
 ‘(I have direct evidence that) It is raining.’

- b. *dopo al vej i faptj*
 afterward CL.SUB come.3.SG the children
 ‘Afterward come the children.’ (from Manzini and Savoia, 2002, 171, ex. 18)

Manzini and Savoia (2002) design a system in which agreement is parasitic on theta-assignment. Thus, when the clitic, the associate and Infl fully agree, this reflects movement of a theta feature to all three positions. Variation in number is accounted for by postulating a second theta-role that is shared only between V and N. In addition, they propose four distinct functional projections associated with subject clitics, [D [N [Num [P]]]]. These positions are only projected when the clitic system reflects their actual content.

To illustrate the system, (39a) shows 3rd person plural agreement on the clitic, the verb and the associate subject. The clitic *i* appears in NumP, and an additional clitic *a* heads DP. In (39b), on the other hand, D is not projected, and *al* is heads NP. In both cases, agreement is triggered by the theta-feature chain. In (40), on the other hand, agreement is not shared across the board, because a second theta-feature, which is not attracted as far as the first theta-feature.⁸

- (39) a. [DP θ_1 -*a* [NumP θ_1 -*i* . . . [IP θ_1 *duar* [VP *i kanais*]]]] (Vito d’Asio Italian)
 CL.SUB 3.PL.MAS sleep the children
 b. [NP θ_1 -*al* . . . [IP θ_1 -*duar* [VP θ_1 -*il fanat*]]]
 CL.SUB sleep the boy (cf. Manzini and Savoia, 2002, 180, 33a, c)
- (40) [DP θ_1 -*al* . . . [IP θ_1 *ϕyge* [VP θ_1 - θ_2 -*i fiɔi*]]] (Premana Italian)
 CL.SUB plays the children

I think the agreement patterns of Spanish existentials can shed some light on the analysis of Italian dialects. In Spanish, existential constructions show associate agreement in the past/future for some speakers, as illustrated in (41). For others, agreement is necessarily 3rd person, as seen in (42) (cf. Rodríguez-Mondoñedo, 2006 and references therein).

- (41) *Habían varios invitados.* (Latin American Spanish)
 were several guests
 ‘There were several guests.’
- (42) *Había/*habían varios invitados.* (Peninsular Spanish)
 was/*were several guests
 ‘There was several guests.’

For all speakers, lack of agreement with the associate is the only possibility in the present. This correlates with the fact that the verb has an unusual verb ending for a 3rd person (-y), as seen (43).⁹

⁸In Manzini and Savoia’s analysis, θ_1 is more precisely an aspectual role, and θ_2 establishes the relationship between D and N, by which N is D’s lexical restrictor.

⁹*Hay* is only used as the present form for existential constructions.

Chapter 8

The status of preverbal subjects in NSLs

As mentioned in 4.3, treating verbal morphology as pronominal means that overt preverbal subjects (PS) are not the ‘real’ subject of a clause in a NSL, but perhaps a dislocated or A’ constituent. Even if one does not accept the pronominal agreement hypothesis, the distribution of PSs seems to be substantially different from that of overt subjects in languages like English. For these reasons, much of the debate over the past thirty years has centered on the status of PS (cf. Rizzi, 1982; Rigau, 1988; Contreras, 1991; Koopman and Sportiche, 1991; Solà, 1992; Zubizarreta, 1994; Cardinaletti, 1997, 2004; Alexiadou and Anagnostopoulou, 1998; Ordóñez and Treviño, 1999; Goodall, 1999; Suñer, 2003; Camacho, 2006; Frascarelli, 2007; Sheehan, 2007, among others). This debate usually conflates logically independent notions: First, whether PS occupy an A or A’-position, second, whether PS have the properties of clitic left-dislocated (CLLD) phrases, and third, whether they are interpreted as topics or not.

These theoretical constructs overlap in many ways, but they are not completely identical. For example, some researchers argue that PSs are CLLD phrases because CLLD phrases cannot be interpreted in the same position as their non-dislocated counterparts (i.e. they do not reconstruct, cf. Cinque, 1990, 66). Thus, according to Cinque, sentences like (1a) and (2a) differ with respect to the possibility of interpreting the first constituent in its base position: The former can be interpreted as its counterpart in (1b), without displacement, whereas the latter cannot, hence it is ungrammatical (with that interpretation). These interpretive differences correlate with the fact that (2a) is a CLLD construction (with a null clitic, as in (2b)), but (1a) is not.¹

- (1) a. PER QUESTA RAGIONE_i, ha detto che se ne andrà t_i . (Italian)
for this reason has said that CL CL will leave
‘FOR THIS REASON he said that he would leave.’
- b. Ha detto che se ne andrà per questa ragione.
has said that CL CL will leave for this reason
‘S/he said that s/he will leave for this reason.’ (from Cinque, 1990, p.64, ex. 13)
- (2) a. *Per questa ragione_i, ha detto che se ne andrà t_i . (Italian)
for this reason has said that CL CL will leave
‘For this reason, he said that he would leave.’ (from Cinque, 1990, 66, ex. 14)

¹In Cinque’s analysis, CLLD does not involve movement, but focalization or wh-questions do.

- b. [PP Per questa ragione]_i [ha detto [CP che [CL_{0-i} se ne andrà]]]

Based on this assumption that CLLD phrases do not reconstruct, Alexiadou and Anagnostopoulou (1998), for example, argue that PSs are also left-dislocated, since they do not reconstruct either (see below). However, it is logically possible that PSs and CLLD phrases do not reconstruct for independent reasons. For example, PSs could fail to reconstruct because they are A'-moved, whereas CLLD phrases do not reconstruct because they are base-generated in their peripheral positions. On the other hand, one could argue that both are base-generated in peripheral positions, but the nature of those positions is different.

CLLD phrases are usually interpreted as topics, and PS are as well, although not always, but once again, these are two potentially independent properties: Topichood is a discourse-related property that does not correlate one-to-one with any given syntactic position. As is well-known, so-called hanging topics (cf. (3)) are topics but have distinct syntactic properties from CLLD phrases or topicalization. For example, there is no syntactic connection between the hanging topic in (3a) and the embedded verb *seguir* 'continue', whereas there is a clear relationship between the dislocated phrase in (3b) and the clitic *la* 'it'. In fact, as we saw in section 6.2, there may be several topic positions within the clause.

- (3) a. En cuanto a la decisión, tenemos que seguir la discusión. (Spanish)
with respect to the decision, have to continue the discussion
'With respect to the decision, we have to continue the discussion.'
- b. La decisión, tenemos que continuar discutiendo-la.
the decision, have to continue discussing-CL.ACC.FEM
'The decision, we have to continue discussing it.'

The strongest version of the PS-as-CLLD proposals assumes that **all** preverbal subjects should be in an A' position and that PSs have exactly the same properties instantiated in CLLD phrases, i.e. as topics (among other things). A weakened version would argue that some subjects may be in A', while others are in A-position (Goodall, 1999; Suñer, 2003; Camacho, 2006; Sheehan, 2007). I will briefly discuss the main points of the debate, following Sheehan's (2007, 50-89) and Camacho's (2006) discussion.

8.1 Preverbal subjects as CLLD phrases

8.1.1 Scope asymmetries

Alexiadou and Anagnostopoulou, 1998, 504 argue that PS have the same scope properties as CLLD constituents, and these are reflected in two specific areas: Their scope interactions with other quantifiers, and with sentential modals and negation. According to them, quantified and indefinite PS in Greek and Spanish only take wide scope with respect to other quantifiers in the clause:

- (4) a. Kapios fititis stihiothetise kathe arthro. (Greek)
some student filed every article
'Some student filed every article.' (not 'For every article there was a student who filed it.')

- b. Algún/un estudiante archivó todos los artículos. (Spanish)
 some/a student filed every the articles
 'Some student filed every article.' (not 'for every article there was a student who filed it.'

This follows if they do not reconstruct, like CLLD phrases (cf. the discussion of (1)-(2) above). However, quantifier scope facts are more complex than examples (4) would suggest. On the one hand, Suñer (2003, 344) points out that PS can have narrow scope, and in fact she argues that this is the preferred interpretation for that example. As Sheehan (2007, 57-58) notes, in addition to Spanish, PS can have narrow scope in European Portuguese, Romanian and Italian.

- (5) a. En la biblioteca departamental, algún estudiante sacó prestado cada
 in the library departmental, some student took loaned each
 libro. (Spanish)
 book
 b. 'In the departmental library, each book was taken out by some student.' (narrow scope subject)
 c. 'In the departmental library, some (specific) student took out each book.' (wide scope subject, dispreferred)

A slightly different version of the scope argument is presented by Barbosa (1995, 36-37), as noted by Sheehan (2007, 58-61). Barbosa argues that indefinite subjects always take wide scope over other scope-bearing elements, unlike in English or French. Thus, Barbosa points out, in a context where I am being reminded of the items that are necessary in order to apply to a certain job, English (6a) and French (6b) are both grammatical, but (7a)-(9a) are not. By contrast, (7b)-(10b) are.²

- (6) a. A letter of recommendation is required. (English)
 b. Une lettre de recommandation est requise. (French)
 a letter of recommendation is required
 'A letter of recommendation is required.' (from Barbosa, 1995, 35, ex. 61-2)
- (7) a. ???Una carta de recomanacio es necessaria. (Catalan)
 a letter of recommendation is necessary
 'A letter of recommendation is required.'
 b. Es necessaria una carta de recomanacio.
 is necessary a letter of recommendation
 'A letter of recommendation is necessary.' (from Barbosa, 1995, 36, ex. 63)
- (8) a. ???Una carta de recomendación es necesaria. (Spanish)
 a letter of recommendation is necessary
 'A letter of recommendation is required.'

²The a) versions of (7)-(10) are appropriate if the letter is interpreted as a specific indefinite ('one of the letters').

- b. Es necesaria una carta de recomendación.
 is necessary a letter of recommendation
 ‘A letter of recommendation is necessary.’ (from Barbosa, 1995, 36, ex. 64)
- (9) a. ???Una lettera di raccomandazione é necessaria. (Italian)
 a letter of recommendation is necessary
 ‘A letter of recommendation is required.’
- b. E necessaria una lettera di raccomandazione.
 is necessary a letter of recommendation
 ‘A letter of recommendation is necessary.’ (from Barbosa, 1995, 36, ex. 65)
- (10) a. ???Uma carta de recomendação é necessaria. (EP)
 a letter of recommendation is necessary
 ‘A letter of recommendation is required.’
- b. É necessaria uma carta de recomendação.
 is necessary a letter of recommendation
 ‘A letter of recommendation is necessary.’ (from Barbosa, 1995, 36, ex. 66)

According to Barbosa, the deviant examples are inadequate because the context is set up for a wide-scope reading of the modal, but given the wide-scope requirements of the indefinite, this is not possible. If both English subjects and NSL PS are derived by A-movement, Barbosa points out, then one needs to suggest some extra machinery to account for the contrast between English/French and the other languages, whereas if PS subjects are left-dislocated, all that remains to be added is an independent theory of why dislocated indefinites take wide scope.

I think both Alexiadou and Anagnostopoulou’s and Barbosa’s contrasts can be subsumed under the criterial freezing account adopted earlier. The relevant generalization from these examples is that indefinite subjects in NSLs take widest scope possible. This suggests that in these languages, the subject is not frozen in Spec, TP, as the Subject Criterion analysis correctly predicts. In English and French, on the other hand, subjects are frozen in Spec, TP so their scope possibilities are limited to that position. If the modal (or another quantifier) can raise to a higher scope position, then the attested scope relations obtain. The difference between Alexiadou and Anagnostopoulou’s examples (where the subject wide scope over the object) and Suñer’s examples (where the object can have wider scope) has to do in part with the fact that distributive quantifiers like *cada* ‘each’ favors widest scope.

From another perspective, Zubizarreta (1994, 1998) shows that CLLD constituents and PS do not pattern uniformly with respect to reconstruction, and CLLD phrases themselves vary depending on whether they are accusative or dative. First, a quantified DP subject can bind a pronoun inside an accusative CLLD phrase if the subject is preverbal, not if it is postverbal (cf. (11a vs. b) and also Suñer, 2003). According to Zubizarreta (1998, 114-15), this is evidence for an intermediate projection for the clitic, above VP but below TP, to which the CLLD phrase can reconstruct, as in (12). Since postverbal subjects are inside VP, the reconstructed phrase will always be higher.

- (15) a. *¿Qué dijiste que quién compró el otro día? (Spanish)
 what said that who bought the other day
 ‘What did you say that who (S) bought the other day?’
 b. ¿Qué dijiste que compró quién el otro día?
 what said that bought who the other day
 ‘What did you say who bought the other day?’ (from O & T, 51, ex. 50)

O & T argue that this same pattern obtains with respect to CLLD wh-objects and wh-indirect objects as well. Thus, in (16), the subject can be extracted if the wh-indirect object is postverbal but not if it is preverbal (cf. (16a) vs b)). Since non-wh CLLD indirect objects do not induce the same effect, as seen in (16c), they suggest that the explanation cannot be due to whether the Spec, IP position is property governed or not.

- (16) a. ¿Quién crees que le va a dar eso a quién? (Spanish)
 who think that CL.3.SG will give that to whom
 ‘Who do you think will give that to whom?’
 b. *¿Quién crees que a quién le va a dar eso?
 who think that to whom CL.3.SG will give that
 ‘Who do you think that to whom will give that?’
 c. ¿Quién crees que a ti te va a dar eso?
 who think that to you CL.2.SG will give that
 ‘Who do you think that will give that to you?’ (from O & T, 50, ex. 52)

Rather, they connect this difference with the scope possibilities found in examples like (17). In (17a), the quantified DP *cada senador* ‘each senator’ only has narrow scope with respect to the wh-phrase, i.e. it is not possible to interpret it as each senator loves a different person, but in (17b), with a postverbal subject, the scope ambiguity reappears.

- (17) a. ¿A quién dices que cada senador amaba? (Spanish)
 to who say that each senator loved
 ‘Who do you say that each senator loved?’
 b. ¿A quién dices que amaba cada senador?
 to who say that loved each senator
 ‘Who do you say that each senator loved?’ (from O & T, 52, exs. 53-4)

O & T propose that freezing of the quantified subject in (17) would follow if it already occupies a topicalized, A'-position, as it would if it were CLLD. However, this logic leads to analyzing (16c) as involving non-cyclic movement of the subject from a postverbal position, otherwise we would predict that it should be ungrammatical if the wh-subject is moved from the pre-verbal A' position. This contrast between ((16a) and c) suggests that the effect is not due to being frozen in an A' position. Rather, a general property of Spanish is that it does not allow wh-words to appear in embedded position with verbs that do not select for wh-questions. Thus, (18) shows the same contrast as (15), but with a single wh-word.

- (18) a. *¿Pedro dijo que quién compró una casa el otro día? (Spanish)
 Pedro said that who bought a house the other day
- b. ¿Pedro dijo que compró una casa el otro día quién?
 Pedro said that bought a house the other day who
 ‘Who is it that Pedro said that bought a house the other day?’

8.1.3 Adverb intervention effects

Alexiadou and Anagnostopoulou (1998, 502) point out that in Spanish and Greek an adverb can follow a PS (cf. (19a)), whereas in French, they cannot (cf. (19b)). Assuming that adverbs adjoin to a maximal projection, the subject and the verb cannot be in the same maximal projection in Spanish. Assuming, further, that the verb is in I, it follows that the subject must be higher. In French, on the other hand, the subject could be in Spec, IP.

- (19) a. Juan ya se fue. (Spanish)
 Juan already CL left
 ‘Juan has already left.’ Sheehan (2007, 51, ex. 48)
- b. Jean *(déjà) veut (déjà) s’en aller. (French)
 Jean already wants already CL-of go
 ‘Jean already wants to leave.’ (from Sheehan, 2007, 51, ex. 50)
- c. Maria *(già) parlava già di lui. (Italian)
 Maria already was-speaking already of him
 ‘Maria was already speaking about him.’ (from Sheehan, 2007, 52, ex. 51)

However, as Sheehan (2007, 52) points out, the corresponding adverb in Italian patterns with the French one, suggesting that the test does not give similar results for Italian and Spanish, although they both are NSLs. Rather than follow from the NSL status, the contrasts in (19) would follow if the verb in Italian and French raises to a higher position than in Spanish.

The strong PS-as-CLLD hypothesis cannot readily explain why the adverb *probabilmente* ‘probably’ in Italian cannot intervene between the negative quantifier *nessuno* ‘nobody’ and the verb, as seen in (20a). If PS uniformly occupy the same A’-position, this is somewhat unexpected. Notice that if *nessuno* is contrastively focalized, then the sequence becomes acceptable (cf. (20b)). Recall that contrastive focus is analyzed by Cinque (1990, 66) as movement, but this would suggest that the ungrammaticality of (20a) follows from a requirement that the quantifier be in the Spec of IP, rather than in the CLLD position (cf. Belletti, 1990, quoted in Sheehan, 2007, 52).

- (20) a. *Nessuno probabilmente telefonerà alle 5. (Italian)
 nobody probably will-phone at-the 5
 ‘Nobody will probably phone at 5 o’clock.’
- b. NESSUNO probabilmente telefonerà alle 5.
 nobody probably will-phone at-the 5
 ‘NOBODY will probably phone at 5 o’clock.’

A similar state of affairs holds in Spanish with negative quantifiers, the adverb *casi* ‘almost’ and a modal (cf. Camacho, 2006). As (21) shows, a negative quantifier in PS position cannot precede a modal when the adverb *casi* ‘almost’ intervenes, contrasting with a non-quantified DPs (cf. (21b) vs. (22a)) and with one in postverbal position (cf. (21b) vs. (22b)). Finally, the PS negative quantifier also contrasts with a CLLD negative quantifier (cf. (21b)) vs. (22c).

(21) a. Nadie pudo avanzar 3 metros. (Spanish)
no one could advance 3 meters
‘No one could advance 3 meters.’

b. *Nadie casi pudo avanzar 3 metros.
no one almost could advance 3 meters

c. *Ninguno de ellos casi pudo avanzar 3 metros.
none of them almost could advance 3 meters

(22) a. La tortuga casi pudo avanzar 3 metros. (Spanish)
the turtle almost could advance 3 meters
‘The turtle could almost advance 3 meters.’

b. Casi no pudo avanzar 3 metros nadie.
almost not could advance 3 meters no-one
‘No one could almost advance 3 meters.’

c. A ninguno de los culpables casi lo descubrieron.
to none of the guilty almost CL discovered
‘None of the guilty, were almost discovered.’

These contrasts suggest that a unified account of PS as A’ CLLD constituents is not fine-grained enough to account for the distributional complexities.

8.1.4 PS incompatibility with quantified CLLD phrases

O & T (p. 45) note that although quantified DPs can appear as CLLD phrases (cf. (23a)), an overt PS intervening between the quantified phrase and the verb makes the sequence ungrammatical (cf. (23b-c)). As the example in (24) shows, this same pattern exists when the PS is quantified and the CLLD phrase is not.

These patterns would be unexpected if *pro* and the overt PS occupy the same position in each of those examples. According to O & T, if PSs and CLLD phrases occupy a similar position, the contrast follows: In (23b-c), the non-quantified phrase blocks the relationship between the quantifier and inflection. Furthermore, this contrast can be taken as evidence in favor of treating inflection as pronominal, if one assumes that that bound morphemes do not block relationships between heads and maximal projections, the contrast between (23a) and b) follows.

(23) a. Nada le debe a sus amigos. (Spanish)

nothing CL.3SG owes.3.SG to his/her friends
‘He/she owes nothing to his/her friends.’

(from O & T, 44, ex. 25)

- b. *Nada Juan le debe a sus amigos.
 nothing Juan CL.3.SG owes.3.SG to his/her friends
 ‘Juan owes nothing to his/her friends.’ (from O & T, 45, ex. 29)
- c. *A nadie Juan le debe la renta.
 to nobody Juan CL.3.SG owes.3.SG the rent
 ‘Juan owes the rent to nobody.’ (from O & T, 45, ex. 30)
- (24) a. *Nadie el libro se lo regaló a Marta. (Spanish)
 no one the book CL.3.DAT CL.3.ACC gave to Marta
 ‘The book, nobody gave it to Marta.’
- b. El libro, nadie se lo regaló a Marta.
 the book, no one CL.3.DAT CL.3.ACC gave to Marta
 ‘The book, nobody gave it to Marta.’

Notice that these facts are very similar to (21), where the adverb *casi* cannot intervene between the negative quantifier and a modal. These two sets of facts suggest that the negative quantifier must be adjacent to an inflectional projection, and that nothing can block that relationship, as O & T suggest. Since the effect is limited to negative quantifiers, and we know independently that negation must be adjacent to Infl (cf. (25)), we can assume with Bosque, 1980; Laka, 1990 that the NPIs must be adjacent to NegP.³

- (25) a. Alfonso casi no sobrevivió al accidente. (Spanish)
 Alfonso almost not survived the accident
 ‘Alfonso almost didn’t survive the accident.’
- b. *Alfonso no casi sobrevivió al accidente.
 Alfonso not almost survived the accident

In any case, given the fact that the content of the quantifier drives the restrictions above, it seems that one cannot generalize from these data that PS in general are always dislocated, but rather that negative quantifiers in preverbal position must be adjacent to an inflectional projection.

8.1.5 The distribution of Bare NPs

One final argument in favor of the PS-as-CLLD analysis relates to the impossibility of having bare NPs as PS or as CLLD phrases, as seen in (26a)-(27). Both bare PS NPs and bare CLLD phrases are ungrammatical. Once again, this would be an expected consequence if PS are CLLD phrases.

- (26) a. *Niños no llegaron esta mañana. (Spanish)
 children not arrived this morning
 ‘Children didn’t arrive this morning.’

³It is possible that the effects described by O & T extend to other categories associated with Laka’s (1990) ΣP.

- b. No llegaron niños esta mañana.
not arrived children this morning
'Children didn't arrive this morning.'
- (27) a. *Niños, no los vi esta mañana. (Spanish)
children not CL.3.PL saw this morning
- b. *A niños, no los vi esta mañana.
to children not CL saw this morning

However, the grammaticality of (28a) complicates the picture, since that example can be analyzed as having a null clitic, as in (28b). In fact, in languages like Catalan, there is an overt partitive clitic in these constructions (cf. (29)). If this is correct, then the contrast between (26a) and (28a) would follow from the distinct specifications of the agreement clitic, on the one hand, and the null clitic on the other.

- (28) a. Niños, no vi esta mañana. (Spanish)
children not saw this morning
'Children I didn't see this morning.'
- b. [NP_i [IP \emptyset_{CL-i} V . . .]]
- (29) a. Nens, no n' he vist al parc. (Catalan)
children, not CL have seen at-the park
'Children, I haven't seen (any) at the park.'
- b. Llibres, no en tinc.
books, not CL have
'Books, I don't have (any).'
- (from Francisco Ordóñez, p.c.)

If bare NPs CLLD phrases are not uniformly impossible but PS are, this weakens the argument for the strong version of the PS-as-CLLD hypothesis.⁴

8.1.6 Section summary

This section has reviewed six types of arguments in favor of considering PS as CLLD phrases. These arguments include their similar scope properties, extraction asymmetries, adverb intervention facts and behavior of bare NPs and quantified CLLD phrases. Like CLLD phrases, PS have been proposed not to reconstruct and to have wide scope, although evidence from Zubizarreta and Suñer suggest that this is not accurate for Spanish.

Regarding extraction asymmetries, both *wh*-subjects, *wh*-direct objects and *wh*-indirect objects block extraction of another *wh*-phrase if they are pre-verbal, and PS quantified subjects have

⁴It would be tempting to derive the behavior of bare NPs from the nature of the clitic: Taking a cue from the grammaticality of (28)-(29), one could argue that the ϕ -features of clitic and Infl in (26)-(27) respectively are incompatible with a bare NP. However, since bare NP subjects are grammatical postverbally, this would entail that Infl only agrees with the subject when preverbally, or alternatively, it agrees in a different way to be determined.

narrow scope with respect to a matrix wh-phrase. All of these facts are taken as evidence that pre-verbal arguments (PS, CLLD objects and indirect objects) occupy a similar, A'-position. However, as noted, scope facts are not consistent for all quantifiers, and wh-extraction facts may be due to a different reason, namely the generalized impossibility to have embedded wh-words with verbs that do not select for questions. With respect to adverb intervention effects, an adverb can intervene between the PS and the verb in Greek and Spanish but not in French. However, it was argued that this is not a consistent behavior of all adverbs and all PSs, suggesting that at least certain PSs are not left-dislocated.

Negative quantified phrases are ungrammatical when some constituent (a CLLD phrase, a PS or certain adverbs) intervenes between them and inflection. We argued that this distribution effectively groups CLLD phrases and PS, but only when they are negative quantifiers, hence these data do not allow us to generalize to all types of CLLD phrases and subjects.

Finally, both bare NPs cannot be used as CLLD phrases or PSs, but as noted, for the case of CLLD phrases, this may be due to the nature of the clitic, since dislocations that have a partitive clitic in Catalan can be bare in Spanish.

8.2 Against PS-as-CLLD

Several researchers have pointed out a number of areas in which PS do not pattern like CLLD phrases. In this section, we will review a few of the arguments that have been given against equating both types of elements.

8.2.1 Quantified PS and CLLD phrases

Rizzi (1986b, 395) points out the following contrasts: In (30), *nessuno* 'nobody' can appear as a PS but not as a CLLD element, and neither can *tutto* 'everything'. O & T (p. 52, fn. 16) argue that this contrast is a matter of incompatibility between the accusative clitic *lo* and the quantifier. They point out that dative clitics do not have this kind of restriction in Spanish, as seen in (31a). The fact that the equivalent of *tutto* 'everything' can be a CLLD phrase in Spanish (cf. (31b-c)) also suggests that each clitic's specification may be different, and this specification may constrain what can be dislocated.

- (30) a. Nessuno è venuto. (Italian)
 nobody has come
 'Nobody came.'
- b. *Nessuno, l' ho visto.
 nobody CL have seen
 'Nobody, I've seen (him).'
- c. *Tutto, lo dirò alla polizia.
 everything CL.3.SG will say to-the police
 'Everything, I will tell (it) to the police.'

- (31) a. A nadie le han avisado todavía el resultado. (Spanish)
 to nobody CL.3.SG have notified yet the result
 ‘They haven’t notified the result to anybody.’
- b. Todo lo he estudiado.
 everything CL.3.SG have studied
 ‘I’ve studied everything.’
- c. A todos los saludé ya.
 to everybody.PL CL.3.PL greeted already
 ‘I greeted everybody already.’

8.2.2 Discourse status of PS and CLLD phrases

Sheehan (2007, 74-77) points out that if PS are CLLD phrases, they should have the same informational status as other CLLD phrases. However, they pattern differently as possible answers to a question like *what happened?* In the answers to this type of question, it is usually assumed that the whole clause is non-presupposed, i.e. not focused (cf. Zubizarreta, 1998, 1-2). In European Portuguese, the answer to that question is obligatorily SV(O), according to Costa (2001), as seen in (32), and left-dislocation is not possible (cf. Rizzi, 2006, 122 for Italian).

- (32) a. O que é que aconteceu? (EP)
 the what is that happened
 ‘What happened?’
- b. O Pedro partiu o braço.
 the Pedro broke the arm
 ‘Pedro broke his arm.’
- b'. #Partiu o Pedro o braço.
 broke the Pedro the arm
- b''. #O braço, o Pedro partiu-o.
 the arm the Pedro broke-CL.3P.SG

As Pilar Barbosa (p.c.) points out, topicalized subjects in European Portuguese can appear in a wide focus context, as illustrated in (33). The question in (33a), uttered in a context in which mother is making a fuss about unexpected news, can be answered with (33b). Note that in this sentence, *seem* is singular, so *as crianças* ‘the children’ does not agree with it, but rather, it is a topic. The contrast between the possibility of subjects and topicalized subjects and CLLD phrases in wide focus contexts (cf. (32) vs. (33)) are somewhat surprising if subjects are CLLD phrases (and CLLD phrases all have the same interpretation).

- (33) a. O que é que aconteceu? (EP)
 the what is that happened
 ‘What happened?’

- b. As crianças parece que afinal vêm hoje.
 the children, seems.3.SG that after all come today
 ‘The children, it seems that they are coming today, after all.’

In Spanish and Catalan, the typical answer to wide-focus questions is also SV(O). Regarding dislocated phrases, both languages show a contrast between animate and inanimate DPs. Inanimate ones (which tend to be accusative) pattern like the EP example in (32b''), but animate dative dislocations do allow for a wide focus reading, as seen in (34). So, for example, suppose that I hear a sudden scream in the next room and I ask (34a). In that context, the answer in (34b) with a dislocated animate (dative or accusative) is perfectly acceptable, as seen in (34), just as the non-dislocated counterparts are. However, in this same context, (34d), with an accusative left-dislocated phrase is not possible.

- (34) a. ¿Qué pasó? (Spanish)
 what happened
 ‘What happened?’
- b. A Marta le dieron la noticia de su premio.
 to Marta CL.3.DAT gave the news of her prize
 ‘They gave Marta the news about her prize.’
- c. A Marta la saludó Clooney por la calle.
 to Marta CL.3.ACC.FEM greeted on the street
 ‘Clooney greeted Marta on the street.’
- d. #La noticia de su premio, se la dieron a Marta.
 the news of her prize, CL CL.3.ACC.FEM gave to Marta

In EP the corresponding examples also show an asymmetry (thanks to Pilar Barbosa, p.c.): The one corresponding to (34b) is more acceptable than the one corresponding to (34d), although neither is very good. Dislocated experiencers, on the other hand, are appropriate in this context, so (35b) is a good answer to a question like (35a) in a context in which *Maria* is yelling on the phone:

- (35) a. O que aconteceu? (EP)
 the what happened
 ‘What happened?’
- b. A Maria não lhe convêm que as crianças venham hoje.
 the Mary not CL is convenient that the children come today
 ‘It’s not convenient for Maria that the children come today.’

Catalan seems to have a pattern similar to the one quoted for Spanish, according to Solà (1992) (quoted in Barbosa, 1995, 26). Thus, both examples in (36) are possible responses to a question like *What is happening?* (thanks to Francisco Ordóñez for discussing the Catalan data).

- (36) a. A en Joan li han robat la cartera. (Catalan)
 to the Joan CL have robbed the wallet
 ‘They have stolen Joan’s wallet.’

- b. Al nen l' ha mossegat una rata.
 to-the child CL has bitten a rat
 A rat has bitten the child.'

(quoted in Barbosa, 1995, 26)

These examples from Spanish, Catalan and EP suggest that animate CLLD phrases (including subjects, overtly topicalized subjects and experiencers) are compatible with a clause-wide non-presupposed interpretation, whereas inanimate CLLD phrases are not. One logical possibility would be that the incompatibility of inanimate CLLD phrases with a clause-wide non-presupposed interpretation follows from the fact that they must be necessarily presupposed, i.e. interpreted, whereas the other three elements (subjects, overtly topicalized subjects and animate CLLD phrases) can be non-presupposed. To the extent that one assumes a uniform and strong analysis of PS-as-CLLD that includes the topic-like properties of CLLD phrases, these contrasts are problematic.

Arregi (2003) points out that in the context of a question like (37a), there are two possible answers, (37b) and (37c), but whereas the first answer is a complete answer to the question, the second one presupposes that the speaker gave things to other people.

- (37) a. ¿Qué le diste a Juan? (Spanish)
 what CL.3.SG gave to Juan
 'What did you give to Juan?'
 b. Le di un libro (a Juan).
 CL.3.SG gave a book to Juan
 'I gave Juan a book.'
 c. A Juan, le di un libro.
 to Juan CL.3.SG gave a book
 'Juan, I gave him a book.'

(from Arregi, 2003, ex. 6)

PS, on the other hand, are not possible answers to a parallel type of question, as seen in (38). Thus, the answer to (38a) could be (38b) with a PS. A postverbal subject, on the other hand, is not appropriate (cf. (38c)) with regular stress assigned to the subject. Note that (37c) contrasts with (38a) in that the first example has the presupposition just described, but the second one does not have anything similar.

- (38) a. ¿Qué le dio Ana a Juan? (Spanish)
 what CL.3.SG gave Ana to Juan
 'What did Ana give Juan?'
 b. Ana le dio un libro (a Juan).
 Ana CL.3.SG gave the book to Juan
 'Ana gave him a book.'
 c. #Le dio un libro Ana.
 CL.3.SG gave a book Ana
 'Ana gave him a book.'

These contrasts suggest that the CLLD phrase in (37c) induces a contrastive reading (i.e. the presupposition was that no one but Juan was given a book, but the answer denies this presupposition), whereas the PS in (38b) does not.

8.2.3 Possible antecedents

Cardinaletti (1997, 44) notes a certain asymmetry between CLLD phrases and PS with respect to what kind of antecedent they can pick out from the preceding clause. In (39a), the subject of the second clause (*il regista* ‘the director’) is left-dislocated, since it appears before a CLLD object *il premio* ‘the award’. In this context, coreference with the object of the first clause *Wim Wenders* is not possible. In the second example, the subject of the second clause, *il regista* ‘the director’ is not dislocated, since it is adjacent to the verb, and coreference with the preceding object *Wim Wenders* is possible.

- (39) a. Ha-nno premiato un film su Wim Wenders_i. Dopo la proiezione, il regista_{*i/j}, il premio, l’ ha ricevuto dal ministro. (Italian)
 have-3.PL given a prize to a film about Wim Wenders after the show the director the award CL.3.SG has received from-the minister
 ‘A film about Wim Wenders_i has been awarded a prize. After the show, the director_{*i/j} received the prize from the Minister.’
- b. Ho visto ieri alla Biennale un film su Wim Wenders_i. Dopo la proiezione, il regista_{i/j} ha ricevuto un premio alla carriera.
 have.1.SG seen yesterday at-the Biennale a film about Wim Wenders after the show the director has received an award to-the career
 ‘I saw a film about Wim Wenders_i yesterday at the Biennale. After the show, the director_{i/j} received a prize celebrating his career.’

(from Cardinaletti, 1997, 44, ex. 42-3)

I believe the same contrast obtains in Spanish, as seen in a simplified version of (39) in (40).

- (40) a. Ayer le dieron un premio a una película de Wim Wenders_i. Después de la proyección, el director_{*i/j}, el premio lo recibió del ministro. (Spanish)
 yesterday CL.3.DAT.SG gave a prize to a film of Wim Wenders. After of the show, the director the award CL.3.ACC.SG received of-the minister.
 ‘Yesterday, they awarded a prize to a film about Wim Wenders_i. After the show, the director_{*i/j} received the award from the Minister.’
- b. Ayer le dieron un premio a una película de Wim Wenders_i. Después de la proyección, el director_{i/j} recibió el premio del ministro.
 yesterday CL.3.DAT.SG gave a prize to a film of Wim Wenders. After of the show, the director received the award of-the minister.
 ‘Yesterday, they awarded a prize to a film about Wim Wenders_i. After the show, the director_{i/j} received the award from the Minister.’

We might try to relate this observation with Arregi’s in the preceding section. In the question/answer pair, the question introduced the IO as presupposed, so dislocating the IO in the answer forced an additional interpretation where some other people were given books. Suppose that

a similar situation takes place in (39a) and (40a), where by the second clause, Wim Wenders is now presupposed. A dislocated subject that is coreference with Wim Wenders would force a “contrastive” interpretation (i.e., something like “other directors were given other awards”), but given the definite determiners and the situation, this yields a contradictory interpretation (“the film’s only director received the night’s only award but there were other awards given to other directors”). If the subject is not dislocated, this interpretation is not forced.

8.2.4 Section Summary

To summarize these sections, there seems to be clear evidence against the strong version of the PS-as-CLLD hypothesis: On the one hand, CLLD phrases show distinct properties that challenge a unified representation as dislocated topics. On the other hand, the distribution of subjects and CLLD phrases diverge in several ways that do not warrant grouping them together with CLLD phrases. The emerging picture is that PS themselves cannot be said to occupy a single position, some may be in SubjP, others in a topic-related position.

8.3 The status of PS

The discussion from preceding sections leads me to conclude that preverbal subjects are not necessarily or typically dislocated. The proposed analysis would lead us to expect just that, if as we have argued, Infl is not pronominal, and *pro* is available as a separate category. Given these results, a NSL like Italian or Spanish can have the following potential representations.

- (41) a. [*pro* Infl]
 b. [DP Infl]
 c. [DP [*pro* Infl]]

In the first one, *pro* is valued by Infl and subsequently discourse-linked by a topic. *Pro* fulfills the Subject Criterion. In (41b), on the other hand, the DP and Infl agree, but VALUATION is vacuous, because both Infl and the DP are specified, and the DP does not need a topic to link it in discourse. The data from subject NPIs in (8.1.4) would fall into this category: the NPI must be adjacent to Infl.⁵ Finally, in (41), *pro* is valued by Infl, discourse-linked by DP, which is dislocated.

⁵For non-subject NPIs, I assume that they satisfy can satisfy the Subject Criterion, but Infl will still license a postverbal *pro*.

Bibliography

- Abney, S., 1987. The English Noun Phrase in its Sentential Aspect. Ph.D. thesis. MIT.
- Alexiadou, A., Anagnostopoulou, E., 1998. Parametrizing AGR: Word order, v-movement and EPP-checking. *Natural Language and Linguistic Theory* 16, 491–539.
- Arregi, K., 2003. Clitic left dislocation is contrastive topicalization. *Penn Working Papers in Linguistics* 9, 31–44.
- Baker, M., 2008. *The Syntax of Agreement and Concord*. Cambridge University Press, Cambridge.
- Baptista, M., 2002. *The Syntax of Cape Verdean Creole: the Sotavento varieties*. John Benjamins, Amsterdam/Philadelphia.
- Barbosa, P., 1995. *Null Subjects*. Ph.D. thesis. MIT.
- Barbosa, P., 2009. Two kinds of subject pro. *Studia Linguistica* 63, 2–58.
- Barbosa, P., 2010. Partial *pro*-drop as null np-anaphora. Presented at NELS 41, UPenn, 2010 and Romania Nova, Campos do Jordão, 2011.
- Barbosa, P., Kato, M., Duarte, M.E., 2005. Null subjects in european and brasilian portuguese. *Journal of Portuguese Linguistics* 4, 11–52.
- Bayer, J., 1984. Comp in bavarian syntax. *The Linguistic Review* 3, 209–274.
- Bayer, J., 1999. Comments on cardinaletti and starke “the typology of structural deficiency”, in: van Riemsdijk, H. (Ed.), *Clitics in the Languages of Europe*. Mouton de Gruyter, Berlin, pp. 235–242.
- Belletti, A., 1990. *Generalized Verb Movement*. Rosenberg and Sellier, Turin.
- Biberauer, T., Holmberg, A., Roberts, I., Sheehan, M., 2010. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge University Press, Cambridge, UK.
- Béjar, S., 2003. *Phi-Syntax: A Theory of Agreement*. Ph.D. thesis. University of Toronto.
- Borer, H., 1986. I-subjects. *Linguistic Inquiry* 17.

- Borer, H., 1989. Anaphoric agr, in: Jaeggli, O., Safir, K. (Eds.), *The Null Subject Parameter*. Kluwer, pp. 69–110.
- Bosque, I., 1980. *Sobre la negación*. Cátedra, Madrid.
- Bosque, I., 1987. Constricciones morfológicas sobre la coordinación. *Lingüística española actual* 9, 83–100.
- Brandi, L., Cordin, P., 1989. Two italian dialects and teh null subject parameter, in: Jaeggli, O., Safir, K. (Eds.), *The Null Subject Parameter*. Kluwer, Dordrecht, pp. 111–142.
- Bullock, B.E., Toribio, A.J., 2009. Reconsidering dominican spanish: Data from the rural cibao. *Revista Internacional de Linguística Iberoamericana* , 49–73.
- Burzio, L., 1986. *Italian Syntax*. Riedel, Dordrecht.
- Cabrera, M.J., 2007. *Null subject patterns in language contact: the case of Dominican Spanish*. Ph.D. thesis. Rutgers University.
- Camacho, J., 2001. On the interpretation of focus features, in: Herschenson, J., Mallén, E., Zagona, K. (Eds.), *Features and Interfaces*. Benjamins, Philadelphia, pp. 39–48.
- Camacho, J., 2003. *The Structure of Coordination*. Kluwer Academic Press, Dordrecht.
- Camacho, J., 2006. Do subjects have a place in spanish, in: Montreuil, J.P., Nishida, C. (Eds.), *New Perspectives in Romance Linguistics*. John Benjamins, Amsterdam/Philadelphia, pp. 51–66.
- Camacho, J., 2008. Variation in generative grammar: State of the art. *Studies in Hispanic and Lushophone Linguistics* 1.
- Camacho, J., 2010a. The null subject parameter revisited: Spanish and portuguese in discourse. Paper presented at the IVth Romania Nova Workshop. Campos do Jordão, Brazil.
- Camacho, J., 2010b. On case concord: The syntax of switch-reference clauses in Shipibo. *Natural Language and Linguistic Theory* 28, 239–274.
- Camacho, J., 2010c. On the asymmetry of root vs. embedded clauses. Evidence from Shipibo 2nd-position clitics, in: Di Sciullo, A.M., Hill, V. (Eds.), *Edges, Heads and Projections*, pp. 215–232.
- Camacho, J., 2011. Chinese-type pro in a romance-type null-subject language. *Lingua* 121, 987–1008.
- Camacho, J., Elías-Ulloa, J., 2010. Null subject systems in Shipibo switch-reference, in: Camacho, J., Guitérrez-Bravo, R., Sánchez, L. (Eds.), *Information Structure in Languages of the Americas*. Mouton deGruyter, Berlin, pp. 65–85.

- Cardinaletti, A., 1997. Subjects and clause structure, in: Haegeman, L. (Ed.), *The New Comparative Syntax*. Longman, London/New York, pp. 33–63.
- Cardinaletti, A., 2004. Towards a cartography of subject positions, in: Rizzi, L. (Ed.), *The Cartography of Syntactic Structures, Volume 2*. Oxford University Press, Oxford/New York, pp. 115–165.
- Cardinaletti, A., Starke, M., 1999. The typology of structural deficiency, in: Riemsdijk, H.V. (Ed.), *Clitics in the Languages of Europe (Empirical Approaches to Language Typology, 20; Eurotyp, 5)*. Mouton, Berlin, pp. 145–233.
- Chierchia, G., 1998. Reference to kinds across language. *Natural Language Semantics* 6, 339–405. 10.1023/A:1008324218506.
- Chomsky, N., 1975. *Reflections on Language*. Pantheon Books, New York.
- Chomsky, N., 1981. *Lectures on Government and Binding*. Foris, Dordrecht.
- Chomsky, N., 1986. *Knowledge of Language : Its Nature, Origins, and Use*. Praeger, New York.
- Chomsky, N., 1988. *Language and Problems of Knowledge: The Managua Lectures*. MIT Press, Cambridge, MA.
- Chomsky, N., 1993. A minimalist program for linguistic theory, in: Hale, K., Keyser, S. (Eds.), *The View from Building 20*. MIT Press, Cambridge, MA, pp. 1–52.
- Chomsky, N., 1995. *The Minimalist Program*. MIT Press.
- Chomsky, N., 2000. Minimalist inquiries: The framework., in: Martin, R., Michaels, D., Uriagereka, J. (Eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. MIT Press, Cambridge, MA, pp. 89–156.
- Chomsky, N., 2001. Derivation by phase., in: Kenstowicz, M., Keyser, S. (Eds.), *Ken Hale: A Life in Language*. MIT Press, Cambridge, MA, pp. 1–52.
- Chomsky, N., Lasnik, H., 1977. Filters and control. *Linguistic Inquiry* 8, 425–504.
- Chung, S., 1998. *The design of agreement: Evidence from Chamorro*. University of Chicago, Chicago.
- Chung, S., 2003. The syntax and prosody of weak pronouns in chamorro. *Linguistic Inquiry* 34, 547–599.
- Cinque, G., 1990. *Types of A-bar Dependencies*. MIT Press, Cambridge, Mass.
- Clamons, R., Mulkern, A.E., Sanders, G., Stenson, N., 1999. The limits of formal analysis: Pragmatic motivation in oromo grammar., in: Darnell, M., Moravcsik, E.A., Newmeyer, F., Noonan, M., Wheatley, K.M. (Eds.), *Functionalism and Formalism in Linguistics, II: Case Studies*. Benjamins. *Studies in Language Companion Series (SLCS)*: 41, pp. 59–76.

- Cole, M., 2000. The Syntax, Morphology, and Semantics of Null Subjects. Ph.D. thesis. University of Manchester. Manchester.
- Cole, M., 2009. Null subjects: A reanalysis of the data. *Linguistics* 47, 559–587.
- Contreras, H., 1991. On the position of subjects, in: Rothstein, S. (Ed.), *Perspectives on Phrase Structure: Heads and Licensing*. *Syntax and Semantics* 25.. Academic Press, San Diego, pp. 63–79.
- Costa, J.a., 2001. Marked versus unmarked inversion and optimality theory, in: Hulk, A., Pollock, J.Y. (Eds.), *Subject Inversion and the Theory of Universal Grammar*. Oxford University Press, Oxford, U.K., pp. 91–106.
- Croft, W., 1988. Agreement vs. case marking and direct objects, in: Barlow, M., Ferguson, C.A. (Eds.), *Agreement in Natural Language*. CSLI, Stanford, CA, pp. 159–179.
- Dasgupta, P., 2003. Bangla, in: Cardona, G., Jain, D. (Eds.), *The Indo-Aryan languages*. Routledge, London and New York, pp. 386–428.
- Déchaine, R.M., Wiltschko, M., . Decomposing pronouns. *Linguistic Inquiry* 33, 409–442.
- Diercks, M., 2009. Subject extraction and (so-called) anti-agreement effects in lubukusu: A criterial freezing approach. Georgetown University.
- Diercks, M., 2010. Agreement with Subjects in Lubukusu. Ph.D. thesis. Georgetown University.
- Duarte, M.E., 1993. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do brasil, in: Roberts, I., Kato, M. (Eds.), *Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil*. Editora da UNICAMP, pp. 107–128.
- Duarte, M.E., 1995. A Perda do Princípio “Evite pronome” no Português Brasileiro. Ph.D. thesis. UNICAMP.
- Duarte, M.E., 2000. The loss of the avoid pronoun principle in brazilian portuguese, in: Kato, M.A., negrão, E. (Eds.), *Brazilian Portuguese and the Null Subject Parameter*. Vervuert-Iberoamericana, Frankfurt, pp. 17–36.
- Elías-Ulloa, J., 2002. Subject doubling in Capanahua. Rutgers University.
- Fassi Fehri, A., 1993. Issues in the Structure of Arabic Clauses and Words. volume 29 of *Studies in Natural Language and Linguistic Theory*. Kluwer, Dordrecht.
- Figueiredo Silva, M.C., 1994. La position sujet en Portugais Brésilien. Ph.D. thesis. Université de Genève. Geneva.
- Frascarelli, M., 2007. Subjects, topics and the interpretation of referential pro. an interface approach to the linking of (null) pronouns. *Natural Language and Linguistic Theory* 25, 691–734.

- Frascarelli, M., Hinterhölzl, R., 2007. Types of topics in German and Italian, in: Winkler, S., Schwabe, K. (Eds.), *On information structure, meaning and form*. John Benjamins, Amsterdam/Philadelphia, p. 87–116.
- Fuß, E., 2005. *The Rise of Agreement*. John Benjamins, Amsterdam/Philadelphia.
- Fukui, N., 1988. Deriving the difference between English and Japanese: A case study in parametric syntax. *English Linguistics* 5, 249–270.
- Gilligan, G., 1987. *A Cross-linguistic Approach to the Pro-drop Parameter*. Ph.D. thesis. University of Southern California. Los Angeles, CA.
- Giupponi, E., 1988. *Pro Drop Parameter und Restrukturierung im Trentino*. Ph.D. thesis. University of Vienna.
- Givón, T., 1983. Topic continuity in discourse: An introduction, in: Givón, T. (Ed.), *Topic Continuity in Discourse*. John Benjamins, Amsterdam/Philadelphia, pp. 1–42.
- Goodall, G., 1999. On preverbal subjects in Spanish, in: Satterfield, T., Christina, T., Cresti, D. (Eds.), *Current Issues in Romance Languages*. John Benjamins, Amsterdam, pp. 95–110.
- Grimshaw, J., Samek-Lodovici, V., 1998. Optimal subjects and subject universals, in: Barbosa, P., Fox, D., Hagstrom, P., Pesetsky, D. (Eds.), *Is the Best Good Enough? Optimality and Competition in Syntax*. MIT Press, pp. 193–219.
- Gutman, E., 2004. Third person null subjects in Hebrew, Finnish and Rumanian: an accessibility-theoretic account. *Journal of Linguistics* 40, 463–490.
- Hale, K., 1983. Papago (k)c. *International Journal of American Linguistics* 49, 299–327.
- Harley, H., Ritter, E., 2002. A feature-geometric analysis of person and number. *Language* 78, 482–526.
- Henríquez Ureña, P., 1939. *Ello*. *Revista de Filología Hispánica* 1, 209–229.
- Henríquez Ureña, P., 1940/1975. *El español en Santo Domingo*. El Taller, Santo Domingo.
- Hernanz, M.L., 1990. En torno a los sujetos arbitrarios: La segunda persona del singular, in: Violeta, D., Garza Cuarón, B. (Eds.), *Estudios de lingüística de España y México*. Colegio de México, México, pp. 151–178.
- Hinzelin, M.O., Kaiser, G.A., 2006. *Das neutrale Pronomen *ello* im dominikanischen Spanisch und die Nullsubjekteigenschaft. Anhang: Korpus und Bibliographie zu “ello” und unpersönlichen Konstruktionen (mit Subjekt) im Spanischen*.
- Hinzelin, M.O., Kaiser, G.A., 2007. *El pronombre ello en el léxico del español dominicano*, in: Mihatsch, W., Sokol, M. (Eds.), *Language contact and language change in the Caribbean and beyond*. Peter Lang, Frankfurt, pp. 171–188.

- Hoekstra, T., 1983. The distribution of sentential complements, in: Bennis, H., van Lessen Kloeke, W. (Eds.), *Linguistics in the Netherlands 1983*. Foris, Dordrecht, pp. 93–103.
- Hoji, H., 1998. Null object and sloppy identity in Japanese. *Linguistic Inquiry* 29, 127–152.
- Holmberg, A., 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36, p533 – 564.
- Holmberg, A., Nayudu, A., Sheehan, M., . Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63, 59–97.
- Holmberg, A., Nikkane, U., 2002. Expletives, subjects and topics in Finnish, in: *Subjects, Expletives and the EPP*. Oxford University Press, pp. 71–106.
- Holmberg, A., Sheehan, M., 2010. Control into finite clauses in partial null-subject languages, in: Biberauer, T., Holmberg, A., Roberts, I., Sheehan, M. (Eds.), *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge University Press, pp. 125–152.
- Huang, C.J., 1982. Move wh in a language without wh movement. *The Linguistic Review* 1, 369–416.
- Huang, C.T.J., 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15, 531–574.
- Huang, C.T.J., 1989. Pro-drop in Chinese: A generalized control theory, in: Jaeggli, O., Safir, K. (Eds.), *The Null Subject Parameter*. Kluwer, Dordrecht, pp. 185–214.
- Huang, Y., 1992. Against Chomsky's typology of empty categories. *Journal of Pragmatics* 17, 1–29.
- Huang, Y., 2000. *Anaphora: A cross-linguistic study*. Oxford University Press, Oxford.
- Huidobro, S., 2005. Phonological constraints on *verum focus* in Argentinian Spanish. SUNY Stony Brook.
- Hyams, N., 1986. *Language Acquisition and the Theory of Parameters*. Reidel, Dordrecht.
- Hyams, N., 1992. A reanalysis of null subjects in child language, in: Weissenborn, J., Goodluck, H., Roeper, T. (Eds.), *Theoretical Issues in Language Acquisition*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 249–268.
- Hyams, N., Wexler, K., 1993. On the grammatical basis of null subjects in child language. *Linguistic Inquiry* 24, 421–459.
- Iatridou, S., Embick, D., 1997. *Apropos pro*. *Language* 73.
- Jaeggli, O., 1980. On some phonologically-null elements in syntax. Ph.D. thesis. Massachusetts Institute of Technology.

- Jaeggli, O., 1982. *Topics in Romance Syntax*. Foris, Dordrecht.
- Jaeggli, O., 1986. Arbitrary plural pronouns. *Natural Language and Linguistic Theory* 4, 43–76.
- Jaeggli, O., 1987. Ecp effects at *lf* in spanish, in: Birdsong, D., Montreuil, J.P. (Eds.), *Advances in Romance linguistics*. Foris, Dordrecht, pp. 113–151.
- Jaeggli, O., Safir, K.J., 1989. The null subject parameter and parametric theory, in: Jaeggli, O., Safir, K.J. (Eds.), *The Null Subject Parameter*. Reidel, Dordrecht, pp. 1–44.
- Jelinek, E., 1984. Empty categories, case and configurationality. *Natural Language and Linguistic Theory* 2, 39–76.
- Johns, C., 2007. *Interpreting Agreement*. Ph.D. thesis. University of Durham.
- Kanno, K., 1997. The acquisition of null and overt pronominals in japanese by english speakers. *Second Language Research* 13, 265–287.
- Kato, M., 1999. Strong pronouns and weak pronominals in the null subject parameter. *Probus* 11, 1–37.
- Kato, M.A., 2000. The partial pro-drop nature and the restricted vs order in brazilian portuguese, in: Kato, M.A., Negrão, E. (Eds.), *Brazilian Portuguese and the Null Subject Parameter*. Vervuert, pp. 223–255.
- Kato, M.A., Negrão, E., 2000. *Brazilian Portuguese and the Null Subject Parameter*. Vervuert.
- Kayne, R., 1980. Extensions of binding and case-marking. *Linguistic Inquiry* 11, 75–96.
- Kenstowicz, M., 1989. The null subject parameter in modern arabic dialects, in: *The Null Subject Parameter*. Dordrecht, Dordrecht, pp. 263–275.
- Kiss, K., 2002. The *epp* in a topic-prominent language, in: Svenonius, P. (Ed.), *Subjects, Expletives, and the EPP*. Oxford University Press, Oxford, UK, pp. 107–124.
- Koopman, H., Sportiche, D., 1991. The position of subjects. *Lingua* 85, 211–258. Special issue edited by J. McCloskey.
- Laka, I., 1990. *Negation in Syntax. On the Nature of Functional Categories and Projections*. Ph.D. thesis. MIT. Cambridge, Mass.
- Lakoff, G., 1986. Frame semantic control of the *csc*, in: et al., A.M.F. (Ed.), *CLS 22 part 2. Papers from the Parasession on Pragmatics and Grammatical Relations*, University of Chicago, Chicago, Ill.
- Larson, R., Luján, M., 1989. *Emphatic pronouns*. Ms. Stony Brook.
- Li, C.N., Thompson, S.A., 1976. Subject and topic: A new typology of language, in: Li, C.N. (Ed.), *Subject and Topic*. Academic Press, New York, pp. 458–489.

- Li, Y.H.A., 2007. Beyond empty categories. *Bulletin of the Chinese Linguistic Society of Japan* 254, 74–106.
- Liceras, J., 1989. On some properties of the pro-drop parameter: Looking for missing subjects in non-native Spanish, in: Gass, Susan & Schacter, J. (Ed.), *Linguistic perspectives on second language acquisition*. Foris, Dordrecht, pp. 109–133.
- Lima, A., 2005. Predicate inversion, person asymmetries and agreement alternations in Brazilian Portuguese. <http://www.linguistics.ucla.edu/people/grads/lima/LASSOHandout.pdf>. Handout from LASSO 34.
- Loriot, J., Lauriault, E., Day, D., 1993. *Diccionario shipibo-castellano*. Ministerio de Educación-Instituto Lingüístico de Verano, Lima.
- Lozano, C., 2002. Universal grammar and focus constraints: Focus, pronouns and word order in the acquisition of L2 and L3 Spanish. Ph.D. thesis. University of Essex. Essex.
- Manzini, M.R., Savoia, L.M., 2002. Parameters of subject inflection in Italian dialects, in: Svenonius, P. (Ed.), *Subjects, Expletives, and the EPP*. Oxford University Press, Oxford, UK, pp. 157–200.
- McClelland, C.W., 1996. Interrelations of Prosody, Clause Structure and Discourse Pragmatics in Tarifit Berber. Ph.D. thesis. University of Texas, Arlington. Arlington, TX.
- McCloskey, J., 1991. Clause structure, ellipsis and proper government in Irish. *Lingua* 85, 259–302. Special issue edited by J. McCloskey.
- McCloskey, J., 1996. Subjects and subject positions in Irish, in: Borsley, R., Roberts, I. (Eds.), *Celtic and Beyond: Papers on Celtic Syntax*. Cambridge University Press, Cambridge, U.K., pp. 241–283.
- McCloskey, J., 1997. Subjecthood and subject positions, in: Haegeman, L. (Ed.), *Elements of Grammar*. Kluwer, Dordrecht, pp. 197–235.
- McCloskey, J., Hale, K., 1984. On the syntax of person-number inflection in modern Irish. *Natural Language and Linguistic Theory* 1, 487–533.
- Menuzzi, S.d.M., 2000. That-trace effects in Portuguese. *Fórum Lingüístico* 2, 13–39.
- Modesto, M., 2000. Null subject without ‘rich’ agreement, in: Kato, M.A., Negrão, E. (Eds.), *The Null Subject Parameter in Brazilian Portuguese*. Verveurt-Iberoamericana, Frankfurt/Madrid, pp. 147–174.
- Modesto, M., 2008. Topic prominence and null subjects, in: Biberauer, T. (Ed.), *The Limits of Syntactic Variation*. John Benjamins, Amsterdam, pp. 375–409.
- Montalbetti, M., 1984. *After Binding*. Ph.D. thesis. MIT.

- Moro, A., 1993. *I predicati nominali e la struttura della frase*. Unipress, Padua.
- Muriungi, P., 2010. An anti-locality restriction on subject wh-phrases in kîtharaka. *Lingua In Press, Corrected Proof*, –.
- Nicolis, M., 2005. *On pro-drop*. Ph.D. thesis. Università degli Studi di Siena. Siena, Italy.
- Nicolis, M., 2008. The null subject parameter and correlating properties: the case of creole languages, in: Biberauer, T. (Ed.), *The Limits of Syntactic Variation*. John Benjamins, Amsterdam/Philadelphia, pp. 271–94.
- Noguchi, T., 1997. Two types of pronouns and variable binding. *Language* 73, 770–797.
- Ordóñez, F., 1997. *Word Order and Clause Structure in Spanish and Other Romance Languages*. Ph.D. thesis. CUNY. New York.
- Ordóñez, F., Olarrea, A., 2006. Microvariation in caribbean/non caribbean Spanish interrogatives. *Probus: International Journal of Latin and Romance Linguistics* 18, 59 – 96.
- Ordóñez, F., Treviño, E., 1999. Left dislocated subjects and the pro-drop parameter: A case study of Spanish. *Lingua* 107, 39–68.
- Ortiz López, L.A., 2009. El español del caribe: orden de palabras a la luz de la interfaz léxico-sintaxis y sintaxis-pragmática. *Revista Internacional de Lingüística Iberoamericana* 7, 75–94.
- Otheguy, R., Zentella, A.C., Livert, D., 2007. Language and dialect contact in Spanish in new york: Towards the forma of a speech community. *Language* 83, 770–802.
- Ouhalla, J., 1988. A note on bound pronouns. *Linguistic Inquiry* 19, 485–494.
- Perlmutter, D., 1971. *Deep and Surface Structure Constraints in Syntax*. Holt, Rinehard and Winston.
- Perrott, D.V., 1972. *Teach Yourself Swahili*. David McKay Company, New York.
- Pesetsky, D., Torrego, E., 2001. T-to-C movement: Causes and consequences, in: Kenstowicz, M., Keyser, S. (Eds.), *Ken Hale: A Life in Language*. MIT Press, Cambridge, MA, pp. 355–426.
- Pesetsky, D., Torrego, E., 2004. Tense, case and the nature of syntactic categories, in: Guéron, J., Lecarme, J. (Eds.), *The syntax of time*. MIT Press, Cambridge, pp. 495–538.
- Pesetsky, D., Torrego, E., 2006. Probes, goals and syntactic categories, in: Otsu, Y. (Ed.), *Proceedings of the Seventh Tokyo Conference on Psycholinguistics*. Hituzi Syobo Publishing Company, Tokio, pp. 25–60.
- Pesetsky, D., Torrego, E., 2007. The syntax of valuation and the interpretability of features, in: Karimi, S., Samiian, V., Wilkins, W. (Eds.), *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*. Benjamins, Amsterdam, pp. 262–294.

- Platzack, C., 1987. The scandinavian languages and the null-subject parameter. *Natural Language and Linguistic Theory* 5, 377–401.
- Platzack, C., 2003. Agreement and null subjects. *Nordlyd* 31, 326–355.
- Poletto, C., 1993. *La sintassi del soggetto nei dialetti dell'Italia settentrionale*. Unipress, Padova.
- Poletto, C., 2000. *The Higher Functional Field. Evidence from Northern Italian Dialects*. Oxford University Press, Oxford.
- Pérez Jiménez, I., 2007. *Las cláusulas absolutas*. Visor Libros, Madrid.
- Rice, K., 1989. *A Grammar of Slave*. Mouton de Gruyter, Berlin.
- van Riemsdijk, H., Williams, E., 1986. *Introduction to the Theory of Grammar*. MIT P, Cambridge, MA.
- Rigau, G., 1988. Strong pronouns. *Linguistic Inquiry* 19, 503–511.
- Rizzi, L., 1982. *Issues in Italian Syntax*. Foris, Dordrecht.
- Rizzi, L., 1986a. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17, 501–557.
- Rizzi, L., 1986b. On the status of subjects clitics in romance, in: Jaeggli, O., Silva-Corvalán, C. (Eds.), *Studies in Romance syntax*. Foris, Dordrecht, pp. 391–419.
- Rizzi, L., 1987. Three remarks on null subjects. Talk given at the Workshop on Dialectology, GLOW X, University of Venice.
- Rizzi, L., 1996. Residual verb second and the wh-criterion, in: Belletti, A., Rizzi, L. (Eds.), *Parameters and functional heads*. Oxford University Press, New York. volume 2, pp. 63–90.
- Rizzi, L., 1997. The fine structure of the left periphery, in: Haegeman, L. (Ed.), *Elements of Grammar*. Kluwer Academic Publishers, Dordrecht, pp. 281–337.
- Rizzi, L., 2006. On the form of chains: Criterial positions and ecp effects, in: Cheng, L., Corver, N. (Eds.), *Wh Movement: Moving on*. MIT Press, Cambridge, MA, p. 97–133.
- Rizzi, L., Shlonsky, U., 2007. Strategies of subject extraction, in: Sauerland, U., Gärtner, H.M. (Eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*. Mouton de Gruyter. *Studies in Generative Grammar (StGG)*: 89, pp. 115 – 160.
- Roberts, I., 2010a. A deletion analysis of null subject, in: Biberauer, T., Holmberg, A., Roberts, I., Sheehan, M. (Eds.), *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge University Press, Cambridge, UK, pp. 58–87.
- Roberts, I., 2010b. Varieties of french and the null subject parameter, in: Biberauer, T., Holmberg, A., Roberts, I., Sheehan, M. (Eds.), *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge University Press, Cambridge, pp. 303–327.

- Roberts, I., Holmberg, A., 2010. Introduction: parameters in minimalist theory, in: Biberauer, T., Holmberg, A., Roberts, I., Sheehan, M. (Eds.), *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge University Press, Cambridge, UK, pp. 1–57.
- Rodríguez-Mondoñedo, M., 2006. Spanish existentials and other accusative constructions, in: Boeckx, C. (Ed.), *Minimalist Essays*. J. Benjamins, Amsterdam, pp. 326–394.
- Rosenkvist, H., 2009. Null referential subjects in germanic languages – an overview. *Working Papers in Scandinavian Syntax* 84, 151–180.
- Rosenkvist, H., 2010a. Null referential subjects in germanic languages – an overview. *Subjects in Diachrony Conference*.
- Rosenkvist, H., 2010b. Null referential subjects in Övdalian. *Nordic Journal of Linguistics* 33, 231–267. Doi:10.1017/S033258651000020X.
- Saab, A., 2009. *Hacia una teoría de la identidad parcial en la elipsis*. Ph.D. thesis. Universidad de Buenos Aires.
- Saab, A., 2010. Null subjects at the syntax-morphology connection. Paper presented at *Romania Nova IV*, Brazil.
- Safir, K., 1985. Missing subjects in German, in: Toman, J. (Ed.), *Studies in German grammar*. Foris, Dordrecht, pp. 193–230.
- Safir, K., 2004. Person, context and perspective. *Italian Journal of Linguistics (Rivista di Linguistica)* 16, 107–153.
- Samek-Lodovici, V., 1996. Structural contrastive focus in Italian with and without emphatic stress. *Universitat Konstanz*.
- Sánchez, L., 1994. On the interpretation of intensified dps and emphatic pronouns, in: Mazzola, M.L. (Ed.), *Issues and Theory in Romance Linguistics. Selected Papers from the Linguistic Symposium on Romance Languages 23*. Georgetown UP, Georgetown, pp. 479–491.
- Schlenker, P., 2003. A plea for monsters. *Linguistics and Philosophy* 26, 29–120.
- Sengupta, G., 1999. Lexical anaphors and pronouns in Bangla, in: Lust, B., Wali, K., Gair, J., Subbarao, K. (Eds.), *Lexical anaphors and pronouns in selected South Asian languages: A principled typology*. Mouton de Gruyter, Berlin, pp. 277–332.
- Sheehan, M., 2007. *The EPP and Null Subjects in Romance*. Ph.D. thesis. Newcastle University.
- Shlonsky, U., 1997. *Clause Structure and Word Order in Hebrew and Arabic*. Oxford University Press, Oxford.
- Shlonsky, U., 2009. Hebrew as a partial null-subject language. *Studia Linguistica* 63, 133–157.

- Sigurdsson, H.A., 2004. The syntax of person, tense, and speech features. *Italian Journal of Linguistics* 16, 219–251.
- Sigurðsson, H.r., 2010. On epp effects. *Studia Linguistica* 64, 159–189.
- Sigurðsson, H.r., 2011. Conditions on argument drop. *Linguistic Inquiry* 42, 267–304.
- Sigurðsson, H.r., Holmberg, A., 2008. Icelandic dative intervention: Person and number are separate probes, in: D’Alessandro, Roberta, F.S., Hrafnbjargarson, G.H. (Eds.), *Agreement Restrictions*. Mouton de Gruyter, Berlin, p. 251–279.
- Silva-Villar, L., 1998. Subject positions and the roles of cp, in: Schwegler, A., Tranel, B., Uribe-Etxebarria, M. (Eds.), *Romance Linguistics: Theoretical Perspectives*. Benjamins. Amsterdam Studies in the Theory and History of Linguistic Science IV: Current Issues in Linguistic Theory (CILT): 160, pp. 247 – 270.
- Sánchez, L., 2010. *The Morphology and Syntax of Focus and Topic: Minimalist Inquiries in the Quechua periphery*. John Benjamins, Amsterdam.
- Solà, J., 1992. *Agreement and subjects*. Ph.D. thesis. Universitat Autònoma de Barcelona. Barcelona, Spain.
- Speas, M., 1995. *Economy, agreement and the representation of null arguments*. University of Massachusetts.
- Speas, M., 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114, 255–276.
- Sportiche, D., 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19, 425–49.
- Suñer, M., 1983. *pro_{arb}*. *Linguistic Inquiry* 14, 188–191.
- Suñer, M., 1994. V-movement and the licensing of argumental wh-phrases in Spanish. *Natural Language and Linguistic Theory* 12, 335–372.
- Suñer, M., 2003. The lexical preverbal subject in a romance null subject language:: Where are thou?, in: Núñez-Cedeño, R., López, L., Cameron, R. (Eds.), *A Romance Perspective on Language Knowledge and Use. Selected Papers from the 31st Linguistic Symposium on Romance Languages (LSRL)*. John Benjamins, Amsterdam, pp. 341–58.
- Svenonius, P., 2002. Subjects, expletives, and the epp, in: Svenonius, P. (Ed.), *Subjects, Expletives, and the EPP*. Oxford University Press, Oxford, UK, pp. 3–28.
- Szczegielniak, A., 1999. ‘that-t effects’ crosslinguistically and successive cyclic movement, in: *Papers on Morphology and Syntax, Cycle One*. Cambridge, MA. volume 33, pp. 369–393.

- Taraldsen, T., 1980. On the nic, vacuous application and the *that-trace* filter. MIT and Indiana Linguistics Club.
- Tomioka, S., 1999. A sloppy identity puzzle. *Natural Language Semantics* 7, 217–241.
- Tomioka, S., 2003. The semantics of japanese null pronouns and its cross-linguistic implications, in: Schwabe, K., Winkler, S. (Eds.), *The Interfaces: Deriving and Interpreting Omitted Structures*. John Benjamins, Amsterdam, pp. 321–340.
- Toribio, A.J., 1993. Lexical subjects in finite and non-finite clauses. *Cornell Working Papers in Linguistics* 11, 149 – 178.
- Vainikka, A., Levy, Y., 1999. Empty subjects in finnish and hebrew. *Natural Language and Linguistic Theory* 17, 613–671.
- Valenzuela, P., 2003a. Evidentiality in Shipibo-Konibo, with a comparative overview of the category in Panoan, in: Aikhenvald, A., Dixon, R. (Eds.), *Studies in Evidentiality*. John Benjamins, Amsterdam, pp. 33–62.
- Valenzuela, P., 2003b. Transitivity in Shipibo-Konibo Grammar. Ph.D. thesis. University of Oregon.
- Vergnaud, J.R., Zubizarreta, M.L., 1992. The definite determiner and the inalienable constructions in french and english. *Linguistic Inquiry* 23, 595–652.
- Wasike, A., 2007. The left periphery, wh-in-situ and A-bar movement in Lubukusu and other Bantu languages. Ph.D. thesis. Cornell University.
- Yamada, K., 2005. The status of the overt pronoun constraint in grammatical theory and sla of japanese. *Essex Graduate Student Papers in Language & Linguistics* 7, 180–201.
- Zubizarreta, M.L., 1994. The grammatical representation of topic and focus: Implications for the structure of the clause. *University of Venice Working Papers in Linguistics* 4, 97–126.
- Zubizarreta, M.L., 1998. *Prosody, Focus, and Word Order*. MIT Press, Cambridge, Mass.
- Zushi, M., 2003. Null arguments: the case of japanese and romance. *Lingua* 113, 559 – 604. *Formal Japanese syntax and universal grammar: the past 20 years*.