

Chapter 1 On the null-subject phenomenon.

Federica Cognola, La Sapienza University, Rome
federica.cognola@uniroma1.it
Jan Casalicchio, University of Utrecht/University of Konstanz
jan.casalicchio@gmail.com

1.1. The null-subject phenomenon*

In his 2004 paper, Newmeyer takes the null-subject parameter as a case study to illustrate the failure of the Principle and Parameter approach to account for typological variation. He writes:

Despite the importance attributed to them [sc. parameters] in the past two decades of work in generative grammar, there is little reason to believe that parametrized principles play a role in the theory of UG. The bulk of their work should be taken over by the interplay of unparametrized UG principles, language particular rules, and processing principles sensitive to structural relations holding among grammatical elements (Newmeyer 2004: 225).

Criticism of the null-subject parameter is also found in Baker (2008: 352) who writes that “[h]istory has not been kind to the Pro-drop Parameter as originally stated” (cf. also Camacho 2013a and Boeckx 2016).

Despite these challenges to the notion of a pro-drop parameter, many researchers working within the framework of Generative Grammar have continued to investigate the null-subject phenomenon. See, for example, Holmberg (2005), Frascarelli (2007), Neeleman and Szendroi (2005, 2007), Biberauer (2008b), Biberauer, Holmberg, Roberts and Sheehan (2010), Camacho (2013a), Barbosa (2013), Duguine (2014) and D’Alessandro (2015), from a synchronic perspective; and work by, among others, Madariaga (2011), Walkden (2012, 2013, 2014), Schlachter (2012), Zimmermann (2014), Kinn (2016a), Volodina and Weiß (2016) and Wolfe (2015), from a diachronic perspective.

This volume contains a collection of twelve papers on the latest research into null-subject languages, in which all aspects relevant to the complex and multifaceted phenomenon of null subjects are addressed and discussed, using novel data and novel theoretical analyses.

This chapter provides an overview of the volume and describes the latest developments with regard to both the particular topic discussed in each chapter, and the null-subject phenomenon itself.

1.2 The null-subject parameter: original formulation and cluster properties

In its original formulation (Perlmutter 1971), the pro-drop parameter was intended to capture the empirical observation that in some languages, but not in others, a definite, referential, pronominal subject must be expressed in all finite clauses (Roberts and Holmberg 2010: 3, and for a complete historical overview of the treatment of null subjects in traditional grammar). Spanish is an example of the first group (1a), the second is exemplified by German (1b).

(1) a. *(Ella/él) se ha enfadado* (Spanish)

* We thank Theresa Biberauer, Ermenegildo Bidese, Mara Frascarelli, Michelle Sheehan, and two anonymous reviewers for very useful comments and suggestions, and Rachel Murphy for editing the English. All errors are our own. Federica Cognola takes responsibility for section 1.1-1.3, Jan Casalicchio for section 1.4.

- she/he refl.pron has got angry
 b. *(*Sie/Er*) *hat sich geärgert* (German)
 she/he has refl.pron got angry
 ‘He/she got angry.’

Overt subjects in non-null subject languages (unlike in null-subject languages) are not necessarily marked from the point of view of information structure, i.e. their being overt is mainly due to structural factors, not to pragmatic constraints. This is evidenced by, for example, the fact that such languages, unlike their null-subject counterparts, require expletives with weather predicates (2).

- (2) a. *Ha llovido* (Spanish)
 has rained
 b. *(*Es*) *hat geregnet* (German)
 EXPL has rained
 ‘It rained.’
 c. *(*It*) *rained*

Empirically, the null-subject parameter has been claimed by Rizzi (1982) to correlate with the following properties (cf. also Chomsky and Lasnik 1977, Kayne 1980, Taraldsen 1978, Jaeggli 1982, Safir 1985, Jaeggli and Safir 1989, Roberts and Holmberg 2010).

- (3) a. The possibility of a silent, referential, definite subject of finite clauses.
 b. Free subject inversion
 c. The apparent absence of complementiser-trace effects.
 d. Rich agreement inflection on finite verbs.

All the properties in (3) are found in Italian (the basis of Rizzi’s work), and they were thus considered part of a cluster that followed from a language’s null-subject character.

It has been known since Rizzi (1982, 1986a), however, that there exist different types of null subject languages, and that the properties exemplified in (3) appear in a cluster only in so-called consistent null-subject languages, like most Romance languages and Greek. Null-subject languages like Japanese, Chinese, Thai, Vietnamese represent the subtype of radical / discourse null-subject languages (Huang 1984, Greco and Trang and Haegeman *this volume*). In these languages objects can also be dropped along with subjects and the null category is not licensed and identified via agreement, since these languages lack agreement inflection on finite verbs (property 3d in the cluster), but via discourse and context (thus: discourse/radical null-subject languages, see section 1.2.2 below).

As the empirical database grew (see Gilligan 1987 for the initial investigation, which has been followed by a huge amount of further study), it became evident that this partition of null-subject languages in two subtypes only differing for the presence (consistent) or absence (radical) of property (3d) in the cluster could not account for all documented null-subject systems. To this typology a third class of null-subject languages was therefore added, that of partial null-subject languages, among which Finnish (Holmberg 2005, Holmberg and Sheehan 2010), European and Brazilian Portuguese, (Barbosa 2013, Holmberg, Nikanne and Sheehan 2009), Hebrew (Shlonski 2009), Russian (Madariaga *this volume*) are found. Typical of this class of languages is that a silent, referential, definite subject is licensed in a restricted number of syntactic contexts, which appear to be to a greater extent language-specific and do not feature among the cluster properties in (3) (see section 1.2.4 below for discussion).

In this chapter, we discuss in detail all the properties ascribed to null-subject languages and address the issue of whether, in the light of recent research, they can still be considered central to the null-subject phenomenon.

1.2.1 On expletives

The idea that null-subject languages do not exhibit overt expletive subjects¹, in contrast to non-null-subject languages (cf. 2) has recently been challenged by work on a number of null-subject languages (cf. Finnish: Holmberg and Nikkanen 2002, Holmberg 2005; European Portuguese: Carrilho 2003, 2005; on Romance languages: see the contributions in Kaiser and Remberger 2009, Corr 2015b, Raposo and Uriagereka 1990, Hinzelin and Kaiser 2007, Toribio 2000, Batra-Kaufmann 2011, Camacho 2013a, Muñoz Pérez 2014; Vietnamese see Greco, Trang and Haegeman *this volume*).

These works have shown that some varieties of familiar null-subject languages, Spanish for example, can exhibit what appear to be expletive subjects, as shown in (4).

- (4) a. *Ello vienen haitianos aquí* (Dominican Spanish, Camacho 2013a: 44)
 EXPL come.3PL Haitianos here
 ‘People from Haiti come here.’
 b. *Ello llegan guaguas hasta allá* (Dominican Spanish, Camacho 2013a: 46)
 EXPL arrive.3PL buses until there
 ‘Buses arrive up there.’
 c. *El chove* (Galician, Camacho 2013a: 47)
 EXPL rain.3SG
 ‘It rains.’

As discussed by Greco, Trang and Haegeman (*this volume*), even radical pro-drop languages (i.e. the most liberal category of null-subject languages in which null subjects and objects can be dropped as long as they can be recovered from the context, see section 1.2.4 below for the discussion of the typology of null-subject languages) can exhibit elements with expletive-like properties.² In their chapter, which presents entirely new findings, the authors demonstrate that in Vietnamese the third person subject pronoun *nó* can appear in the sentence-initial position in some syntactic contexts where the presence of an expletive is generally required in non-null subject languages, such as with weather verbs, existential constructions, and presentational constructions featuring unaccusative verbs (5) (examples from Greco, Trang and Haegeman *this volume*: 11f).³

- (5) a. *Nó mưa bây giờ đấy* (Vietnamese)
 EXPL rain now PRT
 ‘It is about to rain now.’

¹ It has been claimed that null-subject languages exhibit expletive *pro*, see Rizzi (1986a) and Sheehan (*to appear-b*) among others. We do not discuss this type of null subject, examining only referential null subjects.

² In this chapter, we call the elements found in null-subject languages whose properties and distribution superficially resemble those of expletives in non-null-subject languages “expletive-like” elements (see Hinzelin and Kaiser 2007 and Carrilho 2008 for a similar position, and Camacho 2013a for the opposite idea that both types of element should be subsumed to the category of expletives). This terminological choice is based on the fact that expletive-like elements exhibit a series of properties (optionality, a connection with specific semantico-pragmatic interpretations of the sentence in which they appear, the impossibility of their appearing in the inversion position) which are incompatible with those of expletives in non-null-subject languages (see the discussion in this section). We therefore suggest that the expletives of non-null-subject languages and the expletive-like elements of null-subject languages be kept separate – an idea which implies that only the former can be considered true expletives.

³ As Theresa Biberauer (p.c.) has pointed out, the expletive-like element *nó* is not grammatical in Vietnamese in all contexts in which a non-null-subject language like English requires an expletive. For instance, *nó* is ruled out in all contexts in which English needs the anticipatory *it* expletive. Here we have an indication that, despite their similarities, the behaviour of *nó* and that of expletives in non-null-subject languages is not the same. Notice, moreover, that the presence of the expletive in Vietnamese does not correlate with any definiteness effect, since sentence (6c) can involve both definite and indefinite interpretations of the semantic subject “the boy”, unlike in English (where the presence of the expletive leads to an indefinite reading of the semantic subject).

- b. Nó không có cái bút nào (Vietnamese)
 EXPL NEG exist CLF pen any
 ‘There are no pens.’
- c. Nó ngã thằng bé (Vietnamese)
 EXPL fall CLF little
 ‘A/the boy fell.’

Expletives in non-null-subject languages exhibit three properties which, we suggest, can help us establish whether or not the elements appearing in the sentence-initial position in null-subject languages (4-5) are true expletives. A true expletive can be defined as a semantically empty dummy element which is part of the structure because it has to satisfy a formal requirement (related to the need to have a subject in Spec,TP, *i.e.* an EPP feature, see Roberts and Holmberg 2010 among others) which is not met by any other element of the sentence. The three properties are: i) the obligatoriness of the expletive in a series of syntactic contexts; ii) the lack of any marked pragmatic interpretation connected to the presence of the expletive; iii) the need to have the expletive both in the sentence-initial position and after the finite verb. We will show that the expletive-like elements exhibited by null-subject languages do not pattern with the expletives of non-null-subject languages as far as these properties are concerned, which indicates that, according to the definition given above, they are not expletives.

Let us focus on the first property of expletives in non-null-subject languages, *i.e.* their obligatoriness in a set of syntactic contexts (for instance with weather verbs and in existential constructions). The examples in (6) show that in non-null-subject languages the expletive cannot be dropped in these contexts (6a-d), whereas the expletive-like element in non-null-subject languages can be dropped (see Silva-Villar 1998, Toribio 2000, Hinzelin and Kaiser 2007 on Dominican Spanish).

- (6) a. *(Es) hat gehagelt (German)
 EXPL has hailed
 ‘It hailed.’
- b. *(Es) stehen viele Leute im Bus (German)
 EXPL stand.3PL many people on-the bus
 ‘Many people are standing on the bus.’
- c. *(It) rained yesterday
- d. *(There) are many books on the shelf /
- e. Nó mưa bây giờ đấy (Vietnamese)
 EXPL rain now PRT
 ‘It is about to rain now.’
- f. Nó không có cái bút nào (Vietnamese)
 EXPL NEG exist CLF pen any
 ‘There are no pens.’
- g. (Ello) vienen haitianos aquí (Dominican Spanish, 2013a : 44)
 EXPL come.3PL Haitianos here
 ‘People from Haiti come here.’

The data in (6) indicate that the expletive-like element in null-subject languages is optional, which is clearly in contrast with the definition of an expletive as a semantically empty dummy element which is part of the structure because it has to satisfy a formal requirement that is not otherwise met.

The second property of expletives in non-null-subject languages is that their presence is not connected to a marked pragmatic interpretation. This means that the presence of the expletive in (6) does not add anything to the pragmatico-semantic meaning of the sentence, and is needed for purely formal reasons. In null-subject languages, in contrast, the presence of an expletive-like element is

typically connected to a marked pragmatic interpretation of the sentence.⁴ Specifically, the presence of an expletive-like element in null-subject Romance varieties leads to marked discourse-related interpretation. In the examples in (7), for instance, the optional expletive-like element *ele* found in Portuguese dialects, which co-occurs with the syntactic subject *os lobos* and *eu*, has a pragmatic/discourse function (connected with deixis and contextuality according to Silva-Villar 1998: 267, and with the sentence's illocutionary force according to Carrilho 2008).

- (7) a. *Ele os lobos andan com fame* (European Portuguese, Camacho 2013a: 49)
 EXPL the wolves go with hunger
 'Wolves are hungry.'
- b. *Ele eu gosto de socorrer as pessoas!* (European Portuguese, Carrilho 2008: 1)
 EXPL I like.1SG to help the people
 'I like to help people!'

The fact that in the sentences in (7) the presence of the expletive-like element has a pragmatic effect on the sentence's interpretation indicates that these elements are very likely to be hosted in the clausal left periphery, where such interpretations are typically encoded (see Rizzi 1997b, Benincà 2001), rather than in the syntactic subject position (IP or TP), unlike expletives in non-pro-drop languages.

The fact that the expletives-like elements in non-null-subject languages discussed in the above examples differ from those in null-subject languages is further confirmed by the fact that non-null-subject languages can also exhibit optional, pronominal expletive-like elements with a discourse/pragmatic import alongside syntactic expletives (Koenenman and Neeleman 2001, Richards and Biberauer 2005, Haegeman and van de Velde 2006, Haegeman 2008, Biberauer and van der Wal 2014). In (8, from Haegeman and van de Velde 2006: 9f) we illustrate this with data from the non-null-subject language West Flemish, which exhibits both canonical syntactic expletives and the expletive-like element *tet*, which is morphologically identical to the third person singular neuter pronoun. The examples show that *tet* is optional and that it has the same distribution as regular expletives, since it must immediately follow the complementiser and precede the syntactic subject in embedded clauses (8a), and immediately follow the finite verb in main clauses (8b,c).

- (8) a. *da (tet) Marie dienen boek a kent* (West Flemish)
 that TET Marie that book already knows
 'than Marie knows that book already.'
- b. *Dienen boek kent (tet) Marie a* (West Flemish)
 that book knows TET Marie already
 'That book Marie knows already.'
- c. *Kent (tet) Marie dienen boek a?* (West Flemish)
 knows TET Marie that book already
 'Does Marie know that book already?'

Haegeman and van de Velde (2006) claim that the expletive-like element *tet* has neither a syntactic nor a descriptive meaning, but contributes to the sentence's expressive meaning, since it indicates that "the content of the proposition with which it occurs contrasts in some respect with assumptions in the background context" (Haegeman and van de Velde (2006: 9). The authors show that this expletive-like element is not hosted in the canonical subject position Spec,TP, but in a higher position between IP and CP connected to SubjP (Cardinaletti 2004 and Rizzi 2006), which they suggest should be divided into different FPs.

4 Greco, Trang and Haegeman *this volume* review some cases from the literature in which the optional expletive-like element has been claimed to have no pragmatic import. This does not, however, demonstrate that the expletive-like element is a true expletive, since it is still optional, unlike expletives in non-null-subject languages.

The data discussed above indicate that non-null-subject languages exhibit true expletives with a syntactic function and appear in the canonical subject position Spec,TP, and expletive-like elements with a pragmatic/discourse function are hosted in a different position.⁵ In null-subject languages, only expletive-like elements appear to be available.

The third property of expletives in non-null-subject languages is that they are obligatory both pre- (see the examples in 6 above) and post-verbally (9a-c) (see Svenonius 2002, Biberauer 2010 for a complete typology of expletives in Germanic; for more examples, see Camacho 2013a among others).⁶ In null-subject languages, on the other hand, expletive-like elements can never appear in the inversion position (in Spec,TP) (see also Silva-Villar 1998: 254).

- (9) a. *Gestern hat *(es) gehagelt* (German)
 yesterday has EXPL hailed
 'It hailed yesterday.'
 b. *Has *(it) hailed?*
 c. *Are *(there) many books on the shelf?*
 d. **Hay ello arroz?* (Dominican Spanish, Hinzelin and Kaiser 2006: 24)
 has EXPL rice
 'Is there rice?'

Camacho (2013a) (see also Silva-Villar (1998) challenges the generalisation that Romance expletive-like elements are not found in the inversion position (Spec,TP), discussing data from Dominican Spanish which seem to indicate that expletives can appear in Spec,TP in this language. The data supporting Camacho's claim are given in (10). Here it can be seen that in Dominican Spanish the expletive *ello* appearing in the sentence-initial position co-occurs with an NP subject following the finite verb (10a). Unlike in the examples in (7), however, the expletive and the NP subject cannot both precede the finite verb (10b).

- (10) a. *Ello llegan guaguas hasta allá* (Dominican Spanish, Camacho 2013a: 46)
 EXPL arrive.3PL buses until there
 'Buses arrive up there.'
 b. **Ello unas personas llegaron ayer* (Dominican Spanish, Camacho 2013a: 45)
 EXPL some people arrived.3PL yesterday

Camacho (2013a: 45f,65) takes the examples in (10) to indicate that in the null-subject language Dominican Spanish the expletive and the overt subject compete for the same position, Spec,TP/IP; if such is the case, null-subject languages cannot be said to lack true expletives.

However, the general consensus – contrary to Camacho's (2013a) position – is that the elements in (10) are not true expletives (Carrilho 2008, Hinzelin and Kaiser 2006, Frascarelli and Jiménez-Fernández 2016, Sheehan *to appear-b* among others). First, in Dominican Spanish the expletive *ello* is always optional: even in impersonal constructions, or with weather verbs (see Hinzelin and Kaiser 2006) *ello* is not obligatory. This strongly contrasts with the behaviour of expletives in non-null-subject languages, and is incompatible with the idea that *ello* is hosted in Spec,IP/TP. Second,

⁵ In V2 languages, CP expletives can also appear without a pragmatic function to satisfy the purely formal requirements of the V2 rule (the need to have an XP in Spec,CP). It is relevant to the current discussion that these CP expletives are obligatory in structures in which no constituent is fronted. This indicates that they behave like true expletives, i.e. like semantically empty dummy elements which are part of the structure because they have to satisfy a formal requirement (connected to the V2 constraint, not to the expression of the syntactic subject).

⁶ In V2 languages exhibiting CP expletives (i.e. a number of German constructions, including impersonal passives and presentational sentences - see Mohr 2005 for a complete treatment of the constructions involving CP expletives in German - and Icelandic, in all constructions, see Biberauer 2010), the expletives never appear in the inversion position, since their presence is needed to satisfy the purely formal requirements of the V2 rule in CP (to satisfy the EPP feature, or fill the Spec,TopicP position in Icelandic, see Richards and Biberauer 2005). We thank Theresa Biberauer for pointing this out to us.

Ledgeway's (2010) findings indicate that an expletive-like element and a preverbal subject can both appear in the left periphery, i.e. that FPs traditionally considered to involve the IP layer are actually found in the CP area; this has important implications for the analysis of the examples in (7), since it indicates that the expletive-like element and the preverbal subject can both appear in CP (see in this respect the two competing analysis of preverbal subjects in Italian, with Benincà and Cinque 1993 suggesting that they are hosted in a Topic, and others such as Cardinaletti 2004, Moro 1993, and Rizzi 2006, assuming that they appear in AgrSP (Moro) or in SubjP). Ledgeway's analysis for Campanian dialects, according to which both the expletive-like element and the syntactic subject appear in two SubjPs in the Topic area, can also be applied to data like those in (7), in which the expletive *ello* co-occurs with a preverbal NP subject. According to this alternative analysis, both *ello* and the DP subject appear in the left periphery, and no overt material occupies Spec,TP.

The data in (10) might indicate that in Dominican Spanish *ello* and the syntactic subject compete for the same position *within* the left periphery (possibly in the Topic field), not within IP. Subjects are default topics (see Rizzi 2006 among others) and it is plausible to assume that in this variety topical subjects always have to move to the left periphery, whereas non-topical ones cannot do so. When no subject has raised to the left periphery, *ello* can appear in the left periphery to signal the absence of a topic, like expletives in Icelandic (Richards and Biberauer 2005, Biberauer 2010).⁹ The question of why only a single position is available for the expletive-like element and the DP subject in Dominican Spanish, whereas two are available in other Romance languages, needs to be investigated in further research.¹⁰

The above discussion, which has relied mostly on null-subject Romance and non-null-subject Germanic languages, has shown that only the latter exhibit true expletives. Expletive-like elements of null-subject languages are i) optional and ii) hosted in CP: this implies that their presence typically leads to marked pragmatic readings of the sentence, and that they do not appear in the inversion position. The expletives of non-null-subject languages, in contrast, i) are obligatory in all cases in which a subject is not raised to Spec,TP; ii) are not connected to any marked pragmatic/discourse interpretation, and iii) appear both pre- and post-verbally.

1.2.2 On the identification of null subjects

Null subjects (like empty categories in general) need to be identified, and in the original formulation of the pro-drop parameter (see 3 above) verbal morphology was assumed to play a crucial role in the identification of the null category, which was considered to be a phonologically null referential pronoun (*pro*) in Rizzi's original formulation of the parameter (see Roberts and Holmberg 2010 for

⁹ We thank Theresa Biberauer for suggesting this possible analysis.

¹⁰ The status of the expletive-like element across Romance varieties needs to be better understood. In the dialects studied by Ledgeway, the expletive-like element agrees in number, person and gender with the DP subject (and has thus been analysed, on the basis of prosodic evidence, as a referential demonstrative by Frascarelli, see footnote 7), whereas in the Spanish varieties discussed in this section the element *ello* does not bear any number, person or gender agreement with the DP subject (which makes it more similar to a non-referential element). This is a fundamental difference between the Romance varieties under consideration, and may indicate that the analysis suggested by Frascarelli (footnote 8) for constructions involving a demonstrative pronoun and a DP subject are not applicable to Spanish varieties, specifically because *ello* might be (given its non-referential status) morphologically too reduced to introduce an Aboutness-Shift topic. In this respect, an observation by Carrilho (2008:9) is particularly important: she states that in Portuguese varieties the pragmatic import of the expletive-like element *ele* can be paraphrased with emphatic expressions such as “it is true that, indeed, etc”. This seems to indicate that the scope of this expletive is the whole sentence, rather than a single constituent. Moreover, as shown by Jiménez-Fernández and Miyagawa (2014), Spanish does not use a G-topic for given information, but realises it through a strategy of feature-inheritance with *in situ* destressing (we thank Mara Frascarelli for pointing this out to us). This implies that both the prosodic realisation of the syntax-discourse interface across these Romance varieties (Italo-Romance varieties vs Spanish varieties), and its relationship with expletive-like elements (use of what looks like a referential demonstrative vs use of what looks like a non-referential element) vary. Further research is thus necessary to construct a theory that covers both instances. Here the crucial point is that neither *ello* in Spanish varieties, nor the demonstrative-like element of Italo-Romance varieties, can be considered a true expletive of the type found in non-null-subject languages.

a historical overview of the correlation between morphology and null subjects). Rich agreement on the finite verb is thus assumed to allow speakers to reconstruct the person and number features of the missing subject (Taraldsen 1978, Rizzi 1982, 1986, Jaeggli and Safir 1989, Roberts and Holmberg 2010). Agreeing morphology is also seen as central to the licensing of the null subject in many analyses of null-subject languages, which consider it to substitute the subject, possibly because it contains incorporated subject pronouns (Borer 1986, Alexiadou and Anagnostopoulou 1998, Barbosa 2005, 2009b, Koenenman 2006, Frascarelli 2007, Rohrbacher 1999, Müller 2006, and Sigurðsson 2011a, Neeleman and Szendroi 2005, 2007 for the idea that rich morphology involves incorporated subject pronouns). The morphological richness of null-subject languages can satisfy the structural requirement that all sentences have a subject, the so-called EPP property, i.e. the requirement that all sentences have a syntactic subject (see Chomsky 2001 and Holmberg and Platzack 1995, Bidese and Tomaselli *this volume* for the idea that EPP can be associated with both I° or C°, and Öztürk 2008 for the idea that EPP is a property of v°).

However, null subjects are not restricted to languages with rich agreement: they are also possible in languages with poor or no agreement, such as radical null-subject languages (Huang 1984), and many languages with rich morphology are not null subject languages (like Icelandic). To accommodate the data from radical null-subject languages, it has been proposed that null subjects can be identified through discourse, as well as verbal morphology. Specifically, null subjects can be identified through a topic antecedent and be interpreted by feature inheritance.

Since Holmberg (2005), who pointed out that the notion of *pro* cannot be maintained in current syntactic theory (since the introduction of interpretable and uninterpretable features in Chomsky 1995), most current approaches to null subjects propose a combination of the two identification mechanisms discussed above. Cole (2009, 2010), Frascarelli (2007, *this volume*), and Sigurðsson (2011a), among others, propose that both agreement and the presence of an antecedent in the context play a role in the identification of null subjects.

According to Cole (2009, 2010), identification of thematic null subjects takes place through Agree and context identification. Morphologically, the identification of null thematic subjects requires that the maximal agreement required by a specific language to be present on the finite verb. This so-called *Minimal Morphological Threshold* varies from language to language. Therefore, languages do not need the same number of features to be overtly realised by agreement in order to identify null subjects: in some person, number and gender have to be realised, in others only person (see Camacho 2013a, Müller 2006, Kramer 2015 for a formalisation of features and feature ranking within the Distributed Morphology approach, and Baker 2008, Fernández-Salgueiro 2008 for the idea that agreement mechanisms be parameterised). In Cole's system, null thematic subjects also need an antecedent in the discourse to be identified. In radical null-subject languages, in fact, null subjects can be identified without the presence of overt morphology (these Cole calls "contextually strong languages without agreement"); conversely, the presence of rich morphology in a language like Icelandic does not co-occur with null thematic subject (this Cole calls "contextually weak language").

A combination of discourse and formal properties (among which morpho-syntax and intonation play crucial roles) is also central to Frascarelli's (2007, *this volume*) and Sigurðsson's (2011a) approaches (see also Grimshaw and Samek-Lodovici 1998 on Italian).

Using a corpus of Italian spontaneous speech, Frascarelli (2007) shows that context is crucial not only for the identification, but also for the licensing of null thematic subjects in Italian. Frascarelli assumes that in Italian the null category is a null referential subject *pro* (see Rizzi 1982, 1986a, Cardinaletti 2004 and below) appearing in Spec,IP. She suggests that third person *pro* is licensed through an Agree relation between an Aboutness-shift topic (in the sense of Frascarelli and Hinterhölzl 2007) hosted in CP and the null subject in Spec,IP agreeing in person and number with the finite verb (receiving ϕ -features from it).¹¹ Frascarelli (2007) proposes that the Aboutness-shift

¹¹ According to Sigurðsson (2011a) the fact that Italian has rich agreement morphology (i.e. that agreement contains an affixal pronoun in his analysis) implies that *pro* does not depend on antecedents for ϕ -features, but still needs to be C/Edge-linked. In non-null-subject languages, however, null subjects can only be identified via a C/Edge-linker in the

topic merged in the left periphery is endowed with a [+aboutness] edge feature, which she proposes should be considered as an “extended EPP feature”. Identification of the null category takes place in Frascarelli's system through a Topic Criterion (from Frascarelli *this volume*: 2):¹²

(12) Topic Criterion

- a) The high Topic field in the C-domain contains a position in which the [+aboutness] feature (an “extended EPP feature”) is encoded and matched (via Agree) by the local (3rd person) null subject.
- b) When continuous, the [+aboutness] Topic can be null (i.e., silent).

Frascarelli's (2007) approach constitutes a significant advance, since it demonstrates that even in Italian, a prototypical consistent null-subject language, in which agreement has long been considered to be key to the licensing and identification of thematic null subjects, context overrides morphology for both the identification and the licensing of null subjects (this is in line with Cole's 2009, 2010 conclusions). Therefore, Italian third person null referential subjects involve the presence of a null Topic, just as null arguments in radical null-subject languages do, along with the presence of a silent category (null referential pronoun) in Spec,IP which agrees with the finite verb. According to Frascarelli's model, a topic can only remain silent when it is part of a Topic-chain, the first topic of which must be overt, while its copies in the subsequent sentences can remain silent. Therefore, sentences with a null subject in Italian involve the presence of a silent copy of the first topic in the chain, which is maintained in the following sentences. This need for a Topic-chain in order to have a null subject distinguishes Italian from radical null-subject languages, which allow topics to be dropped if they can be inferred from their context. Moreover, Frascarelli (2007) opens up a novel perspective on the nature of the EPP feature and its relation with null subjects, since she assumes that the EPP is no longer associated with I°/T° as in the original formulation of the pro-drop parameter, but should be considered topic-dependent (her approach goes in the same direction of scholars who consider preverbal subjects to be inherently topical in nature: see the discussion in section 1.2.1 above, and also Holmberg and Nikkanen 2002 on Finnish as a topic-prominent language and the relation between topics, *pro* and null subjects).

In her chapter in this volume, Frascarelli broadens the scope of her 2007 theory, in two ways. First, she extends her analysis to first and second person null subjects in Italian, providing empirical evidence that, unlike third person null subjects, they are not licensed and interpreted through an A-chain with an Aboutness-Shift Topic, but through an Agree relation with silent logophoric operators hosted in CP (as proposed by Sigurðsson 2011a). This empirical finding, based on a corpus of spontaneous speech, confirms Frascarelli's approach to null subjects in Italian (i.e. that an edge feature is needed in CP to license and identify *pro* in Italian, and that this edge feature can either be a topic, or a logophoric operator), and the theoretical claim that the mechanisms for the licensing and the identification of thematic null subjects are sensitive to person (see also Cole 2010, Rosenkvist *this volume*, Weiß and Volodina *this volume* and below).

The second new area in which Frascarelli tests her 2007 proposal is that of cross-linguistic comparison. By relying on novel data from both Italian and Finnish, she shows that the mechanism she assumes for the consistent null-subject language Italian is replicated in a very similar (though not identical) fashion in Finnish, which is usually analysed as a partial null-subject language (see

C-domain: this implies that the only strategy to realise null subjects in non-null subject languages is through topic drop.

12 As correctly pointed out by an anonymous reviewer, there are contexts in which a null subject is licensed in Italian and the presence of an Aboutness Shift topic is doubtful. The first involves an arbitrary null subject, like *Bussano alla porta*, “(they) knock.3PL at the door” (see Jaeggli 1986), in which the null category can either refer to some contextually given (which is compatible with the presence of a silent Aboutness Shift Topic) or to a contextually new (which is incompatible with the presence of a silent Aboutness Shift Topic) set of individuals. The second problematic case is represented by cases in which the null subject is interpreted as variable bound to a negative quantifier, such as ‘nobody’, in sentences like *Nessuno_k ti ha detto che pro_k sarebbe venuto*, “Nobody told you that they would come”, where the negative QP has no referential import (and cannot be considered a Topic) and can though function as an antecedent of a null subject.

Holmberg 2005, Holmberg and Nikkanen 2002, and section 1.2.4 below). This indicates that the interaction between identification and licensing through context (topic chain with an Aboutness topic, or logophoric operators) and morphology (*pro* in Spec,IP/TP agreeing with the finite verb) is a property of all null-subject languages.

Rosenkvist's (2015, *this volume*) Distinct agreement hypothesis offers a somewhat different account of the licensing and the identification of first and second person null subjects to that given by Sigurðsson (2011a) and Frascarelli (2007, *this volume*). Rosenkvist proposes that the person split in the distribution of null subjects in several Germanic varieties in which first and second, but not third, persons are null (see also Weiß and Volodina *this volume* for German dialects and Old High German, and Madariaga *this volume* for a similar, though not identical, person split in present-day Russian) can be derived by combining Cole's (2009, 2010) approach with Frascarelli's Topic Criterion. He suggests that, as long as the ϕ -features that are expressed by the agreeing verb form match the ϕ -features that are expressed by the subject pronoun perfectly, morphology is always able to identify and license a null subject. When this happens, verb agreement can be considered to be distinct (13 from Rosenkvist *this volume*: 23).

- (13) Verb agreement is distinct iff:
- a. a specific verb form (F_a) and a pronoun (P) express the same set of ϕ -features
 - b. F and P have the same values for ϕ
 - c. only F_a matches the values for P.

Rosenkvist (*this volume*) shows that according to the definition in (13), only first and second person verbal agreement morphology can be considered distinct in the Germanic varieties he has investigated, since only in these persons is there perfect match between the ϕ -features expressed by the subject pronoun and those expressed by verbal morphology (person, number) realised either on the finite verb or in C° (complementiser agreement, see Weiß and Volodina *this volume*). Since first and second person morphology is distinct, null subjects can be licensed and identified by verbal morphology. In the third persons, conversely, verbal agreement cannot be considered distinct, since the finite verb expresses person and number, while the subject pronouns express person, number and gender (see Kramer 2015 for an analysis of gender and its implications for null subjects). Therefore, morphology is not distinctive for the identification and the licensing of third person null subjects, and null referential subjects are thus impossible in Germanic. Rosenkvist's hypothesis has the advantage of deriving the split between the first and second person, on the one hand, and the third, on the other, from the distribution of null subjects in Germanic. He does this by referring to different degrees of verbal morphological richness, thus eliminating logophoric operators from the system.

According to Rosenkvist, this hypothesis can also be applied to Romance. This would mean that first and second person null subjects rely on a purely morphological mechanism (Distinct Agreement), whereas third person null-subjects are licensed and identified by an Aboutness-Shift topic (Frascarelli's 2007 Topic Criterion: a strategy not available in Germanic varieties due to their V2 character, see also Sigurðsson 2011a).

Rosenkvist's account potentially paves the way towards novel perspectives in the future analysis of referential null subjects in consistent null-subject languages, such as the Romance languages. If Distinct Agreement is shown to be a relevant mechanism for the licensing and identification of null referential subjects in Romance, we can expect it to play a role not only in first and second persons (as in Germanic), but also in those third persons in which gender is expressed on verbal agreement along with person and number, i.e. in the Italian compound tenses which involve a past participle and the auxiliary "to be" or an object clitic. Distinct Agreement also predicts the presence of an asymmetry in the mechanisms of licensing and identifying null referential subjects between languages allowing (Italian) and disallowing (Spanish) for agreement on the past participle.

The validity of Distinct Agreement in Romance may be undermined by the distribution of null referential subjects in Italian subjunctive forms (Cardinaletti 1997). In Italian, singular person subjunctive forms are specified for number, but not for person. According to Rosenkvist's analysis, this implies that verbal agreement is not distinct, and null subjects should therefore only be possible in the third person, where null subjects are licensed through context and information structure rather than through morphology. This prediction is only partially borne out. Third person null subject pronouns are actually possible, and in the second person singular an overt subject pronoun must be realised, as expected within Rosenkvist's Distinct Agreement hypothesis. However, in the first person singular a null subject is also possible, which possibly indicates the presence of a logophoric operator. A logophoric operator also needs to be assumed in the first (*nosotros/nosotras*, “we”) and second (*vosotros/vosotras*, “you”) persons plural in Spanish, where the subject pronouns are distinct for gender (the forms ending in *-os* are masculine, those ending in *-as* are feminine), but the verbal endings are not.

Rosenkvist's Distinct Agreement hypothesis also has important implications for Older Germanic languages. As shown by Axel (2007 relying on Eggenberger's 1961 data) and Schlachter (2012) for Old High German, and by Walkden (2014) for Old English, null referential subjects were possible in these language stages, and they occurred more frequently in the third person. This split, which Schlachter (2012) and Walkden (2014) account for by assuming that null subjects are licensed contextually by a null Aboutness-Shift topic similarly to Frascarelli (2007) indicates that the topic criterion was productive in Older Germanic varieties, unlike in present-day varieties. Why this is so, how the observed person split emerged in Germanic and how this can be accounted for within the Distinct Agreement hypothesis are questions that further research will have to answer.

The last mechanism for the identification of null subjects in finite clauses that should be mentioned is control (see also below). Holmberg (2010b) and Holmberg and Sheehan (2010) argue that null subjects can be derived in two ways. In consistent null-subject languages, null subjects involve the incorporation of a subject pronoun in TP and the null subject is a deleted copy of the chain headed by T (see above). In partial null-subject languages, conversely, deleted pronouns are identified by the fact that they are controlled by the subject of the higher clause (see also Madariaga *this volume* on present-day Russian and Frascarelli *this volume* for the idea that control involves the presence of a topic chain in Finnish). Two examples of controlled null subjects in partial null-subject languages are given in (14 adapted from Holmberg and Sheehan 2010: 131).

- (14) a. *A Maria_i admite que ela_j/pro_j não fala muito bem inglês* (Brazilian Portuguese)
 the Mary admits that she/pro NEG speak.3SG very well English
 b. *Maria_i myöntää ettei hän_j/pro_j puhu englantia hyvin* (Finnish)
 Mary admits that-NEG-3SG she/pro speak.PRS English well
 ‘Mary admits that she does not speak English very well.’

As this section has shown, morphology is still central to research on null-subject languages, as assumed in the original formulation of the null-subject parameter. New contributions (including those contained in this volume) to the debate on the relationship between null subjects and morphology have, however, refined our conception of this connection in two ways: firstly, by providing a more nuanced reading of how rich/poor morphology may be defined, and, secondly, through analyses of the interplay between discourse and morphology.

1.2.3 What is the null category?

In the previous section we discussed a series of approaches to identifying of the null subject. We have not yet, however, actually defined the null category. There are three main approaches within Generative Grammar to this question.

The first is the classic Government and Binding approach (Chomsky 1982a, Rizzi 1986a, which treats *pro* together with other types of empty nominal categories (NP-trace, wh-trace, PRO) as an

instance of an empty nominal with no inherent properties apart from (presumably) nominal categorial features and maximal X-bar level: $[_{NP} e]$, whose features are supplied by Agr/Infl or by a clitic. As discussed in Holmberg (2005), this analysis of the null category involved in null-subject phenomena turned out to be incompatible with the recent development of syntactic theory, in particular with the Minimalist Programme (Chomsky 1995, 2001 and subsequent work) which assumes the ϕ -features on T° to be uninterpretable. This implies that the null category *pro* cannot be analysed as an empty nominal $[_{NP} e]$ unspecified for ϕ -features, because this analysis would not explain how the null category is interpreted given that Agr/Inf is not a supplier for ϕ -features. The alternative proposed by Holmberg (2005) is thus assuming that the null category is a pronoun with interpretable features, i.e. it is specified for ϕ -features, it occupies the Spec,TP position and functions as a weak overt pronoun, which can be thus deleted at PF (see also Perlmutter 1971, Roberts 2010a for this analysis). The fact that deleted pronouns are silent thus becomes a PF matter in this model: *pro* is assumed to enter the derivation as a ϕ -P, and is fully specified for ϕ -features. The ϕ -features set is deleted at PF by the same deletion operation that affects copies in general.¹³

A second approach assumes that there is no null category at all in null-subject languages, i.e. the ϕ -features on Agr/I are interpretable and phonologically expressed as a verbal affix with the same features as a referential, definite pronoun (see Alexiadou and Anagnostopoulou 1998, Borer 1986, Barbosa 2005, 2013). Therefore, there is no need to postulate an empty category because the ϕ -features are supplied by Agr/I (see Sheehan *to appear-b* and Holmberg 2005).

A third approach to null-subject phenomena, originally put forth for radical null-subject languages lacking agreement (Huang 1994, Oku 1998, Kim 1999, Tomioka 2003, Saab 2009, Saito 2004; 2007, Takahashi 2006; 2008) and recently applied also to consistent (Greek: Giannakidou and Merchant 1997; Spanish: Duguine 2008, 2014) null-subject languages, analyses null-subject constructions as cases of argument ellipsis.¹⁴ The ellipsis approach links the availability of null subjects to the independent properties of nouns in the various languages which favour argument ellipsis. For Japanese, a radical null-subject language, Tomioka (2003) shows that bare NPs allow for a wide range of semantic interpretations (indefinite, definite, singular or plural) which depend on discourse factors, and he proposes the following generalisation (Tomioka 2003: 336).

(15) All languages which allow discourse *pro*-drop allow (robust) bare NP arguments.

The above generalisation has been shown to be valid for Japanese, Korean, Mandarin Chinese, Thai, Hindi, Turkish (see Camacho 2013a: 33) and also for Brazilian Portuguese (Barbosa 2013). It is captured through the assumption that the null category in these null-subject languages is the phonologically null version of a bare NP deleted through ellipsis (see Modesto 2008 for the idea that null subjects in Brazilian Portuguese, Finnish and Chinese involve topic drop).

By starting from the developments of recent theories on the nature of pronouns which assume the presence of a phonologically null NP as a complement of D in every pronoun (Elbourne 2005), Barbosa (2013) claims that the idea that the null category in null-subject languages is a minimally specified bare NP, $[_{NP} e]$ should not be abandoned from the theory of grammar. More specifically, she provides arguments in favour of the fact that *pro* can be reduced to the same $[_{NP} e]$ in both radical and partial null-subject languages. Therefore, null subjects (and null objects) in these languages are analysed in Barbosa's system as cases of null NP anaphora, and the differences in the interpretation of null subjects are assumed to follow from the resources available in the languages for the application of the semantic operator of type shifting (see Chierchia 1998). This analysis cannot be applied to consistent null-subject languages, which, according to Barbosa (2013), do not exhibit any null underlying category, i.e. the ϕ -features on Agr/I are interpretable and phonologically expressed as a verbal affix with the same features as a referential, definite pronoun (see Alexiadou and Anagnostopoulou 1998, Borer 1986, Barbosa 2005).

¹³ We thank an anonymous reviewer for pointing this out to us.

¹⁴ Topic drop in Germanic can also be analysed as a type of argument drop, see Sigurðsson (2011a) and van Gelderen (2013).

All three approaches to the nature of null categories are represented in the volume.

In their chapters, Biberauer, Bidese and Tomaselli, Frascarelli, Madariaga and Sheehan assume that the null category is *pro* which they analyse along the lines of the recent developments of Generative syntax as in Holmberg (2005), Biberauer, Holmberg, Roberts and Sheehan (2010). Rosenkvist's and Weiß and Volodina's analyses for German synchronic and diachronic varieties respectively set themselves within the second approach to the null-subject phenomenon, since they assume that the distribution of null subjects can be reduced to the properties of agreement (see section 1.2.2).

Ruda (*this volume*) approach to the null category in null-subject phenomena sets itself within approaches which consider the null category to be a minimally specified nominal (Barbosa 2013; Panagiotidis 2003), namely the categorizing head *n*. Starting from the observation that a language with robust bare NPs like English lacks null-subjects – in contrast to Tomioka's (2003) generalisation – Ruda (*this volume*) revisits arguments from the literature (Hoekstra and Roberts 1993) which show that English has some limited cases of null arguments: in middle constructions (16a,b), with the impersonal *one* (18c,d) and in arbitrary missing-object sentences (16e,f), see the following examples in Ruda (*this volume*: 6, borrowed from Hoekstra and Roberts 1993: 187–188):

- (16) Middle formation
- a. Bureaucrats always bribe easily.
 - b. ?*Bureaucrats bribe easily; even John managed it.
- Impersonal *one*
- c. One always goes out on Sundays.
 - d. *One telephoned last night; it was John.
- Arbitrary missing object
- e. John steals for a living.
 - f. *This evening John stole.

Ruda also provides novel arguments which demonstrate that object drop is possible in English, not only in marked (like in recipes, see Bender 1999; Cote 1996; Culy 1996; Haegeman 1987a,b; Massam and Roberge 1989; Ruppenhofer 2005; Ruppenhofer and Michaelis 2010), but also in unmarked stylistic contexts. According to Ruda, null objects are typically heavily lexically-constrained (see Cote 1996; Fillmore 1986), although natural data suggest that object drop can be acceptable, when a null-object sentence containing a verb which typically requires an overt definite object is appropriately contextualised.

She proposes that null arguments in English follow from the properties of NPs in the language, i.e. from the fact that nouns can be phonologically unrealised because of their internal structure. More specifically, null arguments are not phonologically null nouns, but are analysed as [_{NP} n], i.e. as nouns without the nominalising head *n* (see Ruda *this volume* for arguments for this, and Acquaviva 2009; Embick and Marantz 2008; Embick and Noyer 2007; Panagiotidis 2011, Kramer 2015). Within Ruda's approach, D-less argumental NPs do not need the merge of D to receive referential interpretation, and can be interpreted via a type-shifting operation at the syntax-semantics interface (see Chierchia 1998 for iota ι , the operation of shifting a predicate to an individual, see also Partee and Rooth 1983, Partee 1987 on type shifting in semantics). Post-syntactic type-shifting operations can, therefore, according to Ruda, be a further mechanism for the identification of the null category.

Sheehan's and Egerland's chapters deal with the status of the null categories involved in non-finite constructions. Focusing on particular instances of infinitival clauses (inflected or cased) in Icelandic, Russian and European Portuguese, Sheehan argues that the null category featuring in these infinitival constructions is *pro*, rather than PRO, and that it is obligatory controlled. The notion of “control” has been mostly applied to describe the restrictions on the interpretation of null subjects in embedded finite (Borer 1986, Holmberg, Nikanne and Sheehan 2009, among others, see

example (15) above) and, to a lesser extent, in finite clauses (Madariaga *this volume* for Russian and Rodrigues (2007) for Finnish and Brazilian Portuguese) in partial null-subject languages.¹⁵

Sheehan suggests that this notion can also be applied to embedded non-finite constructions which contain a *cased* obligatorily null subject which is *not* referential but rather behaves like PRO in being obligatorily controlled. In the languages considered, two types of obligatory control are present, which Sheehan calls caseless control and cased control and which involve two different derivations. Sheehan claims that caseless control involves movement (see Hornstein's 1999 *et seq.* analysis of control as movement), while cased control involves Agree between an argument in the higher clause and *pro* in the non-finite embedded clause. Sheehan's system of obligatory control does not include PRO (see also Hornstein 1999, Boeckx and Hornstein 2004, Martin 2001, Pires 2006 for the idea that PRO should be eliminated from the theory): the only available null category is *pro*.

Egerland (*this volume*) approaches null categories in non-finite clauses from a different perspective, examining control and the status of the null category in gerundial constructions in Italian and English.¹⁶ His focus is on gerundial constructions in which the null category refers back to the whole preceding clause, rather than to a single constituent, as shown in the following examples (17 from Egerland *this volume*: 1).

(17) The crops were destroyed by drought, leaving the population in despair.

Egerland shows that the null category that appears in the non-finite constructions above is not *pro*; his analysis builds upon Iatridou and Embick's (1997) proposal that *pro* cannot have a clause as its antecedent because clauses do not have the ϕ -feature specification necessary for the identification of *pro*, and for the existence of *pro* as a category (see Rizzi 1986a, Cardinaletti 1997, Holmberg 2005 among others). His proposal is that the null category is PRO, i.e. a null pronoun lacking ϕ -features, which is licensed through obligatory control by the matrix event. According to Egerland's theory, the properties of the gerundial constructions he considers should be universal, i.e. they should also be valid in those partial null-subject languages for which, on the basis of the data on infinitival and non-gerundial constructions, the presence of *pro* as a null category is assumed. Moreover, clausal control in gerundial constructions is also possible in non-null subject languages like English, which makes the hypothesis that PRO is involved more plausible, since English lacks *pro* (with some exceptions, see Ruda, *this volume*). Considerable further research is needed to ascertain whether or not Egerland's theory is universally valid.

To sum up, null subjects have been variously categorised in the literature: as silent, referential pronouns (*pro*), elliptical gaps (topic drop), or empty nominals [_{NP} *e*] unspecified for ϕ -features. *Pro* subjects are licensed and identified through morphology and discourse, or through control. On the other hand, bare NPs deleted through ellipsis can be identified through context, or through post-syntactic type-shifting operations.

1.2.4 On partial null-subject languages

As briefly mentioned in section 1.2 above, recent work on null-subject languages has pointed to the existence of a third type of null-subject systems beside consistent and radical null-subject languages: partial null-subject languages. These languages allow for a silent, referential, definite subject in a restricted number of syntactic contexts or under certain conditions which are not always

¹⁵ Madariaga's (*this volume*) and Rodrigues's (2007) conclusion that *pro* can be controlled in main finite clauses is at odds with Holmberg, Nayudu and Sheehan's (2009) results based on Finnish, Marathi and Brazilian Portuguese and raises the issue of determining the extent to which Russian really differs from them. We thank an anonymous reviewer for pointing this out to us.

¹⁶ On gerundial constructions in Italian, and more generally in Romance, see Lonzi (1991) and Casalicchio (2013, 2014, 2016b). There is abundant literature on English gerunds, see e.g. Abney (1987), Felser (1999), Pires (2006) on different types of gerundial clause.

easy to define (see Holmberg 2005 on this). In this subsection, we summarise the most common restrictions on the distribution of null-subject in partial null-subject languages and discuss how the observed facts fit into the typology of null-subject languages discussed in (3).

One typical restriction on the distribution of null-subjects in partial null-subject languages involves person. In some partial pro-drop languages, referential subjects tend to be null only in the first and second person and not in the third person (as in Hebrew, Shlonsky 2009; Finnish, Holmberg 2005, Holmberg and Sheehan 2010, but see Frascarelli *this volume* for evidence that third person null subjects are also possible in Finnish; Germanic varieties, see Rosenkvist *this volume*, present-day Russian, see Madariaga *this volume* and Bizzarri 2015 for evidence that also third person null subjects are possible in present-day Russian). In others, however, null subjects are only possible, or are more frequent, in the third person (as in Shipibo, Camacho and Elías-Ulloa 2010, Old English, Walkden 2014 and Old High German, Schlachter 2012 where null subjects are considerably more frequent in the third persons). A third type of restriction on person is found in Marathi, where null subjects are only possible in the second person (Holmberg and Nayudu and Sheehan 2009; see also Bavarian dialects, Fuß 2005, Weiß and Volodina *this volume* and Cognola 2014 on Mòcheno for the same effect). The second restriction found in the distribution of null subjects in partial null-subject languages involves an asymmetry between main and embedded clauses. In Old Romance languages (see Benincà 1984, Cognola 2015 on Old Italian, Adams 1987b, Vance 1989, Roberts 2010a, Zimmermann 2014, *this volume* on Old French, Wolfe 2015 on Old Spanish) and in Old Germanic languages (see Axel 2007, Eggenberger 1961, Schlachter 2012, Weiß and Volodina *this volume*, Cognola 2015 for Old High German; Walkden 2014, van Gelderen 2013, Rusten 2015 on Old English; Håkansson 2008 on Old Swedish and Kinn 2016a on Old Norwegian¹⁷) null subjects appear almost exclusively in main clauses, and are rare in embedded clauses. In other partial null-subject languages, null subjects are highly restricted in embedded clauses because they can only be licensed if controlled (see discussion in section 1.2.2 above and Sheehan *this volume*, Madariaga *this volume*).

Another typical property of partial null-subject languages discussed in the literature concerns the interpretation of generics. As pointed out by Holmberg (2005) and Holmberg and Sheehan (2010) among others, in partial null-subject languages a third person null subject (regardless of whether it is singular or plural) can have an arbitrary, generic interpretation. In particular, it can have an inclusive interpretation similar to that of English *one*. This is the crucial property that distinguishes partial null-subject languages from consistent null-subject languages. In the latter, a plain third person subject can never have a generic interpretation that is meant to include the speaker. A 3rd singular null subject cannot have a generic interpretation unless it is overtly marked with an impersonal pronoun, like Italian *si*. A third person plural null subject can have an arbitrary interpretation, but it has an exclusive interpretation to the effect that it must exclude the speaker. As discussed in Barbosa (2013), there is a split with respect to number morphology within partial null-subject languages. While in Brazilian Portuguese and Finnish, the generic null subject is singular, in Russian (Madariaga (*this volume*: 3f) and Hebrew it is plural (i.e. it triggers plural verbal morphology). But regardless of number morphology, what is crucial is that the inclusive interpretation is available in these languages as opposed to the consistent non-null-subject languages.

The properties assumed to characterise null-subject languages are not always found as cluster properties, and this is typically shown by the distribution of *that*-trace violations. We know that *that*-trace violations are found in some non-null subject languages, like the Scandinavian (see Maling and Zaenen 1978, Lohndal 2009) and Southern German varieties (Haider 1983, 1993, Grewendorf 1988, Featherston 2005, Kiziak 2010, Bayer and Salzmann 2013) and Övdalian

17 See Kinn, Rusten and Walkden (2016) on Old Icelandic, which falls outside the generalisation as it does not exhibit any main-embedded clause asymmetry in the distribution of null subjects. However, this may not be directly linked to the null-subject phenomenon, since present-day Icelandic does not have the kind of main-embedded asymmetries typical of Germanic V2 languages either, and is therefore defined a symmetric V2 language (see Holmberg 2015). We thank Theresa Biberauer for pointing this out to us.

(Rosenkvist 2010). The property's relevance for the null-subject phenomenon has therefore been questioned (see Camacho 2013a).

Gilligan (1987) and Nicolis (2008), among others, have proposed a way to include *that*-trace violations in the cluster of properties ascribed to null-subject languages. They suggest that *that*-trace violations be considered part of a “mini-cluster”, together with expletive null-subjects, but not with free subject inversion, as was implied in the original formulation of the null-subject parameter according to which free inversion is a prerequisite for the absence of *that*-trace effects (see Chomsky 1981, Rizzi 1982, Burzio 1986, Bidese and Tomaselli *this volume*, and Gallego 2013, Ordoñez 1998, Belletti 2004, Sheehan 2010 Mensching and Weingart 2016 for some new proposals). Using data from some Creole languages, Nicolis (2008) shows that the lack of *that* trace effects in these languages correlates with the availability of null expletive *pro*, and not with the presence of free-subject inversion. Therefore, all languages allowing *that*-trace effects are expected to allow expletive null-subjects; this has been confirmed for Southern German varieties which have null expletives in impersonal passives and lack *that*-trace effects (Featherston 2005, Kiziak 2010, Bayer and Salzmann 2013, Cognola 2013).

Roberts and Holmberg (2010: 22) drawing on Gilligan (1987) and Nicolis' (2008) work suggest widening the correlation and propose the implication scale in (20), which states that if a language has free subject inversion, it will also allow for *that*-trace violations and expletive null subjects (appearing in Spec,TP, see the discussion in section 1.2.1 on CP and TP expletives).

(18) free subject inversion → allow *that*-trace violations → expletive null subjects.

As discussed by Roberts and Holmberg (2010: 22), this correlation holds in a variety of languages, although not, it seems, in some Spanish (see Camacho 2013a: 51f) and Rhaeto-Romance varieties (Casalicchio 2016a). Cimbrian (Bidese and Tomaselli *this volume*) exhibits free-subject inversion, does not have *that*-trace effects one overt expletive subject (*da*) which, according to Bidese and Tomaselli, is connected to the realisation of the syntactic subject (it is a morpho-syntactic expletive subject enclitic to Fin/C which absorbs nominative case).¹⁸ A similar argument to that made for expletives in Romance can also be made for Cimbrian, where the status of the expletive element *da* is far from clear. Cognola and Hinterhölzl (2016) provide evidence that Cimbrian *da* has a pragmatic/discourse function. More specifically, the function of *da* is to anchor an utterance to a context when there is no referential subject (the element which typically performs this function) in the left periphery. According to Cognola and Hinterhölzl' s (2016) analysis, therefore, *da* should be analysed as a CP discourse expletive (like the expletives in Romance languages).

To sum up, although the cluster character of the properties distinguishing null-subject languages has been challenged, we suggest that no argument strong enough to invalidate the theory has yet been presented. More specifically, it is indisputable that some null-subject languages, Italian for example, exhibit all the properties of the null-subject cluster. For those languages which do not fully pattern with consistent null-subject languages, the identification of “mini-clusters” of properties and of implications between them, might be an effective means towards balancing descriptive and explicative adequacy – thereby allowing us to further our understanding of the null-subject phenomenon.

1.2.5. How can the properties of null-subject languages be formalised?

The discussion has shown that the null-subject phenomenon is multifaceted, since the cluster of properties identified for Italian (a consistent null-subject language) in the 1980s (Rizzi 1982) is not

¹⁸ Cimbrian also has the expletive element *z* which behaves like a CP expletive and whose distribution is ruled by the constraints of the V2 rule. This expletive, like CP expletives in V2 languages in general, should not therefore be considered in the implicational scale in (18).

found in all languages exhibiting null-subject phenomena. Duguine and Madariaga (2015: 1), in fact, speak of the “chaotic character” of null-subject phenomena.

What does this mean for our understanding of syntax? How can we capture these empirical manifestations of null-subjects?

By starting from the developments of recent theories on the nature of pronouns which assume the presence of a phonologically null NP as a complement of D in every pronoun (Elbourne 2005), Barbosa (2013) claims that the idea that the null category in null-subject languages is a minimally specified bare NP, [_{NP} *e*] should not be abandoned from the theory of grammar. More specifically, she provides arguments in favour of the fact that *pro* can be reduced to the same [_{NP} *e*] in both radical and partial null-subject languages. Therefore, null subjects (and null objects) in these languages are analysed in Barbosa's system as cases of null NP anaphora, and the differences in the interpretation of null subjects are assumed to follow from the resources available in the languages for the application of the semantic operator of type shifting (see Chierchia 1998). This analysis cannot be applied to consistent null-subject languages, which, according to Barbosa (2013, do not exhibit any null underlying category, i.e. the ϕ -features on Agr/I are interpretable and phonologically expressed as a verbal affix with the same features as a referential, definite pronoun (see Alexiadou and Anagnostopoulou 1998, Borer 1986, Barbosa 2005).

Biberauer (*this volume*) provides a new answer to this question within current parametric theory (see Biberauer and Holmberg and Roberts and Sheehan 2010 for a previous formalisation of null-subject languages, in particular Holmberg 2010 for the idea that different parameters are involved in the null-subject phenomenon and Biberauer *this volume* for the shortcomings of this first formalisation). Biberauer (*this volume*) shows that previous attempts to account for null-subjects in terms of parameters fail to capture the complexity behind the phenomenon, in which not only the properties of the licensing category (T, C etc) and its relation with morphology (as assumed in the first formalisation of null subjects in terms of parameters, see above), but also the properties of the pronouns themselves, the conditions on ellipsis and its relation with DP structure, the position of the EPP feature (associated with I° or C°, see Bidese and Tomaselli *this volume*), the nature of control, and discourse-oriented factors like Aboutness (see above discussion) play a role. Biberauer proposes that this empirical complexity can only be captured if we assume that the properties connected to the null-subject phenomenon do not follow from a single parameter (and thus from a single acquisition hierarchy, according to the parametric theory proposed by Biberauer and Roberts 2012b, 2015 a,b, 2016, Biberauer, Holmberg, Roberts and Sheehan 2014) but from different acquisitional hierarchies and, therefore, parameters (see also Duguine 2014 and Duguine and Madariaga 2015 for the idea that variation in null subjects does not result from a single parameter, but rather from the interaction of the properties of a language's grammar). Biberauer (*this volume*) proposes that the single properties connected to the null-subject phenomenon be parameterised and that language-specific systems result from the interaction between the single hierarchies involved in this phenomenon (see also Frascarelli *this volume* for a possible parametrisation of her Topic Criterion).

The approach to null subjects suggested by Biberauer (*this volume*) opens up new perspectives on the description, analysis and formalisation of the null-subject phenomenon both synchronically and diachronically, since it provides a model of how parameter resettings in various superficially unrelated areas of grammar known to play a role in the null-subject phenomenon may bring about variation and change in the null-subject domain (see Biberauer and Roberts 2016 for parametric change in diachrony). This model could, for example, potentially allow us to account for cases of diachronic changes in the realisation of subjects: from partial null-subject languages to consistent null-subject languages (as in the case of Old Italian to present-day Italian, see Benincà 1984, Cognola 2016), from partial null-subject languages to non-null-subject languages (as in the case of Old High German and Old English to present-day Standard German and English, see Axel 2007, Walkden 2014 among others); from non-null-subject languages to partial null-subject languages (as in the case of Cimbrian, Bidese and Tomaselli *this volume*), or from consistent null-subject languages to partial null-subject languages (as in the case of Russian, see Madariaga *this volume*) by

ascertaining which parameter in the hierarchies has been reset and how its resetting affects other parameters involved in the phenomenon.

1.3 The null-subject phenomenon: conclusions and outlook

After this review of the properties of null-subject languages we can go back to Newmeyer (2004), who considers the null-subject parameter as illustrative of the Principle and Parameter approach's failure to account for typological variation.

We argue that far from being an example of failure, the history of the null-subject parameter is a clear example of successful linguistic research. The pro-drop parameter in its original formulation inspired an extremely wide ranging comparative enterprise which involved an enormous number of languages, as linguists looked for the properties that had been identified for consistent null-subject languages and compared their findings. One result of this comparative work highlighted in Newmeyer (2004) was the discovery that the pro-drop parameter in its original formulation could not account for all the attested empirical patterns, and that other factors (like discourse and ellipsis) had to be taken into account. The other result, which has received less attention, was that a finite number of mechanisms seem, from a cross-linguistic point of view, to play a role in the licensing and interpretation of null subjects (morphology, discourse, control etc). This indicates that despite the differences between null-subject languages from different families, most of their properties are shared.

It was only possible to reach this result because syntacticians had a powerful theory, GB, and a parameter within which a series of empirical facts could be united and accounted for. The parameter was revealed to be less powerful than expected, and many facts proved to be more complex than expected, but this result was only obtained because this strong theoretical hypothesis led so many researchers working on so many different languages in the same direction. The pro-drop parameter is, we believe, a clear example of the fact that an incomplete or wrong theory is better than no theory, demonstrating, as it does, the energising effect of a strong theory, which inspires researcher to seek out novel data in order to confirm or refute it.

The pro-drop parameter in its original formulation can be “viewed as a “first-pass” attempt at capturing the morphosyntactic regularities that consistently shape the grammars of human languages” (Biberauer *this volume*: 1). In its original formulation the parameter may have failed,

but it has revealed the depths of complexity underlying null subjects. Our introduction has attempted to summarise this complexity, subject, too, of the present volume; and necessarily subject of much future research, in order to be fully understood.

1.4 The individual contributions

The book is organized two parts: the first focuses on the nature of pro-drop and the parameters that define it. The second is a collection of studies on the different types of null category, and the mechanisms through which they can be identified.

The first part opens with a paper by **Ciro Greco, Trang Phan and Liliane Haegeman**, who discuss the presence of the expletive pronoun *nó* in spoken Vietnamese, a radical pro-drop language. Some scholars (see Camacho 2013a and the discussion in section 1.2.1 above) have divided optional expletive subjects in pro-drop languages into two types (but see the introduction to this volume): TP-expletives, which are held to merge in Spec,TP (like Dominican Spanish *ello*), and CP-expletives, which have discourse related properties (like Balearic Catalan *ell*). However, the expletive pronoun *nó* challenges this bipartition because it shares some properties with both types of expletive: it has no referential content and contributes to the interpretation of the clause, like CP-expletives. On the other hand, it is incompatible with preverbal subjects. The authors account for this fact by positing that *nó* is merged in Spec,TP and has a [+specific] feature, as the analysis of different types of structure, such asthetic and existential sentences, shows.

The role of expletives is also dealt with by **Ermenegildo Bidese** and **Alessandra Tomaselli**. Their contribution focuses on Cimbrian, a little studied German(ic) heritage language spoken in Northern Italy. Unlike most Germanic varieties, Cimbrian has lost the linear V2-restriction, although it maintains V-to-C movement and has three different expletive pronouns, which are used in different contexts. Moreover, although it is a non-null-subject language Cimbrian has some correlates of pro-drop: free subject inversion and the absence of *that*-trace effects. These observations lead the authors to argue that V2 is incompatible with (consistent) pro-drop, and that the loss of V-to-C movement is a necessary precondition for the development of pro-drop.

Old French is another language whose status in relation to the null subject parameter is unclear. In his chapter, **Michael Zimmermann** offers a diachronic analysis based on a large, newly established, corpus of Medieval texts. He considers both referential and TP-expletive subjects: the high frequency of the former and the existence of the latter lead the author to posit that Old French was neither a consistent nor a partial null-subject language. To support this claim, he shows that all the properties typically ascribed to null-subject languages are absent in Old French. As a consequence, Old French should be considered a non-null-subject language. The fact that the instances of pro-drop found in the texts are all explained as instances of left-peripheral focalisation, which exceptionally allows the drop of the postverbal subject pronoun, is consistent with this hypothesis. Significantly, Modern French also allows subject drop in these contexts.

The second part examines the issue of what types of null subject exist, and how they can be identified. It is divided into two sections: the first deals with null subjects in control structures, the second with the role of discourse factors and agreement in the identification of null subjects in finite clauses.

The first section, focussing on control structures, contains chapters by Michelle Sheehan, Nerea Madariaga and Verner Egerland.

In the first chapter, **Michelle Sheehan** compares the control structures of Russian, Icelandic and European Portuguese. These three languages share the availability of Partial Control, besides Obligatory Control, in infinitival clauses. This property correlates with the possibility of the null subject of the embedded non-finite clause carrying an independent case. Based on this observation, the author proposes that there are two syntactic types of obligatory control: cased control, where the null subject is a pro bearing an Agree relation with the controller, and caseless control. In the latter, the null subject has the same case as its controller: the author derives this property from the movement of the controller, which is merged in the embedded subject position and then moves to a projection of the main clause. The languages under investigation thus show that the two most important minimalist approaches to control are required: cased control is best explained through Landau's (2000 et seq.) approach, caseless control by the MTC.

The contribution by **Nerea Madariaga** also deals with Russian: she analyses the diachronic evolution of the Russian pronominal system in embedded clauses. This analysis is particularly interesting because Old Russian was a consistent null-subject language, while present-day Russian has become a partial null-subject language. Madariaga shows that infinitival structures in Old Russian did not display any syntactic control; however nowadays both finite and non-finite embedded clauses have obligatory control. Basing her analysis on the Movement Theory of Control (MTC, Hornstein 1999 et seq.), Madariaga suggests that this evolution is due to a reanalysis: while embedded infinitives used to be able to have a pro subject, learners gradually started to parse the gap not as a pro, but as an NP-trace/copy. This change must be seen in the broader context of the development of Russian from a consistent to a partial null-subject language, and is directly linked to the development of overt weak pronouns.

While the early chapters of this section argued for an MTC analysis, at least in some control contexts, **Verner Egerland** discusses a structure that he argues is incompatible with a movement theory: gerunds with clausal control. This is a type of construction in which the null subject of the gerund is coindexed with the whole preceding clause. The author shows that the MTC is unable to account for this prediction; nor can the analysis of the null subject as pro, because pro can never be coindexed with a clause. An alternative possible explanation, the analysis of clausal gerunds as

reduced relative clauses, is also discarded - both on the basis of cross-linguistic comparison and of data from English. Hence, the author concludes that in these constructions the null subject must be PRO, and that a binding approach to Control is needed: an Event Phrase merged in the matrix vP controls the null subject PRO of the clausal gerund.

The second section of the second part is devoted to analyses of the mechanisms that allow the identification of the null subject. The two most important factors discussed in these chapters are related to discourse and morphology.

Mara Frascarelli compares the identification of pro in Italian, a consistent null-subject language, and in Finnish, a partial null-subject language. The point of departure is her Topic Criterion (Frascarelli 2007), which states that all third person null subjects agree with a [+aboutness] feature in the high Topic field. The author analyses a series of new data on the interpretation of null subjects in different structural contexts, testing the account of the Finnish partial pro-drop proposed in Holmberg, Nayudu and Sheehan (2009). According to these authors, two factors play a role in licensing pro in Finnish: locality and control. Frascarelli's analysis confirms that the locality condition (combined with semantic factors) does indeed play an important role. However, control appears to be a weaker requirement. Frascarelli also restates the difference between null-subject language and non-null-subject language through a macroparameter (the Topic Criterion): all types of null-subject language can have a null Topic-chain, while null-subject s cannot. On the other hand, the difference between consistent and partial pro-drop languages is due to the Interface Visibility Criterion (a mesoparameter), which is present only in partial null-subject languages and which requires that at least one link of the Topic chain be visible at the interface level.

Marta Ruda discusses some instances of subject and object drop in English, Polish and Kashubian. She claims that empty arguments are merely a projection of the nominal categorising head *n* lacking a lexical root insertion. This analysis is a reformulation of Panagiotidis' (2003) and Barbosa's (2013), among others, proposals that null arguments be considered as NPs projected by bare null nouns. Null nouns are possible when a verbal functional head, such as *v* or *T*, has nominal features (Chomsky 1995), thus allowing the interpretation of *nP* as a referential argument. This analysis is reinforced by a comparison of Polish and Kashubian, which are closely related genetically. Although both languages have rich agreement and display all the properties which usually correlate with the pro-drop parameter, Polish is a null-subject language and Kashubian a non-null-subject language. According to the author, this shows that null and overt subjects have a similar syntactic composition, the only difference being in pronunciation.

Next, **Helmut Weiß** and **Anna Volodina** deal with the phenomenon of referential null subjects from Old High German (OHG) to present-day dialects. OHG differs in two respects from modern German dialects: OHG was consistent pro-drop, while the modern dialects are partial pro-drop (only some persons can be pro); in OHG pro-drop is mainly attested in main clauses (partial asymmetry), while it is found in both main and embedded clauses in current dialects. The question they try to answer concerns the intermediate stage: what type of pro-drop do we find in Early New High German (ENHG)? Analysing a small corpus, the authors show that ENHG was like modern dialects. Furthermore, ENHG arguably developed inflected complementisers, which are also present in modern dialects and which allow pro-drop through "Agr-in-C" licensing. This type of licensing is able to account for all instances of pro-drop in the various stages of Germanic, except for embedded pro-drop in OHG, which is considered a "Indo-European relic" that subsequently disappeared.

An alternative approach to pro-drop in German(ic) is offered by **Henrik Rosenkvist**, who analyses a number of dialects in which pro-drop is possible with some persons. He shows that the persons involved are always the first or the second, never the third. Unlike Weiß and Volodina, Rosenkvist proposes an approach that is based on purely morphological requirements: pro-drop is possible when there is Distinct Agreement. Distinct Agreement exists for those persons where the ϕ -features of the verb ending and the correspondent overt pronoun match, i.e. the two elements must have the same features, with the same value. The author suggests that gender is also a relevant feature in Distinct Agreement, which explains why pro cannot usually express the third person in

Germanic. In consistent null-subject languages, on the other hand, when Distinct Agreement fails it is still possible to resort to clause-external contextual identification, as in Frascarelli's (2007 and this volume) proposal. This means that syntactic accounts referring to inflected complementisers or to a Speech Act Participant can be dispensed with.

The contributions to this volume discuss many aspects relevant to the complex and multifaceted null-subject phenomenon. The languages considered vary from major languages, like Russian and Italian, to lesser used languages like Germanic vernaculars and Kashubian. Older language stages, such as Old French, Old Russian and Old High German, are also analysed. Moreover, apart from the main topics, around which the book is structured, the chapters also discuss other arguments connected to the complex theme of the null subject phenomenon: the diachronic development of null-subject languages, the status of expletives (see also section 1.2.1 of this chapter) and the correlates of the null-subject phenomenon (see section 1.2.4 of this chapter).

As we have seen, since Rizzi (1982) the null subject has been one of the most analysed and debated phenomena in Generative Grammar. This volume constitutes an up-to-date contribution to this rich and continuously evolving debate, and the variety of topics discussed draws a detailed picture of the multiple facets to this phenomenon and of the latest approaches to it. Nevertheless, we are aware that this contribution is by no means the last word on the question of null subjects and that much work has still to be done, and thus we hope that this volume will offer an insightful stimulus to future research.