THREE PARTIAL NULL-SUBJECT LANGUAGES: A COMPARISON OF BRAZILIAN PORTUGUESE, FINNISH AND MARATHI

Anders Holmberg, Aarti Nayudu & Michelle Sheehan

Introduction[*]

This paper reports an investigation of a set of languages which fall under the general rubric of partial null-subject languages, that is languages which allow null subjects but under more restricted conditions than consistent null-subject languages. For the languages considered here, the conditions include

(a)   when the subject is a generic pronoun corresponding to English 'one' (exemplified by (1a), from Marathi), and

(b)   when the subject is controlled by an argument in a higher clause (exemplified by (1b), also from Marathi).

(1)   a.   unahlyat   lavkar utthavla jato   [Marathi]

summer-in early   wake   go-PRS-3SM

'In summer one wakes up early'

b.   Ram   mhanala   ki   ghar   ghetla   [Marathi]

Ram   say-PST-3SM   that   house   buy-PST-3SN

'Ram said that he bought a house'

We will argue that property (a) is particularly revealing. It shows that the languages lack a D(efinite)-feature in T. This is the feature which makes a null subject with definite

interpretation possible in consistent null-subject languages. The languages that we will focus on are Brazilian Portuguese (BP), Finnish, and Marathi, three languages which are genetically and areally distant from each other (although BP and Marathi are both Indo-European),[1] but share properties (a) and (b), and we propose an analysis of their sentential featural composition on that basis. When property (b) is looked at more closely, it turns out that the conditions under which the languages allow a controlled, externally licensed null subject vary to some extent. We will show that they nevertheless have enough in common for the control relation to be characterised as a type of relation distinct from both Obligatory Control and Non-Obligatory Control, as found with non-finite clauses, and from the discourse-based antecedence relation characteristic of consistent null-subject languages, according to Samek-Lodovici (1996) and Frascarelli (2007).

1.   Null subjects in partial pro-drop languages

1.1   Null subjects that are not licit in partial pro-drop languages

Impressionistically speaking, null subjects in partial null-subject languages are optional in some contexts where they are obligatory in consistent null-subject languages and excluded in non-null-subject languages, and excluded in some contexts where they are allowed in consistent null-subject languages. The following is an illustration: Consider (2), where John is talking about himself, as indicated by the indexing.

(2) $John_1$ said that he_1 bought a house.

In a non-null subject language such as English the pronoun has to be overt. In a consistent null-subject language, such as Arabic, Greek, Spanish, Turkish, etc., the pronoun has to be null, assuming for the sake of argument that there is no contrast involved. In our partial null-subject languages the pronoun can optionally be null, the null-option exemplified in (1b) for Marathi.

Now imagine a context where another person, call him Bill, is being discussed. One of the interlocutors utters (3) as a contribution to the discussion, where the embedded pronoun refers to Bill, as indicated by the index 2.

(3)   John$_1$ said that he$_2$ bought a house.

In a non-null subject language such as English the pronoun obviously has to be overt. In a consistent null-subject language the pronoun would still typically be null, assuming no contrast or topic-shift. In our partial null-subject languages the pronoun has to be overt in this case.

. We exemplify the difference between the three types with an embedded subject pronoun because a wider range of languages, including at least some partial null-subject languages and non-null subject languages, allow null subjects in main clauses, under certain discoursal conditions (see Haegeman 2000). Even so, the example requires a certain amount of idealization. For example, the exact syntactic role that the NP referring to Bill has in the discourse preceding (3) may affect the interpretation of a null subject in otherwise consistent null-subject languages (see Cole, forthoming). Furthermore, there appear to be languages which have some properties of partial null-subject languages, but allow a null subject even in (3), and languages with some properties of partial null-subject languages which do not allow a null subject even in (2).[2]

Considering just the three languages under investigation in this paper, there is another complication, which is that Finnish allows dropping 1st and 2nd person subjects basically in any context (although more commonly in formal and written Finnish). 3rd person pronominal subjects, though, are subject to constraints which do not apply in consistent null-subject languages, but which closely resemble those which hold for BP and Marathi. We will, from now on, deal with 3rd person subjects only; we return briefly to 1st and 2nd person pronouns in section 2. See also Vainikka & Levy (1999) and Holmberg (2005). It should also be mentioned that many of the data we are reporting from Marathi are from spoken Marathi, written Marathi being more restrictive as regards null subjects (in a sense the opposite of the situation in Finnish).

For the purposes of this paper, we will use the term 'partial null-subject language' strictly for languages that have properties (a) and (b).[3]

The following sections will go through the cases where partial null-subject languages have null subjects, either optionally or obligatorily.

1.2 Null non-thematic subject

With predicates which do not have a theta-marked subject the partial pro-drop languages generally have no overt subject. One such case is weather-predicates.[4]

(4) a. Está   chovendo.   [BP]

         Is   raining

         'It's raining.'

     b. Ulkona sataa.   [Finnish]

outside rains

'It's raining outside.'

1.3   Null generic subject pronoun

One situation where a null subject is licit in finite clauses in our three partial null-subject languages is when the subject is a generic pronoun corresponding to 'one'.

(5)   a.   É   assim   que   faz   o   doce.   [BP]

is   thus   that   makes   the   sweet

'This is how one makes the dessert.'

   b.   Nesse hotel   não   pode   entrar na   piscina   bêbado

in-this   hotel   neg   can   enter   in-the   swimming-pool drunk

'In this hotel it is not permitted to go in the swimming pool drunk.'[5]

[BP, Rodrigues 2004:72]

(6)   a.   Kesällä   herää   aikaisin.   [Finnish]

in-summer wake-PRS.3S   early

'In the summer one wakes up early.'

   b.   Täällä ei   saa   uida.

here   not-3S may-PRS swim

'One must not swim here.'

   c.   Nuorten mielipiteitä kuuluu   arvostaa.

youth's   opinions   should-PRS-3S   respect

'One should respect the views of young people.'

5

(7) a. unahlyat   lavkar utthavla jato   [Marathi]

         summer-in early wake   go-PRS-3SM

         'In summer one wakes up early'

     b. asa   lokan kade   baut dakhavayla   paidze nahi

         like this people towards finger show-INF- DAT should NEG.AUX

         'One should not point at people'.

     c. mulan-chya vicharan-cha aadar   kar-ay-la   paidze

         children's   views   respect do-INF- DAT should

         'One should respect the views of young people'.

Interestingly, in this case consistent pro-drop languages such as Spanish and Greek, do not

allow a plain null subject, but have to resort to some overt strategy. The contrast is seen most

clearly when comparing BP, a partial pro-drop language, with European Portuguese (EP), a

consistent pro-drop language. Compare (5) and (8):

(8) a. É assim que   **se** faz   o doce.   [EP]

         is thus   that   SE makes thesweet

         'This is how one makes the dessert.'

     c. Nesse hotel   não **se** pode entrar na   piscina   bêbado. [EP]

         In-this   hotel   neg SE can   enter in-the swimming-pool drunk

         'In this hotel it is permitted to go in the swimming pool drunk.'

In EP, the generic subject reading requires merging the reflexive clitic *se*. This clitic is either itself the generic pronoun, or serves to somehow license a null generic pronoun. In either case it holds that a plain 3SG null subject with no special morphology is not an option in EP or the other Romance null subject languages, in this case.

It is important to make a distinction in this connection between *generic* and *arbitrary* null subject pronoun. By *generic pronoun* we mean a pronoun best translated into English as either 'one' or 'you', the semantic defining characteristic being that it denotes people in general including the speaker and the addressee. By *arbitrary* we mean a pronoun which is best translated into English as *they*, as in *They speak many different languages in India*, the semantic defining characteristic being that it denotes people in general (in some domain), but excluding the speaker and the addressee. Consistent null-subject languages have an arbitrary null subject (null 'they') but to express a generic subject pronoun, they resort to some overt strategy. Among partial null-subject languages some have a null arbitrary pronoun as well as a generic one (in active clauses without any special morphology), others do not; for instance BP does, but Finnish does not.

Other consistent pro-drop languages use other strategies, including making use of 'generic you' (2nd singular), which will be null, but visible on the verb agreement (9), or a mediopassive form (10), or an overt indefinite pronoun (11).

(9) a. xsse-k   texdem   htta l-xamsa-w-settin   [Moroccan Arabic]

         need-you work-2S until the-five-and-sixty

         'One has to work until the age of 65.'

     b. Den mporeis na empistefteis kanenan   [Greek]

         not can-2S to trust-2S   anyone

'You can't trust anyone.'

(10) La y-usbah-u   hunaa.   [Standard Arabic]

      not 3-swim.PASS-PL here

      'One can't swim here.'

(11) Ewaru-aynaa ii kurcii loo sukham gaa kuurcoo waccu.   [Telugu]

      who   even this chair in comfortably sit   may

      'One/anyone can sit comfortably in this chair.'

None of them employ a null 3rd person generic subject in construction with an active, 3SG-marked verb. There is an obvious functional explanation for this: In a consistent null-subject language a null 3rd person subject will be interpretable as a definite pronoun. Remove *se* from (8a), for example, and the sentence can be read as 'This is how he makes the dessert.' In BP this does not happen, as the language does not have definite null subjects. The reason why consistent pro-drop languages resort to overt strategies to express a generic null subject would thus be to avoid ambiguity. There is a grain of truth in this explanation (but no more than a grain), and we will return to it in section 4.

1.4   Null subject controlled by an antecedent in a higher clause

Another situation where BP, Finnish, and Marathi all allow a null subject is when there is a linguistic antecedent in a higher clause.

(12)a.   O João$_1$ disse   que   (ele$_1$) tinha   comprado uma casa   [BP]

           the João   said   that   he   have-PST.3SG bought   a   house

           'João said that he had bought a house.'

    b.   Os meninos$_1$ ficavam contentes quando (eles$_1$) tinham   um día de folga

           the children were   happy   when   they have-PST.3PL a day of holiday

           'The children were happy when they had the day off.'

    c.   A   Maria$_1$ admite que (ela$_1$) não fala   muito bem inglês.

           The Maria admits that she not speak-PRS.3SG very   well English

           'Mary admits that she doesn't speak English very well.'

(13)a.   Juhani$_1$ kertoi että (hän$_1$) oli   ostanut talon.   [Finnish]

           Juhani   said   that he have-PST.3SG bought house

           'Juhani said that he had bought a house.'

    b.   Lapset$_1$ olivat mielissään kun (he$_1$) saivat   vapaapäivän.

           children were pleased   when they get-PST.3PL off-day

           'The children were pleased when they got a day off.'

    c.   Marja$_1$ myöntää ettei   (hän$_1$) puhu   englantia hyvin.

           Mary admits   that-not-3SG she   speak-PRS English well

           'Mary admits that she doesn't speak English well.'

(14) a.   Ram$_1$ mhanala   ki   (tyani$_1$) ghar ghetla   [Marathi]

           Ram   say-PST.3SM that he   house buy-PST.3SN

           'Ram said that he bought a house'.

    b.   mulan-la$_1$ khushi dzali   dzewha (tyan-la$_1$) shalyat-hun

children-DAT happy happen-PST.3SF when   they-DAT school-from

radza milali

off   get-PST-SF

'The children were happy when they got a off from school'.

c.   Seema₁ kabul karte   ki   (ti₁)   chukli

Seema   agree do-PRS.SF   that   she   mistake-PST.3SF

'Seema admits that she made a mistake'.

This will be discussed in more detail below in section 5.

1.5.   A spurious null subject: Replies to Yes/No questions

Many languages readily drop the subject pronoun in answers to yes/no-questions. In fact, in some languages an affirmative answer to a yes/no-question often consists of just the finite verb. Such is the case in Finnish, as well as in BP and Marathi.

(15) – Pesikö   Jari autonsa?   [Finnish]

        washed-Q Jari car-his

      'Did Jari wash his car?'

   – Pesi.

     washed

     'Yes.'

As shown by Holmberg (2001), for Finnish, this is not a null-subject construction in the usual sense of a construction with a null subject in an otherwise complete finite sentence, but instead is derived by ellipsis of a larger constituent (roughly equivalent to TP) which includes the subject. Somewhat simplifying the analysis in Holmberg (2001), the construction is derived by moving the finite verb to the C-domain, the movement triggered by a polarity focus feature, and then deleting the TP containing the subject as well as the object. A strong indication that this is not a subject pro-drop construction is the fact that it is insensitive to the person of the subject. As mentioned, Finnish allows null 1st and 2nd person subjects freely, but 3rd person only under quite restricted conditions. In replies to yes/no-questions any subject can be null. Another piece of evidence is that pronouncing the object but not the subject in a reply is actually ungrammatical. (16) is not a well formed reply to the question in (15).

(16) – *Pesi auton/sen.

washed car/it

The same holds true of Marathi and BP:

(17) – Tarani   pustak vaacheli?   [Marathi]

Tara-ERG book   read-Q

'Did Tara read the book?'

a.   – ho, vaache.

yes read

'Yes.'

b.   – ho, tini tila vaache.

11

yes she it   read

c.   *– ho, tila vaache.

yes it   read

(18) – Você viu   o   fogo?   [BP]

you   saw the fire

'Did you see the fire?'

a.   – Vi.

saw

'Yes.'

b.   – Sim eu vi   ele.

yes I   saw it

c.   ??– Vi ele.[6]

saw it

This follows if the bare verb reply in (17a) and (18a) is derived by movement of the verb out of

vP (or possibly, as in Finnish, the TP), with deletion of the vP (or possibly the TP) containing

both the subject and the object. (17b) and (18b) are simply ordinary full declarative sentences

(preceded by the affirmative particle in (17b)). (17c) and (18c) are ordinary declaratives with a

null subject, which, as discussed, is not generally allowed by the rules of Marathi or BP.

Answers to yes/no-questions are thus irrelevant to the issue of how to derive and license

null subjects.

2.   The derivation of null subjects

The theory expounded here combines the theory of null subjects in Holmberg (2005) with the theory of incorporation in Roberts (2007) and Frascarelli's (2007) theory of the interpretation of null subjects. According to Holmberg (2005), one of the parameters involved in regulating the pronunciation of subject pronouns is whether finite T does or does not host an inherent, valued feature [D], encoding definiteness. In consistent null-subject languages T hosts a D-feature, in partial null-subject languages and non- null-subject languages it does not. This is, in effect, an updated version of Rizzi's (1982) formulation of the null-subject parameter as a matter of having or not having a feature [+referential] in INFL. We will modify this theory as follows: Instead of a valued D(efiniteness)-feature, T has an *unvalued* D-feature, in consistent null-subject languages, which is valued either by the subject, that is if the subject is a DP marked for (in-)definiteness, or by a null topic in specCP (following Frascarelli 2007).[7] The latter is the case when the subject is a null ϕP. More precisely this works as follows:

Pronouns differ in terms of richness of functional structure. There is wide-spread agreement that there are pronouns which are DPs ('strong pronouns') and pronouns which are not DPs ('weak' or 'deficient' pronouns); Cardinaletti & Starke (1998), Dechaine & Wiltschko (2002). More controversial is the question whether there is further differentiation. We will operate with two varieties: (1) What we will call D-pronouns: These are DPs, made up of valued ϕ-features (person, number, and in some languages, gender), a valued D-feature, and an unvalued Case-feature. (2) Defective pronouns or ϕ-pronouns, made up of valued ϕ-features, an unvalued Case-feature, and nothing else. The value of a D-feature, we assume, is a referential index. A category with an unvalued D-feature needs to copy the referential index of a valued DP, by entering a chain with it, or an anaphoric relation of some sort.

13

Frascarelli (2007) argues that 3rd person null subjects in Italian invariably refer to a person or object introduced as a topic of the discourse. She thereby concurs with Samek-Lodovici (1996) and Grimshaw and Samek-Lodovici (1998). Frascarelli develops this idea by characterising more precisely the type of topic that is involved in the interpretation of null subjects, in the framework of a theory of the left periphery articulated in Frascarelli & Hinterhölzl (2007): It is an *Aboutness-shift Topic*), henceforth abbreviated A-topic. The role the A-topic is to introduce or reintroduce a topic in the discourse, and should be kept distinct from the other types of topic in the typology proposed by Frascarelli & Hinterhölzl (2007). We illustrate this with an Italian example from Samek-Lodovici (1996) (see also Grimshaw and Samek-Lodovici (1998) and Cole (forthcoming):[8]

(19) a.   Questa mattina, la mostra   è   visitata di Gianni. Pìu   tardi *Ø/egli/lui

this morning   the exhibition was visited by Gianni. Later   he/he

ha visitato l'università.

visited   the university

'This morning the exhibition was visitied by Gianni. Later he visited the university.'

b.   Questa mattina, Gianni ha   visitato la mostra.   Pìu   tardi Ø ha visitato

this morning   Gianni visited   the exhibition. Later   visited

l'università.

the university

'This morning Gianni visited the exhibition. Later he visited the university.'

In (19a) a null pronoun is impossible in spite of the (seemingly) unambiguous antecedent in the preceding sentence. Instead an overt pronoun (either the more literary *egli* or the more

colloquial *lui*) must be used, because the pronoun introduces a new A-topic, in Frascarelli's (2007) terms, which it does because the topic of the preceding sentence is 'the exhibition'. Frascarelli argues, furthermore, (a) that an A-topic is always syntactically represented in a designated A-topic position in the articulated C-domain, either overtly (for instance in the Italian Clitic Left-Dislocation construction) or covertly, and (b) that the antecedent of a null subject is a null A-topic base-generated in the C-domain of the clause immediately containing the null subject. This null A-topic is a copy of a (possibly null) A-topic in the preceding discourse. It is thus interpreted by virtue of being in a 'topic-chain' with an A-topic in the discourse, and serves as antecedent of the null subject.[9] This is represented schematically in (20). We represent the null subject as 'ϕP' for reasons to be made clear below, where we will also characterise the index-sharing between the null A-topic and the null subject.

(20) $[_{CP}$ <Gianni$_j$> [questa mattina Gianni$_j$ ha visitato la mostra]].

$[_{CP}$ <Ø$_k$> [pìu tardi ha ϕP$_k$ visitato l'università]]

k = j

We represent the formation of the topic-chain simply as an index-identification relation k = j, without trying to characterise the conditions under which it occurs (see footnote 9). What is important for our purposes is that it is a discourse-grammar phenomenon, applying across sentences in a discourse, not subject to narrow-syntactic structural conditions such as c-command. This explains, in part, the facts exemplified by (2) and (3) above.[10]

We now propose that the index-sharing relation between the null A-topic and the null subject in the second line in (20) crucially involves T in the following way: The A-topic values the uD-feature of T, where we assume that the valuation consists of uD copying the referential

index of the A-topic. We take 'definite' to equal 'has a referential index' (note that an A-topic is always definite).

As before, we restrict the discussion to 3rd person pronouns. A defective 3rd person pronoun (which we label 'φP', even though it is not strictly speaking a phrase) on its own cannot be definite, as it has no D which could have a referential index. Restricting the discussion to subjects, in a language without uD in T such a pronoun can only be interpreted as impersonal, that is either as generic, arbitrary, or non-thematic. In a language with uD in T the defective 3rd person pronoun can be interpreted as definite if it is merged in the domain of a T whose uD-feature is valued by an A-topic, and from there is incorporated in T, in the following manner:

Adopting ideas from Roberts (2007) we take incorporation of a φP in T to be a direct effect of Agree (in the sense of Chomsky 2001). This works as follows: Finite T has a set of unvalued φ-features, and therefore probes for a category with matching valued features (step 1 in (21)). The defective subject pronoun has the required valued φ-features, and therefore values T's uφ-features, which is to say that the φ-feature values of the subject pronoun are copied by T. At the same time T values the subject's unvalued case feature (step 2 in (21)). We assume that NOM case is encoded as a valued feature of T (not an entirely uncontroversial assumption; see Holmberg (to appear)).

(21)

1      $[T, D_k, u\phi, NOM]$ $[vP [3SG, uCase] v …]$ →

2      $[T, D_k, 3SG, NOM]$ $[vP [3SG, NOM] v …]$ →

3      $[T, D_k, 3SG, NOM]$ $[vP$ ~~[3SG, NOM]~~ $v …]$

16

As a result, T shares all of ϕ's feature values. More precisely, T's feature values are a superset of ϕ's values, since T also has uD (valued 'k' by the A-topic) and a tense feature. Effectively, the result is the same as if ϕ had moved, by head-movement, incorporating into T, but without the formal problems which classical head-movement has, including lack of c-command between the links of the head-chain; see e.g. Matushansky 2006). Roberts (2007) proposes that the probe and the goal in this situation form a chain. As such it is subject to chain reduction (Nunes 2004, Bobaljik 2002). The principal rules of chain reduction are

(22) a.   Pronounce the highest chain copy.

b.   Pronounce only one chain copy.

Consequently the subject ϕP is not pronounced (indicated by the strikethrough under step 3 in (21)) As the chain includes the feature [D], by virtue of T's D-feature, and since [D] is valued by the A-topic in specCP, the result is a definite null subject construction, with the referential index of the A-topic. The chain is pronounced only in the form of an affix on the finite verb or auxiliary (following incorporation of V+v into T; see Roberts (2007, to appear) for details).[11]

As for 1st and 2nd person null subjects, Frascarelli (2007) adopts Sigurðsson's (2004) idea that every clause has features representing the speaker and the addressee in the C-domain (in a modern version of Ross's (1970) performative hypothesis). In this way, the speaker and the addressee are always available as local antecedents. We adopt the same analysis.

If the subject is a DP, either a lexical DP or a D-pronoun, T's uD-feature will be valued by the subject's D, as either definite or indefinite, as the case may be.[12] This will preclude the assignment of a referential index to T by a null A-topic in specCP. Instead, apart from existential and other thetic sentences, which do not have any A-topic, the subject DP will be the

17

A-topic of clause. We assume that the subject has priority over the A-topic in specCP as a consequence of bottom-up derivation: The subject is probed as soon as T is merged. If the subject is a DP, it will value T's uD-feature. If the subject is a bare ϕP, the valuation of Ts uD feature has to wait until the A-topic is merged in the C-domain.

A lexical DP or a D-pronoun cannot be incorporated in T as they have feature values not copied by T, as T lacks the requisite unvalued features. To begin with, a lexical DP has a root, which is not copied by T under Agree. Consequently, even though T and the lexical subject DP in specvP share ϕ-feature values through Agree, they do not form a chain, and consequently the lexical subject, being the highest member of a chain, is spelled out and pronounced (unless it undergoes movement to a higher position). It is less clear what features pronouns such as *egli* and *lui* in (20) have that prevent their incorporation in T, perhaps particularly *egli*, which is a deficient pronoun in terms of Cardinaletti & Starke's (1998) typology of pronouns. In our terms even *egli* has a valued D feature, as it can introduce a new A-topic (see (20)), but since T, by hypothesis, has an unvalued D-feature which gets valued by the subject, T and the subject pronoun will share that feature value, too, in addition to sharing ϕ-feature values.

Roberts (2007, to appear) proposes that pronouns which do not incorporate have a case feature blocking incorporation, which incorporating, deficient pronouns do not have. We do not want to adopt this hypothesis here, though, since the null generic subject pronoun in Finnish has been shown to have case, quite incontestably (see Holmberg 2005, Vainikka 1989, and especially Holmberg, to appear). Since the generic pronoun is incorporated in T (see below), case must be compatible with incorporation.

We will essentially leave the question open: We postulate a feature F, a property of the pronouns which end up spelled out in specTP, which prevent their incorporation in T, but we leave open the precise characterisation of F.[13]

18

The result is that the only pronouns that remain null are the ones that are linked to a null A-topic in specCP. The generalisations that 3rd person null subjects in consistent null-subject languages are always definite, and always refer to a person or object already introduced as an A-topic is thereby explained.[14]

An additional important assumption is that the null A-topic in specCP, when it values the uD-feature in T, also checks T's EPP-feature. We thus take issue with the view advocated by Barbosa (1995) and Alexiadou & Anagnostopoulou (1998) that the EPP in null-subject languages is checked by V-movement, or, as in Platzack (2004), that it is checked by incorporation of the subject in T. The case in point is when the subject is a bare ϕP, incorporated (by Agree) in T. In that case the subject is the chain (T,ϕP). This chain cannot itself check the EPP. This is particularly clear if the EPP specifically requires a specifier (in specTP or, in the case of the A-topic, in specCP). The non-head member of the chain, ϕP, cannot move, and the whole chain (T,ϕP) obviously cannot move to specTP. This means that whenever T's uD-feature is not valued by a null A-topic, the EPP needs to be checked by a sentence-internal constituent. There are two cases to consider: When the subject is a DP (including when it is a D-pronoun), and thetic sentences, lacking an A-topic. When the subject is a DP it is not incorporated, so it is spelled-out, and can be attracted by the EPP to specTP (i.e. it will merge a second time, with TP). In the case of thetic sentences, either some non-subject constituent or an expletive will typically merge with TP; see Sheehan (2006, to appear).

So far we have accounted for consistent null-subject languages. In other languages finite T does not have a uD-feature. In a subset of these languages, viz. the partial null-subject languages, the subject can still be null, essentially by the same derivation as in the consistent null-subject languages: T probes for ϕ-feature values. The subject's ϕ-feature values are copied by T, and the subject has its Case-feature valued in return. In the case where the subject is a

19

bare ϕP, T will copy all the feature values of the subject. As a result T and ϕP form a chain, and the subject remains null, by chain reduction. However, in the absence of uD in T, valued by an A-topic, the interpretation of the subject chain cannot be that of a definite pronoun. The remaining alternatives are non-thematic, generic, or arbitrary readings, which is what we find in partial null subject languages.

When the subject is a DP (lexical or pronominal), it cannot be incorporated, and when the subject is incorporated in the partial null-subject languages, it can only be interpreted as impersonal. Yet BP, Finnish, and Marathi have null subjects that are interpreted as definite, as shown in section 1, so there must be an alternative derivation of null subjects.

We have also said that a subject which is not incorporated is attracted by the EPP to specTP. The prediction is, then, that the definite null subject in partial null-subject languages is in specTP and checks the EPP, while the generic null subject is in specvP and does not check the EPP. This prediction can be shown to be right. Consider (23a,b):

(23) a.   Jari sanoo että tässä istuu mukavasti.   (Finnish)

           Jari says that here   sits   comfortably

           'Jari says that one can sit comfortably here.'

           ≠ 'Jari says that he sits comfortably here.'

     b.   Jari sanoo että (hän) istuu mukavasti tässä

            Jari says   that   he   sits comfortably here

            'Jari says that he sits comfortably here.'

            ≠ 'Jari says that one can sit comfortably here'.

20

Finnish has an EPP condition which is mostly satisfied by the subject, but can be satisfied by certain other categories, too, including circumstantial adverbials. In (23a) the 3SG subject has not undergone movement to specTP; instead the place adverbial has. The subject is null, by hypothesis due to incorporation by Agree with subsequent chain reduction. In the absence of uD in T, it cannot have definite interpretation.[15] In (23b) the subject has moved to specTP, satisfying the EPP. In this position the subject pronoun cannot be null by virtue of incorporation in T, but only by virtue of having a local antecedent, hence the interpretation (see Holmberg, to appear).

A similar situation is found in BP:

(24) a.   João me contou que na praia vende cachorro quente

João me told that at.the beach sell-3Sg dog hot

'João told me that hot dogs are sold at the beach '

$\neq$ 'João told me that he sells hot dogs at the beach'

b.   João$_1$ me contou que (ele$_1$) vende cachorro quente na praia

João me told that sell-3Sg dog hot at.the beach

'João told me that he sells hot dog at the beach'

$\neq$'João told me that hot dogs are sold at the beach'

[BP, Rodrigues (2004:142)]

In Marathi the prediction cannot be so easily tested since due to its SOV syntax all arguments and adjuncts precede the finite verb anyway.

It appears, then, that the definite null subjects in BP, Finnish, and Marathi are DPs which have been second-merged with specTP. See Barbosa (to appear), who reaches essentially the

same conclusion as us regarding BP (and other partial null-subject languages), and a similar, though not identical conclusion regarding consistent null-subject languages.

A corpus-investigation of the kind that Frascarelli reports in her recent works remains to be done for BP, Finnish, or Marathi. Our impression is, though, that the antecedents of their definite null subjects are A-topics, and that the null subject carries over this A-topic to the embedded clause. However, while in consistent null-subject languages licensing a null subject only requires that there be an A-topic in the immediately preceding discourse (in a higher clause or in an independent sentence) which can be, indirectly, the antecedent of the null subject, this is not sufficient in partial null-subject languages. In those languages the antecedent argument must actually c-command the null subject (barring one case which will be discussed in section 5).

If Frascarelli (2007) is right, the relation between the null subject and its antecedent in consistent null-subject languages is indirect: The antecedent is the A-topic of its clause, which means that typically it has a null copy in specCP (in the designated A-topic position). The clause with the null subject also has a null A-topic in specCP which enters a topic chain with the antecedent A-topic, which means, we assume, that the inherent referential indices of the two categories are collapsed. This referential index is then copied by the uD feature of T, and finally, through Agree, by the null subject. It is therefore irrelevant whether the antecedent argument does or does not c-command the null subject.

In partial null-subject languages, by hypothesis, the indirect relation is impossible due to absence of a uD-feature in T which could transmit the antecedent topic's index to the null subject. Why can the fronted subject not be controlled by an A-topic in specCP directly? If it could, we would not expect to see any differences between consistent and partial null-subject languages regarding the relation between a null subject and its antecedent in the linguistic or

22

situational context. Modesto (to appear) in a comparative study of BP and Finnish argues that the controlled null subject in finite clauses is itself in topic position (based on the notion that BP and Finnish always move an argument, usually but not necessarily the subject, to topic position). While we recognize the 'topic-prominent' character of BP and Finnish (see Holmberg & Nikanne (2002) on Finnish), we cannot say whether the final position of the null DP subject in BP and Finnish actually is Frascarelli/Hinterhölzl's A-topic-position.

What features does the fronted null pronoun have in partial null-subject languages? It seems reasonable to assume that a nominal argument which must enter a structurally determined relation with another argument to be interpretable must be somehow deficient; see Shlonsky (to appear), discussing control into finite clauses in Hebrew, for a similar conclusion. We suggest, therefore, that it lacks a D-feature value, and therefore must enter a structurally defined control relation with a valued DP antecedent. That is to say, either the D-pronoun comes with an inherent D-feature value (a referential index), in which case it will be spelled out/pronounced in specTP, being the head of an A-chain, or it has an unvalued D-feature, and enters a control relation with a valued DP, and remains null as a result of an extended version of chain reduction ('extended' since the control relation does not qualify as a chain in the strict sense). A third possibility, irrelevant here, is that it undergoes A-bar movement to some higher position.

That is to say, partial null-subject languages and consistent null-subject languages differ with respect to the distribution of the uD feature: In consistent null-subject languages finite T has this feature, while in partial null subject-languages pronouns may have the uD-feature. The prediction is that there may be languages which have neither: They would have null impersonal pronouns (in finite sentences) given lack of uD in T, but they would not have null, controlled

subject pronouns in finite sentences. Icelandic may be such a language: see Sigurðsson and

Egerland (to appear).

3.   **The role of agreement**

What makes a language a partial pro-drop language, as opposed to a consistent pro-drop

language? What role does the agreement paradigm play? Consider first BP: When compared

with European Portuguese (EP), BP has a reduced verb agreement paradigm with only three

forms (across all tenses and moods). It is generally taken for granted that this is a factor behind

the differences the two display with regard to pro-drop (Duarte 1995, 2000, Figueiredo Silva

2000, Modesto 2000a, 2000b, Ferreira 2004, Rodrigues 2004). The table (from Duarte 2000)

shows the evolution of the pronoun-verb paradigm in the 20th century (see also Barbosa, to

appear).

(25)

Person & number	Pronoun	Paradigm 1	Paradigm 2	Paradigm 3
1s	Eu	Amo	Amo	Am**o**
2s	Tu	Amas	*	*
2s	Você	Ama	Ama	Am**a**
3s	Ele/Ela	Ama	Ama	Am**a**
1pl	Nós	Amamos	Amamos	*
1pl	A gente	*	Ama	Am**a**
2pl	Vós	Amais	*	*
2pl	Vocês	Amam	Amam	Am**am**

3pl	Eles/Elas	Amam	Amam	Am**am**

As the table shows, the change has in fact affected primarily the system of pronouns, which has lead to a simplification of the agreement paradigm. Thus the 2S *tu* and the 2PL *vós*, each of which triggered a distinct agreement form on the verb, have been substituted by the erstwhile polite forms *você* and *vocês*, which do not trigger agreement distinct from the 3S and the 3PL respectively. And 1PL *nós*, which triggered distinct 1PL agreement, has been substituted by *a gente* (literally 'the people') triggering 3S (or a form homonymous with it) on the verb. The net effect is, however, an agreement paradigm with only three forms. Moreover, Duarte's (1995, 2000) empirical studies of the evolution of BP shows that this simplification has gone hand in hand with an increased use of overt pronouns in spoken BP.

Likewise, Marathi has syncretism between 1st and 3rd person forms, in both plural and singular, throughout the agreement system. Only the 2nd person singular is unambiguously marked.

(26) Marathi *gana* 'sing'

	SG	PL
1	gat-o (M)	gat-o
	gat-e (F)	
2	gat-os	gat-at
3	gat-o (M)	gat-at
	gat-e (F)	

25

Finnish, on the other hand, has a verbal agreement paradigm which is almost completely free from syncretism. As mentioned, Finnish allows free pro-drop in the 1st and 2nd person (particularly in formal varieties), so we might expect those to be morphologically distinct. If 1st and 2nd are unambiguously distinct, then 3rd is, too, by opposition to 1st and 2nd. The Finnish 3rd person subject agreement is a null form in some tenses and moods, morphologically marked in others, but always identifiable by opposition to the 1st and 2nd. The only syncretism is that the 3rd singular and 3rd plural have the same form in a common variety of colloquial Finnish. In other varieties, including formal and written Finnish but also some dialects, 3rd singular and 3rd plural are clearly distinct. As far as we know, this variation regarding number marking in the 3rd person has no effect on the null-subject facts summarized in section 1 (but this remains to be investigated).

(27) Finnish *laula* 'sing'

	SG	PL
1	laula-n	laula-mme (or: lauletaan)
2	laula-t	laula-tte
3	laula-a	laula-vat (or: laula-a)

Note that what is characteristic of partial null-subject languages, if we are right, is that they do not have incorporation of a subject pronoun with definite interpretation. In the theory adopted here based on Roberts (2007) it holds that a pronoun cannot be incorporated in T if it has a feature which T does not have. If, for instance, T lacks a person feature, this will rule out incorporation of a pronoun with a specified person feature. Can the difference between consistent and partial null-subject languages be understood in these terms?

26

Consider first BP: If we ignore the 1SG (*amo*), subject-verb agreement distinguishes only between singular (*ama*) and plural (*amam*). Thus the uφ-feature system of T may be characterised as a two-feature system: [u1, uNr]: Only *amo* is a realisation of [+1], all the other forms are realisations of [-1], distinguished only by number. The pronouns also distinguish between 2nd (*voce, voces*) and 3rd person (*ele/ela, eles/elas*), though, and even if *a gente* is grammatically 3SG, *eu* is obviously 1st person. Thus, apart from *e,* whose features are copied by T, all the pronouns would have a person feature which T does not have, ruling out incorporation and forcing spell-out.

In Finnish, too, there are indications of a weakening of the φ-feature system of T (in widely spoken varieties): In those varieties there is no number distinction in the 3rd person, and likewise in the 1st person, there is no featural opposition between singular and plural: The form *lauletaan* is the passive form, which is uninflected for φ-features. The pronouns, on the other hand, are all clearly marked for person and number. However, in the 2nd person a distinction is consistently made between 2SG (*laulat*) and 2PL (*laulatte*). Furthermore, it does not seem to be the case that varieties of Finnish which maintain a distinct 3PL form, or a distinct 1PL, of T would have properties of consistent null subject languages; for instance, all varieties of Finnish have a null generic pronoun, which, as we have seen, is uncharacteristic of consistent null-subject languages. Therefore it seems at least as plausible that the reduction of the verbal paradigm is a consequence, rather than a cause of the partial null-subjecthood of Finnish.

And when we turn to Marathi, even though the φ-feature system of T is somewhat reduced, it does not seem possible to reduce it to a two-feature system. The paradigm clearly has person (singling out the 2nd person), number (distinguishing 2SG and 2PL), and gender, as does the pronominal paradigm, even though there are syncretic forms of T.

27

Furthermore, the discussion of BP above presupposes that the only ϕ-features that T has are those that are distinguished morphologically. But this would predict that, for instance, Italian and Spanish would also be, at best, partial null-subject languages, disallowing incorporation of a subject pronoun , because T never makes a distinction between feminine and masculine, while the 3rd person pronouns consistently do make that distinction. Given Roberts (2007), the fact that 3rd person subject pronouns do incorporate (with definite interpretation) in Italian and Spanish means that T has gender features, even if they are not morphologically expressed.

Roberts (to appear) suggests that definite interpretation requires specification of number and person. In the case of T this implies that a valued D-feature, a property characteristic of consistent null-subject languages, as we have argued, presupposes specified/valued Number and Person features. This does not in itself entail that the specification must be morphologically expressed, though. Recall, however, that T in construction with a null subject is the head of an A-chain, in consistent null-subject languages. Suppose that the head of a chain with definite value (a referential index) must be spelled out, as a matter of UG.[16] As long as the D-feature of T has no morphological expression of its own, the only way it can be spelled out is by spelling out the person and number features. This might explain, at least in part, why consistent null-subject languages tend to have morphologically expressed person and number. We do not, then, expect to see a consistent pro-drop language with a seriously defective agreement paradigm, while partial null subject languages are more varied in this regard: from the richly articulated paradigm of Finnish to the completely agreementless system typical of many East Asian languages (see next section).

4. Non-null subject languages and discourse pro-drop languages

Non-null subject languages typically have a poor subject-verb agreement system, or no subject-verb agreement at all. Again, it is tempting to explain this in terms of Roberts's (2007) theory: These are languages where T has fewer features than pronouns, consequently pronouns cannot be incorporated but must be spelled out. If T in these languages has even fewer features than in partial null-subject languages, say, if T does not even have a number feature, then this might explain why even a generic pronoun cannot be incorporated.

However, to begin with, some non-null-subject languages have agreement systems which are at least as diversified as the systems found in BP, Finnish, and Marathi: This is the case for German as well as French, two well-studied non-null-subject languages. Second, we still need to account for why non-null-subject languages do not even allow null subjects that are controlled from a higher clause, the way partial null-subject languages do. That is to say, not only do they not permit incorporation of a generic pronoun, but must spell it out and (consequently) move it to specTP to check the EPP, as in (28a), but they also do not allow a controlled null pronoun in specTP (or specCP), as in (28b).

(28) a.   Jari sa   att *(man) sitter bekvämt här.   (Swedish)

            Jari said that   one   sits   comfortably here

      b.   Jari sa   att *(han) sitter bekvämt här.

            Jari said that   he   sits comfortably here

We therefore concur with Holmberg (2005) that non-null-subject languages accept neither impersonal null subjects nor controlled null subjects because they have a 'phonological EPP' in T (labelled [P] in Holmberg (2000) and Landau (2006)). In principle any head may have this

29

feature, as a parametrized property. In the case of T, the effect of [P] is that the subject probed by T is spelled out, i.e. assigned a phonological form. Provided with a phonological form it cannot be incorporated, and so will normally be targeted by the EPP and moved to specTP.

There is a class of languages which have no subject-verb agreement at all, yet allow null subjects (and null pronouns more generally), interpreted by recourse to an antecedent in the discourse. These languages are particularly common in East Asia (Chinese, Japanese, Korean, Thai, etc.), but also many Dravidian languages belong to this class. Consider the following examples, from Cantonese Chinese and Japanese:

(29) a.   Ah John waa hai Jinggwok jiu   gong Jingman   [Cantonese]

            Prt John say   in England   need speak English

            'John says that one/he needs to speak English in England.'

       b.   John-wa kono beddo-de-wa yoku nemu-reru-to iu.   [Japanese]

            John-TOP this bed-in-TOP well sleep-can-COMP say

            'John says that one/he can sleep well in this bed.'

Apparently these sentences are truly ambiguous between a generic and a personal, controlled reading. The fact that null subjects are possible at all means that they do not have [P] in T. The fact that a generic reading of a null subject is possible means that they do not have uD in T, while the fact that they can have null controlled pronouns means that their pronouns have a uD-feature. In other words, they are partial null-subject languages, by our definitions.

(29a,b) indicate that the absence of a generic 3SG null subject in consistent null-subject languages is not about avoiding ambiguity (as was suggested in section 1.3). Ambiguity is

30

tolerated, here as in many other contexts. Instead, if we are right, it is an effect of the uD-feature in T, which will always assign a definite interpretation to a subject φP.

5. Control of null subjects in finite clauses

As discussed, BP, Finnish, and Marathi all allow null subjects in finite clauses controlled by an antecedent in a higher clause. But unlike the situation for consistent null-subject languages, the antecedent argument controls the null subject directly (rather than via a null-topic chain). In this respect it is more similar to control of PRO in non-finite clauses. Indeed, Landau (2004) has shown that control into finite clauses in Hebrew, another partial null-subject language, is Obligatory Control (OC), as familiar from non-finite clauses. As we will show, while the conditions for control into finite clauses vary to some extent between the three languages under investigation here, none of them conform to OC. In the following we will first show what the languages have in common, listing the contexts where controlled null subjects are allowed in all three languages, and subsequently show how they differ from each other. For reasons of space we will use a minimum of example sentences. See Holmberg, Nayudu & Sheehan (to appear) for a more thorough exposition of control into finite clauses in these three languages.

5.1. Similarities between BP, Finnish, and Marathi

5.1.1. Complements of verbs of saying, thinking, and perceiving. They all allow subject control into complements of verbs of saying, thinking, and perceiving.

(30) a. O João$_1$ disseque (ele$_1$) tinha comprado uma casa .

31

the João said that he had bought a house

'João said that he had bought a house.'

b. Marja$_1$ luulee että (hän$_1$).on hyvän näköinen.

Marja thinks that she is good looking

c. Seema$_1$ kabul karte ki (ti-chya-ni$_1$) chuk dzali

Seema agree does that she-ERG mistake made

'Seema admits that she made a mistake'.

In all three languages the embedded clauses can be temporally independent of the matrix clause, as illustrated here by Marathi.

(31) Seema$_1$ kabul karte <u>aaj</u> ki (ti-chya-ni$_1$) <u>kal</u> chuk dzali

Seema agree does today that she-ERG yesterday mistake made

'Seema admits today that she made a mistake yesterday'.

This is different from Hebrew, investigated in Landau (2004). In Hebrew, control into finite clauses presupposes that there is a temporal dependency between the null subject clause and the antecedent clause similar to the dependency characteristically holding between an embedded subjunctive clause and a matrix clause. Indeed, Landau argues that the null subject-containing finite clauses in Hebrew are 'covertly subjunctive'.

5.1.2. Adjunct clauses. They all allow control into finite adjunct clauses (SUBJ = subjunctive).

.

(32) a. O João pode vir, desde que ∅ termine o trabalho.

32

The John may come, since that ∅ finishes-SUBJ the job

'John can come, if he finishes the job (first).'

b.   Eeva saa tulla mukaan jos (hän) lupaa   olla hiljaa.

Eeva may come along if   she   promises be   quiet

b.   John₁ khush hota   karan   (tya-la)₁   pushkar bheti milyala

John happy   be-past3sm because (he-ACC) very   gifts receive-past3plf

'John was happy because he received many gifts.'

5.1.3. Indirect questions. BP, Finnish, and Marathi allow control into embedded questions, at least marginally.[17]

(33) a.   O João perguntou se (ele) podia dormir aqui.

The John asked   if   he could sleep here

'John asked if he could stay the night.'

b.   Jari haluaa tietää saako   (hän) jäädä yöksi.

Jari wants know can-Q   he   stay night-TRA

'Jari wants to know whether he could stay the night.'

(c)   John-ni   vicharle   ki (to) ratri rahu shakto ka?

John-ERG ask-past-3SN that (he) night stay happen-PRS-3SM Q

'John asked whether he could stay the night.'

5.1.4. Locality. None of the three languages allow a control relation across another subject, even if that subject has features incompatible with the null subject (visible on the embedded

33

finite verb) . Marathi is possibly somewhat less strict than BP and Finnish in this regard, though.

(34) a.   O João$_1$ disse   [que   os moleques$_2$ acham   [que   *(ele)$_1$ é   esperto]]

The João   said   that   the   kids   think   that   he   is   smart

'João said that Maria thinks that *he/she is smart.'

b.   Jari sanoo [että lapset   uskovat   [että *(hän)

Jari says   that children believe-PST-3PL that   he

kävi   tohtorilla]]].

visited-PST-3SG doctor-ADE

'Jari says that the children believe that he went to see a doctor.'

c.   Ram-ni$_1$ mhantlə   ki Mary-la   watlə   ki ?(to)$_1$ doktaran-

Ram-ERG say-PST-3SN that Mary-ACC think-PST-3SN that he doctor-

kade gela

to   go-PST3SM

'Ram said that Mary thought that he went to the doctor'.

However, when the intervening clause contains no argument all three languages tolerate control by an argument which is more than one clause away.

(35) a.   A Maria$_1$ disse [que é verdade [que (ela)$_1$ entornou   o copo]]

The Maria said that is true   that   she   knocked.over the glass

'Maria said it's true that she knocked over the glass.'

b.   Jukka$_1$ sanoi [että oli onni   [että (hän)$_1$ oli   voittanut arpajaisissa]].

34

Jukka said   that was fortune that he   had-3SG won   lottery-INE

'Jukka said that it was fortunate that he had won in the lottery.'

<center>(based on Rodrigues (2000)</center>

c.   Ram-la$_1$ watto   [ki he changla dzala   [ki (to)$_1$

Ram-DAT think-PRS-3SM that this good   happen-PST that he

doktaran-kade gela]]].

doctor-to   go-PST-3SM

'Ram thinks that it was good that he went to the doctor'

The judgments in the case of BP and Finnish are quite subtle, though, and subject to lexical variation. In both languages even an implicit argument will act as intervener. Thus a null subject is not possible in (36), conceivably because the predicate 'obvious' takes an implicit experiencer argument 'to us/ anybody', which acts as intervener.

(36) Jukka sanoo [että on ilmeistä [että *(hän) on voittanut arpajaisissa]].

Jukka says that   is obvious   that   he   has won   lottery-INE

5.1.5. Sloppy identity and bound variable reading. A well known test for Obligatory Control (OC) is whether ellipsis requires sloppy identity, as in (37a). Another one is whether *only* scoping over the antecedent induces a bound, as opposed to a coreferential reading of the null subject   (Landau (2000), Hornstein (1999)):

(37) a.   John expects to be invited, and so does Mary.

b.   Only John expects to win.

(37a) can only mean that Mary expects that she will be invited, and (b) can only mean 'John is the only x who expects x to win' (the bound reading), and cannot mean 'John is the only x who expects John to win' (the coreferential reading). These are characteristic properties of OC, as opposed to Non-Obligatory Control (NOC).

In the case of control into finite clauses in Finnish we do not see these effects:

(38) a. Marja luulee että (hän) on ovela, ja niin luulee Jarikin.

Marja thinks that   she is clever and so does Jari-too

'Marja thinks that she is clever, and so does Jari.'

b. Vain John uskoo että Ø voittaa vaalit.

only John thinks that   wins   elections

'Only John thinks that she will win the elections.'

(38a) allows a strict or a sloppy reading, regardless whether the pronoun is overt or covert. (b) allows the bound reading but also allows the coreferential reading.[18]

Similarly in Marathi, (39) allows a strict or a sloppy reading, regardless whether the pronoun is overt or covert.

(39) Seema-la wattə ki (ti-ni)   hi pustak vaachli aahe ani

 S-ACC   thinks that she-ERG this book   read   is   and

tasach   Ram-la   pan wattə

similarly Ram-ACC also thinks

And (40) allows a bound as well as a coreferential reading, regardless whether the pronoun is overt or covert.

(40) ho, Ram-la   wattə ki (to) jinkel

     yes Ram-ACC thinks that he win-FUT

     'Yes, Ram thinks that he will win'

Thus it can be the answer to either the question (41a) or (41b):[19]

(41) a.   kontya-hi umiddhwar-la wattə   ka ki to jinkel?

           who-EMPH candidate-ACC think-PRS-3SN QM that he win-FUT

           'Does any candidate think that he will win?'

     b.   kontya-hi umiddhwar-la wattə   ka ki Ram jinkel?

           who-EMPH candidate-ACC think-PRS-3SN QM that Ram win-FUT

           'Does any candidate think that Ram will win?'

We return to BP below, after considering the case of split antecedence.

5.1.6. Split antecedents. A well known difference between OC and NOC is that only NOC allows split antecedents (Landau 2000, Hornstein 1999).

    In BP a null subject in a finite clauses can have split antecedents where one is plural.[20]

(42) a.   A Maria1 disse que o João2 acredita que *(eles)1+2 vão morar juntos

           the Maria said that the João believes that will-3pl live-INF together

37

'Maria said that João believes they will live together'

b. O Zé₂   convenceu   os meninos₁   que (eles)₁₊₂   tinham que ir embora.

The Ze   convinced   the kids   that they   had   that go away

'Zé convinced the kids that they had to leave.'

In Marathi, too, a null subject in a finite embedded clause can have split antecedents.

(43) Mary-ni Lucy-la sangitl   ki (te) ekatr   jau shaktat

M-ERG   L-ACC   say-PST-3SN that they together go happen-PRS-3PL

'Mary told Lucy that they can-3PL travel together. '

In Finnish, our informants consider the split antecedents in (44a) to be degraded when compared with the non-split antecedent in (44b).

(44) a.   Marja kertoi Jarille etteivät   *?(he) voi matkustaa yhdessä.

Marja told Jari   that-not-3PL they can travel   together

'Marja told Jari that they can't travel together.'

b.   Marja kertoi Jarille ettei   (hän) voi matkustaa hänen kanssaan.

Marja told Jari that-not-3SG she   can travel him with

'Maraha told Jari that she can't travel with him.'

We contend, nevertheless, that the unacceptability is not of the order expected if the control relation were a case of OC, as in (45), for example.

(45) *Mary told John to leave together.

We will see another fairly acceptable example of split antecedents in Finnish in section 5.2.4 below. We therefore, tentatively, conclude that Finnish is not significantly different from Marathi and BP in this respect.

As regards sloppy identity and bound variable reading with 'only' in BP, Rodrigues (2004) and Modesto (2000, to appear) citing Negrão (1999), both claim that only sloppy identity is allowed in (46) whereas (47), with an overt pronoun, becomes ambiguous between a strict and sloppy reading:

(46) A Maria$_1$ encucou que e_1 estava grávida e o Paulo também   (sloppy only)

the Maria worried-3Sg that was-3Sg pregnant and the Paulo too

'Maria got worried that she was pregnant and Paulo did too'

(47) A Maria encucou que ela estava grávida e o Paulo também   (strict/sloppy)

the Maria got.worried-3Sg that she was-3Sg pregnant and the Paulo too

'Mary got worried that she was pregnant and Paulo did too'

Rodrigues takes this as evidence that the control relation is OC (derived by movement, following Hornstein's (1999) theory of control). Modesto takes it as evidence that it is A'-binding.

Our informants are not quite so categorical regarding (46) and (47). Note also that OC and A'-binding are both quite incompatible with split antecedence. Thus it seems that still more

research is called for before regarding control into finite clauses in BP. We conclude, perhaps prematurely, that the three languages are not significantly different with respect to the tests for OC vs.NOC.

5.2 Differences among BP, Finnish, and Marathi

5.2.1 Noun complements and factive clauses. BP and Finnish both allow control into the finite clausal complement of a noun selected by a verb, Marathi does not. (49b) exemplifes a very common form of verb complementation in Finnish, where the complement is headed by the pronoun *se* 'it' (appropriately inflected).

(48) a. O João$_1$ se esquece do   fato de que (ele)$_1$ vai ganhar menos no novo emprego.

The John SE forgets of.the fact of that   will earn less   in.the new job

'John forgets about the fact that he'll earn less in his new job.'

b. O presidente$_1$ negou os rumores de que (ele)$_1$ tinha

The president   denied the rumors of that he   had

recebido   dinheiro de   empresários

received   money from businessmen.'

'The president denied the rumours that he had received money from businessmen.'

[example from Modesto (2000b:99)]

(49) a. Anu ei usko väitteeseen että   (hän) olisi   maksanut autostaan liika.

Anu not believes claim that   she   has-CON paid   car-ABL-her too-much

'Anu doesn't believe the claim that she would have paid too much for her car.

40

b.   Jari valitti   meille siitä   että (hän) joutui maksamaan autostaan

        Jari complained us-ALL it-ABL that   he   had-to pay   car-ABL-his

        liikaa.

        too-much

        'Jari complained to us about the fact that he had to pay too much for his car.'

(50) John-la   [hi goshta ki   *(tyan-ni) ti-la   dukhavla] mahiti   nahvti

       John-acc the fact that   he-ERG she-ACC offended was-aware not

       'John wasn't aware of the fact that he had offended her.'

BP and Finnish also allow control into finite factive clauses. Marathi does not.

(51) a.   A Maria$_1$ sente muito que (ela$_1$) tenha   chegado/chegou tarde.

        the M   feels a lot that she has-SUBJ arrived/arrived late

        'Maria regrets that she has arrived late.'

    b.   Jari$_1$ katuu   että (hän$_1$) tuli takaisin.

        Jari regrets that he came back

    c.   Ram-la   pashchyatap hota ki *(to) parat aala

        Ram-ACC regret   was that he back came

        'Ram regretted that he came back.'

These two cases can be collapsed. It seems fairly uncontroversial to assume that factive clauses are underlying nominal clauses, headed by an abstract noun meaning 'fact'. If so, the null subject in (51c) is ruled out for the same reason that the null subject in (50) is.

41

5.2.2. Directive verb complements. As for object control into complements of directive verbs, Finnish allows it (somewhat marginally), BP and Marathi do not. [21]

(52) a.　O José　recomendou　ao　　João$_1$ que *(ele$_1$) devia　lavar　a louça.

　　　the Jose recommended to-the J　　that　he　should wash the dishes

　　　'José recommended to João that　he should wash the dishes.'

　　b.　Jari$_1$ ehdotti　Tarjalle$_2$　[että (hän$_{*1/2}$) ostaa　uuden telkkarin].

　　　Jari　suggested Tarja-ALL　that　she　buys　new　TV

　　　'John suggested to Tarja that she should buy a new TV.'

　　c.　Ram-ni　　Arun-la　　shikawlə ki *(tya-ni)　mothyaan-cha aader　karavə

　　　Ram-ERG　Arun-ACC taught　　that　he-ERG　elders-of　　respect　do-SUBJ

　　　'Ram taught Arun that he should respect the elders'.

This is a context where Hebrew happily allows control, provided that the embedded clause is temporally dependent (future relative to the time of the matrix clause) and generally has a subjunctive-like dependent relation to the matrix clause (Landau 2004; see also Gutman 2004).

5.2.3. Object control with verbs of communication In complements of verbs of communication ('tell (somebody something)', 'inform', 'convince', 'remind', 'warn') Marathi allows object control, somewhat surprisingly. In BP and Finnish this is a marginal possibility, at best.

(53)a.　John-ni$_1$ Mary-la　　kalav-l-ə　　　　　ki　Ø$_1$/ to$_{1/2}$ parikshet pas

　　　John-ERG Mary-DAT　informed-PST-3SN that　　　he　exam-in　pass

dza-l-**a**

happen-PST-3SM

'John informed Mary that ec passed the test'.

b.   John-ni   Mary-la$_1$ kalav-l-ə   ki   Ø$_1$/ ti$_{1/2}$ parikshet pas

John-ERG Mary-DAT inform-PST-3SN that   she   exam-in pass

dza-l-**i**

happen-PST-3SF

'John informed Mary that ec passed the test'.

c.   Lucy-ni$_1$ Mary-la$_2$ kalav-l-ə   ki   Ø$_{*1/2}$ / ti$_{1/2/3}$ parikshet pas

Lucy-ERG M- DAT inform-PST-3SN that   she   exam-in pass

dza-l-**i**

happen-PST-3SF

'Lucy informed Mary$_2$ that she$_2$ passed the exam'.

In (53a) the embedded verb is inflected for masculine, consequently only the subject is a possible controller of the null subject. In (b) the inflected verb is feminine, and the only possible controller is the matrix object. In (c) there are two c-commanding DPs with features compatible with those of the null subject. In this case the closer one, that is the object, is the antecedent.

In Finnish, on the other hand, the subject is the preferred controller in the case where the embedded verb inflection is compatible with both subject and object control: see (54a). However, object control is possible when the inflection on the embedded verb is incompatible with subject control, as in (54b). An alternative reading, in this case, is the split antecedent reading.

43

(54) a.   Pekka$_1$ muistutti Juhania$_2$ että Ø$_{1/?2}$ oli luvannut   leikata nurmikkoa.

              Pekka reminded John   that   had promised mow   lawn

              'Pekka reminded John that he had promised to mow the lawn.'

      b.   Pekka$_1$ muistutti lapsia$_2$ että Ø$_{*1/2/1+2}$ olivat   luvanneet   leikata nurmikkoa.

              Pekka reminded children that   had-3PL promised-PL mow   lawn

              'Pekka reminded the children that they (the children or him and the children) had

                  promised to mow the lawn.'

BP exhibits a similar situation: subject control is preferred but object control or split
antecedents are possible when subject control is ruled out.

(55) a.   O Pedro$_1$ convenceu o   João$_2$ que Ø$_{1/*2}$ tinha que ir embora. (Modesto 2000)

              the Pedro convinced the João   that   had to go away

              'Pedro$_1$ convinced João that he$_1$ had to leave.'

      b.   O Zé$_1$ convenceu os meninos$_2$ que Ø$_{*1/?2/1+2}$ tinham que ir embora

              The Ze convinced the kids   that   had   to go away

              Ze convinced the kids that they (the children or him and the children) had to leave.'

An interesting observation, due to Modesto (2000a, 2000b) is that object control is also
preferred in BP if the object is wh-moved or topicalized.

(56) Quem$_2$ que   o Pedro$_1$ convenceu t$_2$ que EC$_2$ tinha que ir embora?

      Who$_1$ that the Pedro$_2$ convinced   that   had   that go away?

'Who did Pedro convince that he had to leave?

As discussed by Modesto (to appear) Finnish exhibits a similar, though not identical, situation: Object control improves if the object is wh-moved or topicalized, but only marginally.

(57) ?Ketä$_1$ Pekka muistutti t$_1$ että (hän$_1$) oli luvannut leikata nurmikkoa.

   who   Pekka reminded   that   he   had promised mow lawn

'Who$_1$ did Pekka remind that he$_1$ had promised to mow the lawn?'

The clearest difference, then, is the preference for controller when there are two potential controllers: The object in Marathi, the subject in BP and Finnish. An interesting possibility is that this is related to an obvious parametric difference between Marathi on the one hand, BP and Finnish on the other hand: OV vs. VO (see Nayudu 2008). Two assumptions are required: (a) Clausal complements of this class of verbs are adjuncts, possibly as a result of movement (Ferreira 2004), and (b) OV order is derived by object movement out of VP (Kayne 1994, Julien 2002)). The result is that, in the unmarked case, the object will c-command the clausal complement (in its derived position) in the OV language Marathi, but not in the VO languages BP and Finnish. Thereby the object is a potential controller of the null subject, and in fact the preferred one, presumably because it is the closest one. This is supported by Modesto's observation concerning object movement in BP and Finnish: When the object moves out of VP, it becomes the preferred controller in BP and Finnish, too.

This obviously raises a number of questions. Perhaps the most pressing one is how it is possible for the object to control the null subject in (54b) and (55b)? We will leave this question and other related questions for future research.

45

5.2.4. Relative clauses. Only Finnish allows control into relative clauses.

(58) a. Pekka rikkoi maljan jonka (hän) oli saanut lahjaksi   sukulaisiltaan.

Pekka broke vase   which (he) had got   gift-TRA relatives-ABL

'Pekka broke the vase that he had got as a gift from his relatives.'

b. John-ni$_1$ masale tya sauce-madhe misalavale dzo *(tyan-ni)$_1$ aadhi-cha

John-ERG herbs that sauce-in   mixed   that   he-ERG earlier-EMPH

kelela.

prepared

'John mixed the herbs in the sauce that he had prepared earlier.'

In the case of Marathi this falls under the generalization that control into nominal

complements.is not allowed. For BP there must be a different explanation, though, since BP

allows control into nominal complements

5.2.5. C-command. An interesting difference separating Finnish from Marathi and BP is that

Finnish allows control into a finite clause in the following constructions:

(59) a. Jarin$_1$ suunnitelma oli [että (hän$_1$) ajaisi   yhdessä päivässä

Jari's plan   was that he   drive-CON one-INE day-INE

 Helsingistä   Ouluun].

Helsinki-ABL Oulu-ILL

'Jari's plan was that he would drive in one day from Helsinki to Oulu.'

46

b. Se on Juhanin₁ oma vika että (hän₁) sai potkut.

   it is John's   own fault   that   he   got kicks

   'It is John's own fault that he got the sack.'

c. Se oli Eevalle₁   pettymys   ettei   (hän₁) voinut jatkaa opintojaan.

    it was Eeva-ALL disappointment that-not she   could continue studies-her

   'It was a disappointment to Eeva that she couldn't continue her studies.'

Neither BP nor Marathi allow control in corresponding constructions (the counterpart to (60c)

is not constructed as a predicative construction in Marathi.)

(60) a. *A ideia do   José₁ era que *(ele₁) ia   embora imediatamente.

   the idea of-the José was that   he   went away   immediately

   'José's idea was that he would leave immediately.'

   b. *Foi uma pena   para a Maria₁ quando *(ela₁) não passou o exame.

    was a   shame for the Maria   when   she not passed the exam

    'It was a shame for Maria when she did not pass the exam.'

(61) a. John-ch   vichaar hotə   ki *(to) lawkar nighel

   John-GEN plan   be-PST that he early leave-FUT

   'John´s plan was that he would leave early.'

   b. Mary-ch   lakshy   aahe   ki *(ti) prasidh abhinetri honaar.

   Mary-GEN intention/aim be-PRS that she famous actress happen-FUT

   'Mary's intention is that she will become a famous actress.'

The antecedent obviously does not c-command the null subject in (59). Generally speaking c-command is a requirement for control into finite clauses in Finnish, too, though. Consider (62a,b), where *Jari* in (62a) cannot control the null subject even though pragmatically it is the only possible antecedent, but can do so in (62b), where it c-commands the null subject.

.

(62) a.   [Jarin puhe] teki   selväksi ettei   *(hän) ole syyllinen..

              Jari's speech made clear   that-not he   is guilty

              'Jari's speech made clear that he isn't guilty.'

      a.   Jari teki puhessaan   selväksi ettei (hän) ole syyllinen.

              Jaro made speech-INE-his clear   that-not he is guilty

              'Jari made it clear in his speech that he isn't guilty.'

Infinitival clauses in many languages, including English, allow control in constructions corresponding to (59).

(63) a.   John's plan was [PRO to drive to Edinburgh].

      b.   It's in Jane's interest [PRO to be on time].

Characteristic of the construction is that it involves the copula. Apparently this is a case of control as an effect of 'connectivity under specificational predication'. It is well known that certain relations can hold between the terms of copular specificational predication which otherwise require c-command; see Heycock and Kroch (1999) and references there. Heycock and Kroch argue that the type of predication found in the constructions in question is equative predication. They discuss specifically pseudoclefts, but the theory can be extended to control

48

constructions; see Lyngfelt (2002). At some level of representation, the expressions in (59) and (63) consist of two constituents, each made up of predicate and an argument, which are related by equative predication. One of the constituents has a specified subject, the other a null subject. In this situation, control may occur between the specified subject and the null subject.

(64)   [$_{XP}$ John [plan]] = [$_{XP}$ PRO [drive to Edinburgh]]

       PRO = John

Control via connectivity under specificational predication would then seem to be an option made available by UG. We have no idea why Finnish avails itself of this option in finite clauses while BP and Marathi do not.

5.3   Conclusions of section 5

There are differences among the three languages as regards control of a null subject in embedded finite clauses. Marathi, for some reason, does not allow control into complements of nouns. Furthermore, Marathi prefers object control in at least one case where BP and Finnish do not; we suggested ascribing this difference to the fact that Marathi is an OV language while BP and Finnish are VO languages. The general impression is, though, that we are dealing with variations on a theme: Control of a null subject in finite clauses is essentially the same phenomenon, subject to the same constraints, in the three languages, with variations that are (at least some of them) due to independent parametric differences among the languages.

A comparison with control of PRO in non-finite clauses yields an unambiguous conclusion: Control of a null subject in finite clauses in BP, Finnish, and Marathi is not OC.

49

The locality conditions are not as strict as for OC, the possibility of split antecedents is incompatible with OC, as is the strict reading under ellipsis and the coreferential reading in the *only*-construction (clearer in Finnish and Marathi than in BP).

This suggests that control of a null subject in finite clauses in BP, Finnish, and Marathi might be NOC, thus closely related to the control relation in, for example, (65):

(65) John thinks that [PRO shaving himself] is not important.

However, c-command (or connectivity under predication) is not a requirement for NOC.

(66) John's outburst made it clear that [PRO behaving himself in public] is not important to him.

Informally speaking, control of a null subject in finite clauses in BP, Finnish, and Marathi is stricter than NOC but not as strict as OC.

6. **Conclusions**

BP, Finnish, and Marathi share the following properties:

(a)   They do not have a [P] feature paired with [φ,T]. This means that they allow null subjects in contexts where non-null-subject languages such as English, French, Sindhi, etc. require a pronounced subject.

(b)   They do not have a [uD] feature in T which could receive a value (a referential index) from a null A-topic (itself part of an A-topic chain), which it could pass on to a φP subject via

50

Agree and concomitant incorporation (Roberts 2007), thereby deriving a definite null subject chain headed by T and linked indirectly to an A-topic in the discourse context.

(c)   On the other hand, the absence of a [uD] feature in conjunction with property (a) means that a ϕP subject can enter a chain headed by T (i.e. be incorporated in T, in Roberts' (2007) sense), thus ending up as a null copy of the ϕ-features of T, but only with a generic or impersonal interpretation. In this case another category is needed to satisfy the EPP.

(d)   Instead of being incorporated in T, a pronoun may move and re-merge with TP, satisfying the EPP. Given Roberts (2007), this pronoun must be more richly specified than the incorporating one; we assume that it is a DP. In specTP it will be spelled out, as the highest member of a chain, unless it is moved to an even higher position. Alternatively it may remain null, in which case it must be controlled by a local antecedent. This null pronoun, we have argued, is a DP with a uD-feature, therefore uninterpretable unless it is controlled by a c-commanding argument, typically, but not necessarily, in the next clause up (connectivity under specificational predication is an alternative to c-command in Finnish).

(e)   There is some variation among the three languages regarding the details of the control relation. However, in all three languages it holds that the relation is neither OC nor NOC, but a third type of control relation, whose precise nature is not well understood.

References

ALEXIADOU, A. & E. ANAGNOSTOPOULOU.1998. Parameterizing AGR: Word Order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16, 491-539.
BARBOSA, P. 1995. *Null Subjects*. MIT, Cambridge MA. Ph.D. diss.
BARBOSA, P. To appear. Two kinds of subject *pro*. To appear in *Studia Linguistica*.

BOBALJIK, J. *2002*. A-chains at the PF-interface: Copies and 'covert' movement. *Natural Language and Linguistic Theory 20, 167-267.*

CARDINALETTI, A. & M. STARKE 1999. The typology of structural deficiency: A case study of the three classes of pronouns. *Clitics in the languages of Europe*, ed. H van Riemsdijk, 145-235. Berlin: Mouton de Gruyter.

CHOMSKY, N. 1995. *The Minimalist Program.* Cambridge MA: MIT Press.

CHOMSKY, N. 2000. Minimalist inquiries: the framework. *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik,* eds. R. Martin, D. Michaels, and J. Uriagereka, 83-155. Cambridge MA: MIT Press.

CHOMSKY, N. 2001. Derivation by phase. *Ken Hale: A Life in Language*, ed. M. Kenstowicz, 1-59. Cambridge Mass.: MIT Press.

CHOMSKY, N. 2008. On Phases. *Foundational issues in linguistic theory*, eds. R. Freidin, C. P. Otero & M.L. Zubizarreta.. Cambridge. MA: MIT Press.

COLE, M. Forthcoming**.** Null subjects and accessibility.

DÉCHAINE, R-M. and M. WILTSCHKO. 2002. Decomposing pronouns. *Linguistic Inquiry* 33, 409-442.

DUARTE, E. 1995. *A perda do principio "evite pronome" no Português Brasileiro*. Instituto de estudos da linguagem, UNICAMP. Ph.D. diss.

DUARTE, M. E. L. 2000. The loss of the 'Avoid Pronoun' principle in Brazilian Portuguese. *Brazilian Portuguese and the null subject parameter*, ed. M. Kato and E Negrão, 17-36. Frankfurt: Vervuert.

FERREIRA, M. 2004. Hyperraising and null subjects in Brazilian Portuguese. *MIT Working Papers in Linguistics* 46, Collected papers on Romance syntax.

FIGUEIREDO SILVA, M. 2000. Main and embedded null subjects in Brazilian Portuguese. *Brazilian Portuguese and the null subject parameter*, eds. M. Kato and E. Negrão. Germany: Vervuert.

FRASCARELLI, M. 2007. Subjects, topics, and the interpretation of referential pro. *Natural Language and Linguistic Theory* 25, 691-734.

FRASCARELLI, M. and R. HINTERHÖLZL. 2007.Types of topics in German and Italian. *On information structure, meaning, and form*, eds. S. Winkler and K. Schwabe, 87-116. Amsterdam/Philadelphia: John Benjamins.

GRIMSHAW, J. & V. SAMEK-LODOVICI. 1998. Optimal subjects and subject universals. *Optimality and competition*, eds. P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky., 193-220. Cambridge Mass.: MIT Press.

GUTMAN, E. 2004. Third person null subjects in Hebrew, Finnish, and Rumanian: an accessibility theoretic account. *Journal of Linguistics* 40, 463-490.

HAEGEMAN, L. 2000. Adult null subjects in non-pro-drop languages. *The acquisition of syntax* eds. M.A. Friedemann & L. Rizzi,129-169. Harlow: Longman.

HEYCOCK, C. and A. KROCH. 1999. Pseudocleft connectivity: Implications for the LF interface level. *Linguistic Inquiry* 30, 365 – 397.

HOLMBERG, A. 2000. Scandinavian stylistic fronting: How any category can become an expletive. Linguistic Inquiry 31:445–483.

HOLMBERG, A.. 2001. The syntax of yes and no in Finnish. *Studia Linguistica* 55, 141-175.

HOLMBERG, A. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36, 533-564.

HOLMBERG, A. To appear. The silence of the generic pronoun in Finnish. To appear in *Null subjects and parameters in a minimalist perspective*, eds. T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan. Cambridge: Cambridge University Press.

HOLMBERG, A. and U. Nikanne. 2002. Expletives, subjects and topics in Finnish. *Subjects, expletives, and the EPP*, ed. P. Svenonius, 71-105. Oxford and New York: Oxford University Press.

HOLMBERG, A., A. NAYUDU and M.SHEEHAN. To appear. Control into finite clauses in Brazilian Portuguese, Finnish, and Marathi. To appear in *Null subjects and parameters in a minimalist perspective*, eds. T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan. Cambridge: Cambridge University Press.

HORNSTEIN, N. 1999. Movement and control. *Linguistic Inquiry* 30, 69-96.

HUANG, Y. 2000. *Anaphora. A cross-linguistic study*. Oxford and New York: Oxford University Press

JULIEN, M. 2002. *Syntactic heads and word formation*. Oxford/New York: Oxford University Press

KAYNE, R. 1994. *The antisymmetry of syntax*. Cambridge Mass.: MIT Press.

KATO, M. 1999. Strong and weak pronominals in the null subject parameter. *Probus* 11:1-37.

KATO, M. 2000. The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. *Brazilian Portuguese and the null subject parameter*, eds. M. Kato and E Negrão. Frankfurt: Vervuert.

LANDAU, I. 2000. *Elements of control. Structure and meaning in infinitival constructions*. Kluwer, Dordrecht.

LANDAU, I. 2004. The scale of finiteness and the calculus of control. *Natural Language & Linguistic Theory* 22:811–877.

54

LYNGFELT, B. 2002. OT semantics and control. *Rutgers optimality archive* 411.

MARTINS, A. M.. 2006. Emphatic affirmation and polarity. Contrasting European Portuguese with Brazilian Portuguese, Spanish, Catalan, and Galician. *Romance languages and linguistic theory* 2004, eds. Jenny Doetjes and Paz Gonzales. Amsterdam: John Benjamins.

MATUSHANSKY, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37, 69-109.

MODESTO, M. 2000. *On the identification of null arguments*. University of Southern Califonia, PhD.diss.

MODESTO, M. To appear. Topic prominence and null subjects. To appear in T. Biberauer (ed.) *The limits of syntactic variation*. Amsterdam: John Benjamins.

NAYUDU, A. 2008. *Issues in the syntax of Marathi*. A minimalist apporach. Durham University. Ph.D. diss.

NEGRÃO, E. 1997. Asymmetries in the distribution of overt and empty categories in Brazilian Portuguese. *Clitics, pronouns and movement,* eds. J. R. Black and V. Motapanyase, 217-235. Amsterdam & Philidelphia: John Benjamins.

NUNES, J. 2004. *Linearization of chains and sideward movement*. Cambridge MA: MIT Press.

PLATZACK, C. 2004. Agreement and the person phrase hypothesis. *Working Papers in Scandinavian Syntax* 73: 83-112. Lund University.

RIZZI, L. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

RIZZI, L. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17, 501-557.

ROBERTS, I. 2007. Clitics, head movement, and incorporation. Ms. University of Cambridge.

ROBERTS, I.. To appear. A deletion theory of null subjects. To appear in *Null subjects and parameters in a minimalist perspective*. Ed. T. Biberauer, A. Holmberg, I. Roberts, M. Sheehan. Cambridge: Cambridge University Press.

RODRIGUES, C. 2000. *Deriving referential null subjects from movement*. University of Maryland, Ph.D. diss.

RODRIGUES, C. 2002. Morphology and null subjects in Brazilian Portuguese. *Syntactic effects of morphological change*, ed. A Lightfoot, 160-178. Oxford: OUP.

RODRIGUES, C. 2004. *Impoverished morphology and A-movement out of case-domains*. University of Maryland. Ph.D diss.

ROSS, J.R. 1970. On Declarative Sentences. *Readings in English Transformational Grammar*, ed. R.A. Jacobs and P.S. Rosenbaum. Waltham, Mass.: Ginn.

SAMEK-LODOVICI, V. 1996. *Constraints on subjects: an optimality theoretic analysis*. Rutgers University, Ph.D. diss.

SHEEHAN, M. 2006. The EPP and null subjects in Romance. Newcastle University, Ph.D. diss.

SHEEHAN, M. To appear. Free inversion in the Romance languags: a case of micro-parametric variation.

SHLONSKY, U. To appear. Hebrew as a partial null-subject language. To appear in *Studia Linguistica*.

SIGURÐSSON, H.A. 2004. The syntax of person, tense, and speech features. *Rivista di Linguistica/Italian Journal of Linguistics* 16, 219-251.

SIGURÐSSON, H.A. and V. EGERLAND. To appear. Impersonal null subjects in Icelandic and elsewhere. To appear in *Studia Linguistica*.

VAINIKKA, A. 1989. *Deriving syntactic representations in Finnish*. University of Massachusetts, Amherst, Ph.D. diss.

VAINIKKA, A. & Y. LEVY. 1999. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory* 17: 613-671.

Anders Holmberg

School of English Literature, Language, and Linguistics,

Newcastle University,

Newcastle upon Tyne, NE1 7RU

UK

anders.holmberg@newcastle.ac.uk

Aarti Nayudu

103, Summit-2, near Shaligram-2,

Prahladnagar Road, off 100 Feet Road,

Ahmedabad 380051

Gujarat, India

a_nayudu@yahoo.com

Michelle Sheehan

School of English Literature, Language, and Linguistics,

Newcastle University,

Newcastle upon Tyne, NE1 7RU

UK

michelle.sheehan@newcastle.ac.uk

Abbreviations used include: ABL 'ablative', ALL 'allative', ILL 'Illative', INE 'Inessive', TRA 'translative', SF 'singular feminine', CON 'conditional', SM 'singular masculine', SN 'singular neutre'.

[1] Marathi is an Indo-Aryan language spoken by about 60 million people, mainly in the Indian state of Maharashtra.

[2] Bengali and Hindi, two languages closely related to Marathi, and apparently similar to Marathi with respect to the null generic pronoun (exemplified in (1a)), appear to be more permissive than Marathi, BP and Finnish in the case of (2) and (3) , while Icelandic, which also has a null generic pronoun does not allow a null pronoun even in (2).

[3] The term 'semi-pro-drop language' is also in current use, typically applied to languages which only have non-referential null subjects. See Huang (2000: 51ff.) on the typology of null subjects. A further distinction among the semi-pro-drop languages, proposed by Rizzi (1986) and discussed by Huang (2000), is between languages which allow null subjects with weather

59

verbs (quasi-argumental null subjects, for example Icelandic) and those which only allow purely expletive null subjects (German). As it happens, German is not a partial null subject language in our sense, as it does not have a null generic subject. Icelandic is, as it has a null generic subject, although it does not have controlled null subjects in finite clauses; see Sigurðsson & Egerland (this volume). See end of section 2 for a comment on Icelandic.

[4] Expletives are not excluded in principle, though. Finnish employs an expletive subject in certain constructions as an alternative way to satisfy the EPP; see Holmberg & Nikanne (2002).

[5] The counterpart of (7c) and (8c) is not possible in BP with a null subject. Either a clitic or DP like 'the people' is needed in such circumstances:

(i)   *(A gente/se) deve respeitar as opinões   dos jovens.

        the people/SE must respect the opinions of-the young

This is presumably because, unlike Finnish and Marathi, BP cannot have the object DP satisfy the EPP. We will leave aside a discussion of such differences in this paper.

[6] An interesting complication is that (i) is acceptable (see Martins 2006):

(i)   Vi ele sim.

        saw it yes

        'Yes I did.'

Neither Finnish nor Marathi accepts a corresponding construction.

[7] The reason for postulating an inherent, valued D-feature in T in Holmberg (2005) is that it accounts for why a null subject in consistent null-subjuect languages is invariably definite: It acquires the definiteness feature through the agree-relation with (finite) T. In languages without D in T, a null subject is interpreted as indefinite. A persistent problem with the idea that T has

an inherent, valued D-feature (or that INFL has a referential feature, in terms of Rizzi 1982), is that the subject can be an overt indefinite phrase. Indefinite subjects, too, agree with T, so the question is what happens to the inherently valued D-feature in T in that case. This problem does not pertain to the present theory.

[8] Frascarelli (2007) and Frascarelli & Hinterhölzl (2007) take all their examples from spoken corpora, in part because intonation plays an important part in their theory as an independent criterion of topic type. For reasons of space and presentation we do not quote their examples, but instead rely on Samek-Lodovici's constructed examples.

[9] Frascarelli (2007) denies that there is a topic chain: "/.../ this account does not postulate a Topic chain across sentences, but a copying of referential features in different C-domains (through Merge of silent copies), till a new Aboutness-shift Topic is proposed." (fn. 28). This seems like a rhetorical trick; if the null topics are copies, they do thereby form a kind of chain, subject to certain locality conditions, for example.

[10] As noted by Samek-Lodovici (1996), and taken up by Cole (forthcoming), there is variation among consistent null-subject languages regarding the relation between the null subject and the discourse antecedent. Thus Cole shows that some null-subject languages in fact allow a null subject in a situation corresponding to (19). If we assume, with Frascarelli (2007), that a definite null subject (in consistent null-subject languages) must have an antecedent in the local C-domain (a null A-topic), then the variation must concern the interpretation (or 'licensing') of the null A-topic. Some languages are stricter than others when it comes to null A-topics.

[11] Roberts (2007, to appear) does not, in fact, himself apply the mechanism of incorporation by Agree to the analysis of null subjects. His reason for rejecting this analysis is that incorporation of

61

a pronoun into a probing head H by Agree is, in his theory, incompatible with H having an EPP-feature. T has an EPP-feature in the consistent null-subject-languages he is considering, if not universally. This EPP-feature requires movement of the goal to spec,HP. Consequently it is incompatible with incorporation of the goal in H: Incorporation will leave the EPP-feature unchecked. In Roberts (to appear) he argues that the subject pronoun in Italian (which is the consistent null-subject language he focuses on) in fact undergoes movement to specTP, where it is deleted (see Barbosa, to appear, for discussion of the position of the null subject pronoun). We do not accept that the EPP-feature must remain unchecked, in the case where the subject is incorporated. We do accept that the subject cannot be incorporated and at the same time check the EPP. What seems to happen, though, is that the EPP in this case can be checked by some other category. See the text below on the role of EPP in null subject constructions in consistent and partial null-subject-languages.

[12] We ignore the complications required to accommodate uD-feature valueing by an indefinite subject, given that we have said that a referential index means definite interpretation.

[13] See Holmberg (to appear) for more discussion.

[14] We also ignore the precise relation between the null A-topic and T. An interesting possibility is that the null A-topic is a property of the head C, in which case the index-copying would be a consequence of the inheritance of C's features by T, proposed by Chomsky (2008).

[15] It is not entirely obvious why an A-topic in specCP cannot provide a ϕP subject with a D-feature value directly. This may be because the ϕP simply does not have the required uD-feature. It must be linked to such a feature first, to receive a definite interpretation. Consistent

null-subject languages provide this possibility, partial null-subject languages do not. Alternatively it is a matter of 'defective intervention' (see Chomsky 2000) by T.

[16] This is saying that there are no null D-pronouns; if they are null, they are deficient.

[17] Control into indirect questions is possibly less marginal in BP than in Finnish and Marathi.

[18] These judgments are different from the ones reported in Rodrigues (2000). We have checked the judgements carefully, though, and independently, with four speakers of Finnish, all of them linguists. All agree with the judgments reported above, so we are confident that we are right.

[19] The test from 'only NP' has been adapted to Marathi this way because for some reason the focusing adverb *nusta* 'only' scoping narrowly over the matrix subject prevents control of a null subject in the embedded clause.

[20] This is a surprising finding, given the trend in recent literature on control into finite clauses in BP to assimilate it to either OC (Rodrigues 2002, 2004) or to A'-binding (Modesto 2000, to appear). We have checked the data with a range of speakers, though, including Marcello Modesto.

[21] Directive verbs normally take non-finite complements in Marathi, but can somewhat marginally be construed with a finite subjunctive complement. This is also the case in BP, while several directive verbs in Finnish take a finite complement (e.g. *ehdottaa* 'propose', *suositella* 'recommend', *sanoa* 'tell (someone to do something')')