Null subjects in Finnish and the typology of pro-drop

Anders Holmberg

1. Introduction

Subject pro-drop in Finnish finite clauses differs from subject pro-drop in many other well-known pro-drop languages including Arabic, Spanish, and Greek in that (a) it makes a difference between first and second person on the one hand, third person on the other, and (b) it is optional, in a way that is different from the other pro-drop languages. (1) exemplifies the subject agreement paradigm in Finnish. Third person is represented here by the human-referring pronouns hän and he (see Holmberg and Nikanne 2008 on third person pronouns in Finnish).

(1) SG PL

1 (Minä) laula-n. (Me) laula-mme.
I sing-1SG we sing-1PL

2 (Sinä) laula-t.
you.SG sing-2SG you.PL sing-2PL

he/she sing-3SG they sing-3PL

---

1 This research is funded by European Research Council Advanced Grant No. 269752 “Rethinking Comparative Syntax” (ReCoS). Thanks to two anonymous reviewers and the two editors for their constructive critique of a previous version of the paper.
The following examples show that in Finnish a third person subject pronoun cannot be dropped (cannot be null) in a finite clause even though it has an unambiguous antecedent in the preceding sentence, while in Jordanian Arabic it can be, and even has to be, dropped in that context in the unmarked case.²

(2) Onko Leila kotona? Ei, *(hän) lähti ulos. [Finnish]
   Is:Q Leila home no she went:3SG out

(3) Layla bi-l -beet? La, halla (#hiya) tšil'at [Jordanian Arabic]
   Layla in-the -house? no just.now she went:3SG:F
   'Is Layla home? No, she just went out!'

Finnish does allow even third person pro-drop in the right context, though, as in (4), where it is optional.

(4) Leila ei sanonut minne (hän) oli menossa. [Finnish]
   Leila not said where she was going
   ‘Leila didn’t say where she was going.’

Informally, the difference is that the pronoun in (4) has a c-commanding antecedent in the same complex sentence.

² Thanks to Marwan Jarrah for the example. According to Jarrah, the natural interpretation if the pronoun is pronounced is that the speaker resents that fact that she went out.
This paper will present and discuss the properties of null subjects in Finnish, their interpretation and distribution, locating the Finnish system in a typology of null argument systems.

2. **Three pro-drop/null subject types**

Holmberg (2005) proposed a three-way distinction among pro-drop languages: consistent, partial, and radical (or discourse) pro-drop languages. This typology is discussed in more detail in Holmberg, Nayudu, Sheehan (2009), Biberauer, Holmberg, Roberts, Sheehan (2010, esp. chapters 2 and 3), Thampoe (2016), Barbosa (to appear), Frascarelli (to appear). A precursor is Vainikka and Levy (1999) where partial pro-drop was initially discussed. Jordanian Arabic, in (4), is a representative of the consistent pro-drop class of languages, while Finnish would be a partial pro-drop language. In consistent pro-drop languages pro-drop is dependent on agreement, in partial pro-drop languages it is also dependent on agreement at least in part, but in discourse pro-drop languages it is not, and cannot be since most of them have no agreement. These are all distinguished from non-pro-drop languages like English and French, which allow subject pro-drop only under very strict syntactic and stylistic conditions (see Haegeman 2000).

In the following I will first go through the properties that characterise consistent pro-drop languages, then partial pro-drop languages, then discourse pro-drop languages. Thereafter I will put Finnish pro-drop to the test, to see how it behaves with respect to the criterial properties.

2.1 **Consistent pro-drop**

The languages that are called consistent pro-drop languages or consistent null subject languages in the literature have the following properties:
(a) They have subject pro-drop in all persons, all styles, and all clause types, main and embedded, whether the subject is initial in the sentence or not. (3) exemplifies subject drop in a main clause when preceded by a negation particle.

(b) Pro-drop can be dependent on an antecedent (a ‘controller’) in a higher clause, or the antecedent can be extrasentential, as in (3). When there is a controller, it does not need to be strictly local. For instance, in (5), the antecedent of the null subject in the second conjunct can be Gianni, even though Paolo is structurally a closer antecedent.

(5) Gianni₁ non ha detto niente, ma Paolo₂ ha detto che Ø₁/₂ vuole mangiare. [Italian]
'Gianni₁ hasn’t said anything, but Paolo₂ has said that he₁/₂ wants to eat.'

(c) When an embedded subject pronoun is overt, it is typically interpreted as non-coreferential with the subject of the higher clause. If the antecedent is a quantifier, the embedded subject must be null (Montalbetti’s 1986 Overt Pronoun Constraint).

(6) a. Paolo₁ ha detto che lui₁/₁ vuole mangiare. [Italian]
'Paolo has said that he wants to eat.'

b. Cada estudiante₁ dice que él⁴₁₂ tiene poco dinero. [Spanish]
'Every student says that he has little money.'

(d) The subject agreement system is rich, distinguishing at least three persons and two numbers.

(e) There is no null third person inclusive generic pronoun corresponding to English one in One can see the door from here. This function is therefore expressed by a 2SG pronoun,
(which is typically null) or some dedicated generic word, often derived from a word meaning ‘human’, or by a passive, among other means (see Holmberg 2010b).

Examples of consistent pro-drop languages are Arabic, Greek, Italian, Spanish, Polish, Persian. Among the Uralic languages, at least Hungarian, Mari, and Udmurt are consistent pro-drop languages. The Samoyedic languages Enets, Ngasanan and Selkup have pro-drop in all persons (see this volume), and are probably in this class, too.

2.2 Partial pro-drop

(a) Pro-drop may be restricted to some persons or verb forms. In Finnish, as shown, pro-drop occurs freely with first and second person subjects but in the third person only when embedded, with a close enough antecedent. The same is true in Brazilian Portuguese and Hebrew, except that even first and second person pro-drop does not occur in the present tense in Hebrew. (Vainikka and Levy 1999, Shlonsky 2014).

Pro-drop may also be stylistically restricted. In Marathi pro-drop is common in colloquial spoken language, but is more restricted in the written form (Nayudu 2008: 229-230). In Finnish it is the other way around for first and second person pro-drop: It is more characteristic of written than spoken language (see section 3.1).

(b) With third person subjects referential pro-drop requires an antecedent in a higher clause (a ‘controller’), so it is only found in embedded clauses (as exemplified by Finnish (2) and (4)).

(c) Only the closest c-commanding noun phrase can be a controller.

(d) Pro-drop is always optional.

3 Thanks to Eva Dekany, Elena Vedernikova, Yulia Speshilova, and Orsolya Tánczos. See Dalmi (2014) for arguments that Hungarian does not have all the properties expected from a consistent pro-drop language.
(e) The subject agreement system can be deficient.

(f) There is a null third person inclusive generic pronoun.

These properties are exemplified by Finnish in section 3. Examples of partial pro-drop languages are Finnish, Brazilian Portuguese, Hebrew, Marathi (and several other Indo-Aryan languages), and Russian.4

2.3. Discourse pro-drop

Discourse-pro-drop languages, finally, have the following properties:

(a) Pro-drop of subjects and non-subjects, in all persons and tenses.

(b) Pro-drop is not dependent on control; the antecedent can be extrasentential.

(c) Often there is no agreement, but there can be (for instance Tamil is a language with rich agreement, which is nevertheless has the properties of discourse pro-drop languages; Thampoe 2016).

Example languages are Japanese, Korean, Thai, Sinhala, Tamil.

3. Subject pro-drop in Finnish

3.1 Subject pro-drop in root clauses

Pro-drop of first and second person pronouns in root clauses is common in written and formal Finnish, but much less common in colloquial Finnish. It is prescribed in schools and its scarcity in the spoken language is much lamented. The following is a translation from a blog, expressing a sentiment that is fairly typical.

4 On Russian, see Barbosa (to appear), who puts Russian squarely in the partial pro-drop languages category, mainly on the basis of Lindseth (1998). West and South Slavic languages behave more like consistent pro-drop languages, though.
In the last few days I have been worrying about my egocentricity. I was horrified when I read through my previous texts, and I realized that I still use “I”-phrases, representing bad Finnish. Instead of writing “Olen ajatellut” [have.1SG thought - AH], too often I write “Minä olen ajatellut” [I have.1SG thought - AH] emphasizing myself twice in the same sentence. Oh God!

There is one context where pro-drop appears to occur regularly and frequently, that is in answers to yes-no questions.

(7) Question:  Halua-a-ko Leila teetä?  
    want-3SG-Q Leila tea.PRT
    ‘Does Leila want some tea?’

Answer:  Halua-a.  
    want-3SG
    ‘Yes.’

In many languages, including Finnish, a standard form of an affirmative answer to a yes-no question echoes the finite verb of the question (Holmberg 2001, 2016). These answers are not derived by pro-drop, though. One indication of this is the fact that the person of the subject makes no difference. Note that the subject in (7) is third person, which, as mentioned above, cannot otherwise be null in main clauses, in Finnish.

Another indication that something other than pro-drop is going on is that pronouncing the subject in the answer is unpreferred. (8) is not an entirely well-formed answer to (7). Pro-drop in Finnish is never obligatory.
8

Verb-echo answers are derived by ellipsis, but not ellipsis of the subject pronoun as such. Instead the whole sentence is deleted, including the subject, except the finite verb; see Holmberg (2001, 2016) for details. This explains why the person of the subject is irrelevant, and also explains why (8) is not a viable alternative to (6). Verb-echo answers are common among the languages of the world. See Holmberg (2016: ch. 3) on how to distinguish between answers that are derived by pro-drop and answers that are not.

3.2. Subject pro-drop in embedded clauses

In embedded finite clauses null subjects are common, even third person subjects, as long as they have an antecedent (a controller) in the superordinate clause.

(9) a. Jussi₁ sanoo, että (hän₁) aikoo ostaa uuden auton. [Finnish]

Jussi says that he intends buy new car

‘Jussi says that he intends to buy a new car.’

b. Leila₁ muistutti Mattia₂, että Ø₁/₂ oli jo käynyt ullakolla.

Leila reminded Matti:PAR that she/he had already visited attic:ADE

‘Laila reminded Matti that she/he had already been in the attic.’

c. Marjaa₁ harmittaa, ettei (hän₁) voinut tulla aikaisemmin.

Marja:PAR annoys that:NEG she could come earlier

‘It annoys Marja that she couldn’t come earlier.’
d. [Jussin\textsubscript{2} äiti\textsubscript{1}] sanoo, että \(\emptyset\)\textsubscript{1/2} aikoo ostaa uuden auton.

Jussi’s mother says that she/he intends buy new car

‘Jussi’s mother says that she intends to buy a new car.’

Characteristically, the antecedent/controller of the embedded null subject is the subject of the superordinate clause. Example (9b) shows that the controller of an embedded null subject, notated here as \(\emptyset\), can be an object, although in the absence of a context the preferred reading is that it is the subject. (9c) shows that the controlling superordinate subject need not have the same case as the null subject. (9d) shows the role of locality and c-command: the reading in which the embedded null subject is coreferential with the subject NP ‘Jussi’s mother’, rather than the possessive NP ‘Jussi’ is strongly preferred, to the point where the other reading is unavailable. This is a locality effect: the bigger NP c-commands the null subject, and is thereby structurally closer it than the possessor NP is. If spelled out, the embedded subject pronoun can equally well refer to either NP.

(10) [Jussin\textsubscript{2} äiti\textsubscript{1}] sanoo, että hän\textsubscript{1/2} aikoo ostaa uuden auton. [Finnish]

Jussi’s mother says that she/he intends buy new car

‘Jussi’s mother says that she/he intends to buy a new car.’

See section 5 on variation between consistent and partial pro-drop languages as regards pro-drop of embedded subjects.\(^5\)

\(^5\) It was mentioned in section 2.1 that consistent pro-drop languages have obligatory pro-drop of an embedded subject which is a variable bound by a quantifier (as first discussed by Montalbetti 1986).
3.3. **Subject pro-drop in adjunct clauses**

There are indications that subject pro-drop is particularly common and natural in adverbial clauses, among other finite clauses. Compare (11), containing a complement clause, and the clauses under (12), featuring adverbial clauses.

(11) Eeva väittää, että (hän) puhuu ranskaa sujuvasti.  
Eeva claims that she speaks French well

(12) a. Jussi tulee mielellään, jos (hän) voi.  
Jussi comes with pleasure if he can

‘Jussi will gladly come, if he can.’

In Finnish this is generally not the case (likewise in Brazilian Portuguese, another partial pro-drop language; Duarte 2000). In (i) the pronoun can be overt even as a bound variable.

(i) Jokainen ruotsalainen pelkää, että Ø₁/hän₁/₂ joutuu saunaan.  
every Swede fears that he gets.put sauna.ILL

‘Every Swede is afraid that he will end up in the sauna.’

An exception is when the binder is a negative unrestricted quantifier. In this case many speakers find pro-drop to be obligatory.

(ii) Kukaan₁ ei tiedä milloin Ø₁/hän₁/₂ kuolee.  
no-one NEG knows when he dies

‘No-one knows when they will die.’

This is based on a small informant survey carried out by Saara Huhmarniemi. Thanks also to Satu Manninen for data and discussion.
Some evidence comes from an investigation reported in Alsaedi (2016), comparing L2 acquisition of the English subject system by Finnish, French, and Arabic learners. In grammaticality judgment tests the Finnish learners accepted null subjects in English adverbial clauses more than in complement clauses (both are ungrammatical in English), which is plausibly due to interference from Finnish. This tallies with reports from Finnish informants that the null subject is even (mildly) preferred over the overt one in (12a,b), which is not the case in (11).6

(13) contains a relative clause, where the null subject is also highly natural.

(13) Jari rikkoj maljan, jonka (hän) oli saanut lahjaksi sukulaisiltana. [Finnish]
Jari broke vase which he had got present:TRA relatives:ABL:3SG
‘Jari broke a vase that he had received as a present from his relatives.’

Interestingly a similar generalization holds in Hebrew, another partial pro-drop language. Third person null subjects are always well-formed in subjunctive complement clauses in Hebrew, if they have a close enough antecedent. In non-subjunctive contexts third person null subjects are well formed in adverbial and relative clauses, only, but ungrammatical in complement clauses corresponding to (11); see Shlonsky (2014).

6 It is also interesting that the interpretation of embedded third person null subjects in Italian and Finnish is more similar in adverbial clauses than in complement clauses; see Frascarelli (to appear), discussed in section 5.
3.4. **Non-local pro-drop not accepted**

Consider Italian (5) again, repeated here as (15), and compare it with Finnish (16).

(15) Gianni$_1$ non ha detto niente, ma Paolo$_2$ ha detto che $Ø_{1/2}$ vuole mangiare. [Italian]

‘Gianni$_1$ hasn’t said anything, but Paolo$_2$ says that he$_{1/2}$ wants to eat.’

(16) Jussi$_1$ ei ole sanonut mitään mutta Pauli$_2$ sanoo, että $Ø_{*1/2}$ haluaa syödä.

‘Jussi$_1$ hasn’t said anything, but Pauli$_2$ says that he$_{*1/2}$ wants to eat.’

In Italian, in the absence of any context, either reading is equally good: the embedded null subject may have a local or non-local antecedent. In Finnish, a partial pro-drop language, only the local reading is possible, or at least it is highly preferred. See section 5, though, for some more discussion of these data.

3.5. **Non-referential pro-drop**

Expletive subject with extraposition is common in spoken Finnish, but never obligatory.

(17) (Se) oli hyvä, että Jussi tuli auttamaan. [Finnish]

‘It was good that Jussi came to help

‘It was good that Jussi helped out.’

Expletive subject se ‘it’ with weather verbs occurs in spoken Finnish, but is never obligatory and is strictly avoided in written Finnish.
A formal account of this is that Finnish has a null ‘it-type’ expletive. Finnish also has a ‘there-type’ expletive, which is the partitive form *sità* of *se* ‘it’, in impersonal sentences where, typically, the sentence has a subject which does not fill the subject position (Holmberg and Nikanne 2002). This expletive is not optional in the manner of the *se*-expletive, i.e. it does not have a null exponent. The effect can be seen in (19). In (19a), the expletive *sità* fills the subject position, leaving the subject in a lower position. (19b), which has neither a fronted subject nor an expletive, is ungrammatical.\(^7\)


there can horse kick you in:the:head

‘You can get kicked in the head by a horse.’

b. *Voi hevonen potkaista sinua päähän.

can horse kick you in:the:head

3.6. *The inclusive generic null subject (null ‘one’)*

Finnish has a null inclusive generic subject pronoun. ‘Inclusive’ means that the pronoun refers to people in general including the speaker and the addressee. The agreement on the

\(^7\) There is a grammatical reading of (19b), that is when the subject is fronted to the IP-initial position but then the verb is fronted to a sentence-initial focus position (Holmberg 2001, Holmberg and Nikanne 2002).
verb indicates that the null pronoun has 3SG features, like English *one* or German *man*. This pronoun has no overt form.

(20) a. Tässä tuolissa istuu mukavasti. [Finnish]
    This:INE chair:INE sit:3SG comfortably
    ‘One can sit comfortably in this chair.’

    b. Kesällä herää aikaisin.
    Summer:ADE wake:3SG early
    ‘In the summer you wake up early.’

According to Holmberg (2005, 2010a,b), Holmberg, Nayudu and Sheehan (2009), and Holmberg and Sheehan (2010) this is a characteristic of partial pro-drop languages. See Holmberg (2010b) for a summary of arguments that there is a null 3SG pronoun in the syntactic structure which triggers 3SG agreement on the finite verb in these constructions, i.e. they are not subjectless. The subject/topic position needs to be filled with overt material, often a locative or other adverbial, or with the expletive *sitä*. Compare (20a) and (21a,b), showing that the inclusive generic null subject cannot satisfy the EPP in Finnish (see Vainikka 1989, Vainikka and Levy 1999, Holmberg 2010b) (and the expletive *sitä* does not have a null counterpart, as pointed out in section 3.4).

(21) a. *Istuu mukavasti tässä tuolissa. [Finnish]
    sits comfortably this:INE chair:INE

    b. Sitä istuu mukavasti tässä tuolissa.
there sits comfortably this:INE chair:INE
‘One can sit comfortably in this chair.’

4. The formal difference between consistent, partial, and discourse pro-drop languages

The formal nature of null subjects is obviously a highly controversial issue, much debated especially since the early eighties, when the idea of the pro-drop or null subject parameter was introduced (Rizzi 1982, 1986, Chomsky 1982, Huang 1984, Jaeggli and Safir (eds.) 1989), Vainikka and Levy (1999), Y. Huang (2000), Holmberg (2005), Biberauer et al. (2010). This is not the place to elaborate on this issue, but I will, nevertheless, summarize the theory articulated in Holmberg (2010a,b) and Roberts (2010) to explain the differences between consistent and partial pro-drop languages. The general idea is that there is a difference in the agreement system between the two types of pro-drop languages. Consistent pro-drop languages have a D(efinite)-feature in T, along with a set of unvalued $\phi$-features (person, number, and in some languages gender). The unvalued $\phi$-features are assigned a value by the subject, and the subject gets assigned nominative case in the process. This is agreement. The effect of the D-feature is that a null subject will always be interpreted as a definite pronoun. Partial pro-drop languages, too, have subject-verb agreement, i.e. they, too, have unvalued $\phi$-features in T, which are assigned value by the subject, but they do not have a D-feature in T. This means that a third person null subject will not be interpretable as a definite DP unless it is controlled by a definite DP in a higher clause. This explains why the only interpretation a 3rd person null subject can have in a main clause is generic, and why it is dependent on a local controller when it occurs in embedded clauses.8

8 See Barbosa (to appear) for an alternative account of the variation between consistent and partial pro-drop languages. Barbosa notes that the consistent pro-drop languages discussed in the literature
5. A comparison of null subject use in Finnish and Italian

Frascarelli (to appear) reports and discusses the results of an online questionnaire-based investigation testing grammaticality judgments of sentences with null subjects, with 273 Finnish and 128 Italian respondents. The objective is to test whether the facts and the predictions of the theory in Holmberg (2005), Holmberg, Nayudu, and Sheehan (2009), and Biberauer, Holmberg, Roberts, and Sheehan (2010) are right, specifically the distinction between consistent pro-drop languages (represented by Italian) and partial pro-drop languages (represented by Finnish), with focus on third person null subjects.

The overall results show that Italian is more permissive than Finnish as regards the distribution and interpretation of third person null subjects, as we would expect. The difference is not quite as radical, though, as predicted by Holmberg (2005), Holmberg, Nayudu & Sheehan (2009), and Biberauer, Holmberg, Roberts and Sheehan (2010). Consider the following data. The question is how the respondents interpret the embedded null subject: as coreferential with the main clause subject, or as referring to somebody else, or as ambiguous between the two interpretations.

She then develops an idea first put forward by Jayaseelan (1999) and Tomioka (2003), according to which pro-drop may be an effect of noun-deletion (or NP-deletion), as in English I didn’t see a bear, but John saw one, where the noun bear has been deleted in the second conjunct (and one would be a strong form of the stranded indefinite article a/an). If a language has no article, noun deletion would be total, indistinguishable from pro-drop. This, Barbosa (to appear) argues, is essentially how so called pro-drop, of subjects as well as objects, is derived in partial and discourse pro-drop languages.
The theory does not predict any particular preferences in the case of Italian, and that is confirmed by the results. For Finnish, the theory predicts a categorical preference for the coreferential (controlled) reading, as Finnish is supposed not to allow third person referential subject pro-drop except when the subject is locally controlled. A clear majority of the respondents did indeed assign that reading to the null subject. However, as many as 1/3 of them allowed the non-coreferential reading, too, and 14% gave that as their preferred or only reading.

Frascarelli’s (to appear) results indicate that the standard view on third person subject pro-drop in Finnish needs to be modified. Third person subject pro-drop is an option, at least for some speakers, even without a local, c-commanding antecedent, as a more or less marked alternative. The difference between Italian and Finnish is still clear enough, though, evident in the answers to just about every question asked in Frascarelli’s experiment. We may
continue to assume that this difference is a reflection of a more general difference between consistent and partial pro-drop languages.\(^9\)

Frascarelli (to appear) takes her results to show that consistent and partial pro-drop languages all rely on a topic chain to interpret third person null subjects. In particular, they all have a topic feature in the C-domain, which may be abstract, which controls the null argument via the \(\varphi\)-features of T. There is a parameter which determines whether the topic chain can be entirely abstract or whether it needs “(at least) one link of the Topic chain [to] be visible at the interface levels”, i.e. to be pronounced. The latter would be characteristic of partial pro-drop languages.

6. Conclusions and suggestions for future research

The paper has given an overview of null-subject-related properties in Finnish. They are mainly properties which are claimed in the literature to be characteristic of partial pro-drop languages, which is expected as the definition of this type of pro-drop is based in part on the properties of Finnish, in Holmberg (2005), Holmberg, Nayudu, and Sheehan (2009), and Biberauer et al. (2010).

Some recent findings have been discussed, which enrich and complicate the picture. In particular, Frascarelli (to appear) has shown that there is variation regarding the

\(^9\) The option of having a third person null argument interpreted by recourse to a non-c-commanding antecedent, as a marked option, is also found in possessive noun phrases, according to Huhmarniemi and Brattico (2015). In standard Finnish there is a possessive suffix which can stand alone in the first and second person, but needs a spelled out, controlling (c-commanding) antecedent in the third person. Huhmarniemi and Brattico demonstrate that even the third person suffix can sometimes stand alone, in the sense of not having a c-commanding antecedent.
interpretation of null subjects, in that third person null subjects can have a non-controlling, non-local antecedent, at least for many speakers, given the right context.

As for future research, an obviously interesting task is to carry out systematic comparison between Finnish and its relatives, including the very close relative Estonian and the more distantly related but syntactically very similar North Saami, as well as other Saami varieties, and all the other Uralic languages. As always, comparison of closely related varieties has the potential to provide information about variation and universals which cannot easily be had by comparison of unrelated languages; see Kayne (2005).

References


http://essentia.vuodatus.net/blog/category/naurattaisi+jos+ei+vituttaisi