Tarja Riitta Heinonen

Null Subjects in Finnish: from Either-Or to More-Or-Less

1. Background

Finnish is a Null Subject language in the sense that subjects are not obligatorily present in all finite clauses.[1] However, exactly when empty subjects are allowed is not easy to determine. Written standard Finnish resembles Romance languages in that pronoun subjects are avoided, even though only in the first and second person. In contrast, spoken colloquial Finnish favors overt subjects to the extent that they seem obligatory. The picture is further complicated by the existence of subjectless sentence types in all varieties of Finnish.

In this paper, I will address the question of subjectless clauses in Finnish by using attested examples as my data, and the Optimality-theoretic framework to express the conditions on their use. Therefore, I will not contribute to the recent discussion on Null Subjects in terms of Government and Binding theory. While in GB the question is often what kind of languages allow Null Subjects in principle, I will deal with the syntactic and pragmatic conditions under which Null Subjects are actually used. This is not to say that I will present some new information about such conditions. To the contrary, I will make use of the constraints I have come across in the literature, and cast them into a new unified form.

---

[1] I thank two anonymous SKY referees for their comments and encouragement. I am also indebted to Lauri Carlson, Auli Hakulinen, and especially Maria Vilkuna for their criticism and advice. I hope that the result does not prove all their efforts wasted.

The basic idea in Optimality Theory is to allow the grammar to generate all kinds of candidate strings, and then select the best of them as the output. For evaluating a candidate set, a set of ordered constraints is used. (Prince and Smolensky 1993: 4.) For instance, a language may disprefer epenthesis, and syllables without onsets. The best candidate would have neither of them, if possible. If either constraint is to be violated, the one is chosen that is less important in the language in question (p. 25-26). All constraints hold for all languages, but their ranking may vary (p. 5). If a constraint is ranked very low in a certain language, its effects are not necessarily observable.

The present paper has three major sections: First I will illustrate what kind of Null Subject language Finnish is (2). Then I will introduce the conditions I will be using, and assign them Optimality-theoretic interpretations (3). Finally, I will turn to some residual issues and discuss the place my study has in the general setting (4).

One terminological point is in order here: I will use the GB-associated terms Null Subject, empty subject, and (Subject) pro-drop all through the paper. The decision has been made for the sake of convenience, not for any theoretical reasons.

2. Finnish Varieties and Null Subjects

2.1. The Basic Pattern

First, let us see how Colloquial Finnish (CF) differs from Standard Finnish (SF) according to Vainikka (1989). In the following table, potential Null Subjects are in parentheses.

On the basis of this pattern, Vainikka considers CF a non-pro-drop language, but SF a partial pro-drop language (partial because the third person pronouns are not optional). Her main point is to show in which persons and registers pro-drop is licensed, not to pin down the exact contexts in which Finnish speakers actually omit optional subjects.

	Standard Finnish		*Colloquial Finnish*	
Num/Pers	SG	PL	SG	PL
1	(minä) luen	(me) luemme	mä luen	me luetaan
2	(sinä) luet	(te) luette	sä luet	te luette
3	hän lukee	he lukevat	se lukee	ne lukee

Table 1. Subject pro-drop in two main registers according to Vainikka 1989: 185, 188. The verb *lukea* ('to read') in the present tense indicative.

There are many counter-examples to Vainikka's table, however. First, third person finite verbs without subjects actually abound in Finnish data. So called generic sentences or weather verbs provide examples. (See section 2.2. for further information and data.) Second, Null Subjects do occur also in CF and not only in generic sentences or as Null Expletives, but also when the subject is referential. Thus, it seems that finding out which persons allow Null Subjects is just part of the problem since all the persons turn out to allow Null Subjects to some degree. The real issue is to figure out what conditions regulate the actual selection between overt and covert subjects.

Recently, Vainikka has developed her ideas further (Vainikka and Levy 1995). Basically, her main concern is to explain how a mixed system (such as SF) is possible at all. The fact that the Finnish inflectional endings in the first and second person resemble the respective personal pronouns is the explanation for the Null Subjects in these persons (Vainikka and Levy 1995: 11)[2]. Vainikka and Levy admit that they do not know how to interpret the overt subject pronouns that do occur despite the fact that they are optional (p. 9 fn. 5). In typical Null Subject

[2] The CF paradigm is considered less transparent, and thus empty subjects are to be avoided (Vainikka and Levy 1995: 18).

languages, these pronouns are used for a contrastive focus or distinct reading (Rigau 1986: 145). In my mind, such pragmatic factors should not be left aside when determining how pro-drop works in a language.

Vainikka and Levy follow a common practise in trying to reduce the optionality of subjects to transparent verbal inflection. However, to my knowledge, nobody has been able to define what kind of agreement paradigm actually predicts pro-drop. Other factors, such as proper case-government, are often called for as well. (See for example Rohrbacher 1994: 13.)

What seems to be more important than the relative transparency of inflectional endings in Finnish is the fact that subjectless third person forms are used as predicate verbs in so called generic zero sentences (or "missing person sentences", see Hakulinen and Karttunen 1973). These sentences are translated into English with a pronoun *one* or *you* as a subject[3]:

(1) Haminassa **voi** tavata outoja
Hamina+INE can+3SG meet+INF strange+PL+PAR

ihmisiä.
people+PL+PAR

'One/You **can** meet strange people in Hamina.'

Thus, speakers might use the subject pronoun with non-generic third person predicate verbs to make the referential reading explicit. (This observation has been made for example by Maria Vilkuna p.c., and, in a specific context, by Hakulinen 1976: 93.)

If this is true, one would expect that empty subjects do occur with plural verb forms as long as the forms differ from the third person singular. I do not know of any difference between singular and plural in SF, but in CF, there might be. In some

[3] In the examples, the subjectless finite verb is in the bold face. The morphological codes are given in the Appendix 1.

dialects at least, the singular form of the verb is used together with pronoun *ne* ('they'), but if the pronoun is missing, the verb bears the ending for the third person plural (Vilkuna and Laitinen 1993: 36 and references cited there).

2.2. Contexts for Third Person Null Subjects in Finnish

To illustrate the exceptions to the non-pro-drop in the third person, let us have a look at some examples from the data[4].

2.2.1. Generic Zero Sentences

Generic zero sentences are common in all registers, but especially CF makes use of this construction also in case the understood subject is actually the speaker himself:

(2)

Kun	**ei**	oo	omaa	siskoo
when	not+3SG	have+NEG	own+PAR	sister+PAR

nähnyt	ainakaan	neljään	kuukauteen	niin
seen	at least	four+ILL	month+ILL	so

alkaapi	jo	pikkuhiljaa	kaivata	0.
begin+3SG	already	little by little	miss	

'When you haven't seen your own sister for at least four months, you start missing (her)'. [From a personal letter written by a 21-year-old man.]

Notice that in (2), there is also an instance of Object pro-drop in the same clause (marked with 0). Pro-drop is therefore a wider phenomenon than the existence of Null Subjects.

However, generic zero sentences are essentially a different phenomenon from Subject pro-drop: it cannot be claimed that there is a corresponding overt pronoun that has been

---

[4] Appendix 2 gives a list of the data sources used for this study.

omitted in these[5]. The next four subsections will illustrate what real Subject pro-drop cases are like.

2.2.2. Subordinate Clauses

An empty subject may refer back to the subject of the previous clause. Normally this option is limited to coordinated clauses as in English sentence (3), but SF has expanded this usage to subordinate clauses as well (4).

(3) David smiled and **closed** the door behind him.

(4) Mies on kuin punaraitainen pingviini, kun
man is like red-striped penguin when

semmoisen **vetää** päälleen.
such+ACC pull+3SG on+ALL+3SG

'A man looks like a red-striped penguin when he **puts** such a thing on.' [Suomen Kuvalehti 1987/1]

Strictly speaking, not only the subject of the previous clause can serve as an antecedent. Any topic will do. For example, in (5), the noun in the adessive case is the default topic of the construction (of default topics, see Vilkuna 1989[6]), and thus a

---

[5] The pronoun *joku* ('somebody') might be such. In CF, generic you, *sä*, has been adopted from English, but the verb is then in the second person singular.

[6] Vilkuna (1989) makes a terminological distinction between default topics (DT) and topics (T) to keep the first arguments in front of the finite verb, eg. *minulla* in

(i) *Minulla on kylmä*
I+ADE is cold, 'I am cold.'

separate from material that can neutrally occupy the same position, such as *metsässä* in

(ii) *Metsässä satoi.*
woods+INE rained, 'It rained in the woods.'

(i) is an example of DT, (ii) of T. Pre-verbal subjects are then just a sub-group of default topics.

legitimate antecedent for the empty subject. In (6), the antecedent is the object of the passive clause. It takes up the topic position.

(5) No joo, Galbraithilla on tietysti kova tarve
Well yes Galbraith+ADE is of course hard need

päästä julkisuuteen, kun **ei** **ole**
get+INF publicity+ILL when not+3SG have+NEG

moneen monituiseen vuoteen siellä
many+ILL many+ILL year+ILL there

enää ollut.
any more been

'Of course, Galbraith has a great need to get publicity, since he **has not** gotten any for years.' [Suomen Kuvalehti 1987/4]

(6) Häntä pidetään mukavana miehenä, vaikka
He+PAR consider+PASS nice+ESS man+ESS although

on **kotoisin** ruotsinkieliseltä Pohjanmaalta ---
is from Swedish-speaking+ABL Ostrobothnia+ABL

'He is considered a nice man, even though he originally **comes** from Swedish-speaking Ostrobothnia.' [Helsingin Sanomat 6/22/95]

2.2.3 Indirect Requests

Another specific type of referential Null Subject occurs in embedded requests. In the main clause, the NP referring to the one who is asked to do something occurs in the oblique case, as a recipient of a request. Therefore, only a limited number of main clause predicate verbs seem plausible in this construction.

(7) Sano emännälle, että **hakee**
say+IMP+2SG wife+ALL that fetch+3SG

ruusut kellarista.
roses+ACC cellar+ELA

'Say to your(?) wife that she **gets** the roses from the cellar.' [Suomen Kuvalehti 1987/3]

This usage reflects an old-fashioned, polite way of addressing the hearer in the third person singular. A sort of hortative form of this was achieved by omitting the subject. Some native speakers consider even (7) to carry a somewhat out-dated flavor. If there are two potential antecedents in the main clause, as in (8), native speakers judge it to be ambiguous. Either *Pekka* or *emäntä* is the one to take the roses from the cellar.

(8) Pekka sanoo emännälle, että **hakee**
Pekka say+3SG wife+ALL that fetch+3SG

ruusut kellarista.
roses+ACC cellar+ELA

'Pekka says to his wife that he/she **gets** the roses from the cellar.' [Modification of (7)]

2.2.4. Style or Genre

As in English, letters and diaries provide a context in which otherwise obligatory subjects are omitted. The only difference is that in Finnish it is the third person empty subjects that are licensed this way. This kind of Null Subject serves to strengthen the link between the previous and current clause.

(9) Hellevi täytti eilen 50 v.
Hellevi turned yesterday 50 yrs.

Todennäköisesti **on** ollut jossain reissussa.
probably has been somewhere trip+INE

'Hellevi turned 50 years yesterday. Probably she **has** been traveling somewhere.' [A personal letter, a 49-year-old woman]

Unfortunately, rubrics such as 'diary style' or 'letter style' clearly fall short of covering the whole scope of usage. For instance, one can easily find a number of similar occurrences in any issue of

the weekly *Suomen Kuvalehti*. In the present paper, I cannot go further in classifying different subtypes. The following is rather similar to (9):

(10) Kun pyydän luetteloa tulevista esiintymisistä,
when ask+1SG list+PAR forthcoming+ELA performances+ELA

Jorma kysyy, mitä merkitystä sillä
Jorma ask+3SG what+PAR significance+PAR it+ADE

tiedolla on.
piece-of-information+ADE is.

Hakee kuitenkin kalenterinsa.
fetch+3SG however calendar+ACC+3

'When I ask for a list of (his) forthcoming performances, Jorma asks what significance it has. Nevertheless he **brings** his calendar.' [Suomen Kuvalehti 87/49]

2.2.5. Replies

As a final example, let us consider answers to yes-no questions in Finnish.

(11) A: Onks ne jo tullu?
have+Q they already come
B: **On**.
have

'A: Have they come already? B: Yes (they **have**).'

(12) A: Haluavatko he lisää?
want+3PL+Q they more
B: **Eivät**.
no+3PL

'A: Do they want more? B: No (they **don't**).'

Keep in mind that instead of a negative particle, Finnish has a negative verb that inflects in person and number (12). If the subject is present, it follows the verb: *On ne*. Since this order

reads as Contrast-Topic sequence (see Vilkuna 1989 for details), it does not function as a neutral reply any more, but serves to explicitly deny the expectations for the opposite. Usually, some further expansion is required as well — *On ne joo.* ('Yes, they have, yes.'), *Eivät he enää.* ('No, not any more.')

Notice that this last context is not restricted to the third person only.

I do not claim that these five groups (2.2.1—2.2.5) cover all types of third person Null Subjects to be found. Still, the list hopefully gives the reader some idea of the range of the phenomenon. Sections 2.2.2 through 2.2.5 listed contexts in which an empty subject may refer to an antecedent in the previous clause. Replies seem to be a special case: the antecedent is given in the previous turn, and subjects are dropped also because of the word order restrictions.

Hereafter, I will regard the generic zero form (2.2.1) as a separate form in the paradigm. Just like the passive, the generic zero verb has its first argument suppressed, and thus its omission is a different phenomenon from pro-drop.

2.3. Null Subjects in Non-Third Person

I will not go into the details of the non-third person Null Subjects. It is a well-known generalization that non-contrastive first and second person pronouns are not used in SF. In CF, this option is also available. In my recorded material, especially plural pronouns are omitted if the inflectional ending is unambiguous (i.e. *-mme*, *-tte*, *-vAt*), but since their number is very small, no conclusions can be drawn. (13) and (14) provide examples of Null Subjects in the first person, and (15) is a typical case in the second person singular. In (16), the second person plural subject pronoun is missing.

(13)	mul	oli	se	maitopurkki	miäles	mut
	I+ADE	was	it	milk can	mind+INE	but

en oikeen sit tienny
not+1SG really then know+NEG

'I had the milk can in my mind, but I **didn't** really know (whether I should have taken it with me)' [A tape recording, a 60-year old woman]

(14) tota, nyt joten kuten **nään** ilmankii
well, now so and so see+1SG without

lasii hätätilas
glasses+PAR emergency+INE

'Well, I sort of **see** even without glasses if I have to.' [A tape-recording, a 60-year old woman]

(15) [A and B are giving instructions to C:]
A: älä siint kiskase
don't there+ABL pull

B: **työnnät, työnnät** vaa
push+2SG push+2SG only

'A: Don't pull from there! B: You **should** only **push**.'
[A tape-recording, B is a 47-year old man.]

(16) myä aateltii vaa että **voitte** koristella
we thought+PASS/1PL only that can+2PL decorate+INF

jotenkii tät
some way this+PAR

'We just thought that you **can** decorate this (room) some way.'
[A tape recording, a 44-year old woman]

On the basis of my material, it seems that the first person subject may be missing in contexts similar to subordinate clause pro-drop (13). However, it is likely that also several subsidiary factors are at play. The negative verb is one of them (13), a fronted adverbial is another (14). They appear to favor pro-drop, but it is not yet clear to me what their role really is. (15) is an example of the Null Subject in an instruction, a kind of alternative to the imperative form. (16) might be interpreted accordingly. In what follows, only the empty first and second person subjects will be

taken into account that appear in instructions (or requests) and replies. Otherwise, more study is needed.

2.4. Tendencies Rather than Clear-Cut Distinctions

I have already shown that Null Subjects occur in all registers and in all persons. Therefore, what makes Table 1 meaningful at all?

According to Vainikka, CF does not allow pro-drop in the first person singular, but in my 45-minute tape recording of an everyday conversation, there were as many as 23 instances of this (only clear cases included). 5 of them were replies; the rest must be motivated otherwise. On the other hand, the total number of first person singular verbs rose as high as 197. Thus, only 1 out of 10 subjects were empty. By comparison, in 18 personal letters, the number of empty subjects in the first person singular was 146 out of the total 160 finite verbs. See Table 2 below.

Even though personal letters are not as formal as newspaper texts, Subject pro-drop is clearly favored. When an overt subject occurs, it is usually contrastive.

However, in CF, the percentage of phonologically reduced forms of the pronoun *minä* is overwhelming. It is likely that the pronoun is turning into a clitic in CF - in fact, there are plenty of cases in which the pronoun is blended with an auxiliary or a conjunction: *emmä* ('I don't'), *mioon* ('I am', used in the dialect where my examples are from), *kummä* ('when I'). Therefore, most overt subjects in CF are essentially of a different nature than those in SF.

	Conversation/CF	*Letters/SF*
Null Subj	23/197	146/160
minä	5	10
mä	169	0
itse (alone)	0	4

Table 2. Null Subjects in a colloquial conversation and 18 personal letters. *mä* stands for all reduced forms of *minä* ('I'). *itse* ('self') is used sometimes instead of *minä* in SF. About the data, see Appendix 2.

3. Constraints on Null Subjects

In this section I will introduce a collection of explanations suggested by other authors for why Null Subjects may or may not occur in certain contexts. I will also refer back to some explanations we have already encountered. Since one of the basic assumptions of Optimality Theory is that all constraints are shared by all languages, any piece of evidence from any language should be relevant when determing what are the factors affecting pro-drop.

(17) Conditions against Null Subjects:

(a) Null Subject leads to ambiguity.

> [In Portuguese,] subject pronouns are duplicated by verb inflection -- and are frequently omitted, especially in the unambiguous first and second person forms. Third person forms are more ambiguous. The use of third person grammatical forms as the main form of address restricts the omission of pronouns to clear cases of anaphora or address. Otherwise, subjectless third person verbs are interpreted as having indefinite subjects: *é horrível* 'it's terrible', *dizem que é proibido* 'they (people) say that it is forbidden'.
>
> Parkinson 1988: 160

(b) Null Subject Fails to Fill the Topical Position/Contrast the Previous Item

> [In Finnish, there is] some kind of surface constraint or condition on the canonical form of a declarative sentence that it start with a nominal rather than a verbal element, to distinguish it from other sentence types.
>
> Hakulinen 1976: 143-144

> the Finnish discourse configuration: K T V-field
>
> Vilkuna 1989: 37
>
> A non-imperative sentence should have a T [Topic/Theme] if possible.
>
> Vilkuna 1989: 40
>
> Another [other than marking K =Contrast position] example of the centrality of T is the existence of T dummies in colloquial spoken language
>
> Vilkuna 1989: 41

(c) Informal Register

> As in Chinese, a Hebrew speaker should have no reason to use an overt pronoun when coreference is intended, for zero is usually allowed. However, overt coreferent pronouns occur quite often, their popularity depending on the genre (much more so in spoken than in written discourse).
>
> Ariel 1994: 13

(18) Conditions that favor Null Subjects:

(a) Avoid Pronoun/ Avoid Distinct Reference.

> (4i) John would much prefer his going to the movie
> (4ii) John would much prefer his (own) book
> ---
> Thus in (4i), where PRO [=Null Subject] may appear, the overt pronoun is taken as distinct in reference from John; but in (4ii), where PRO may not appear, the overt pronoun is free in reference. Principle (5) [Avoid Pronoun] might be regarded as a subcase of a conversational principle of not saying more than is required, or might be related to a principle of deletion-up-to-recoverability, but there is some reason to believe that it functions as a principle of grammar.
>
> Chomsky 1988 [1981]: 65

> We can restate the Avoid Pronoun Principle as (i):
> (i) Avoid full pronoun.
>
> Rigau 1986: 161

(b) Avoid Unintended Emphasis and Do Not Contrast the previous item. See examples of the latter in section 2.2.5, Replies to yes-no questions.

(c) Make a cohesive link to the previous clause.

> Li & Thompson (1979) quote the following example from Chinese:
>
> (13) (a) This *Wang-Mian* was gifted.
> (b) 0 (=he) was not more than twenty years of age.
> (c) 0 (=he) had already mastered everything in astronomy, geography, and classics.
> (d) However, *he* had a different personality.
> (e) Not only did 0 (=he) not seek officialdom...
>
> Note that in Chinese, references to highly accessible entities, as 'Wang-Mian' above must be, are preferably made by zeros. This happens in three of the four non-initial references above, and a pronoun in any of them would have indeed favored a disjoint reading. -- Li & Thompson, however, suggest that clause (in)dependency is not a syntactic matter only. It is also pragmatic in nature, and pragmatic cohesion plays a role in anaphoric interpretations.
>
> Ariel 1994: 13

(d) Give an order or instruction. (See examples 7, 8, 15 above.)

Most of the above conditions could be phrased in two ways, either as conditions for empty subjects or against them. For example, the converse of (18c) is used by Ariel to motivate some occurrences of overt full pronouns at points where there is a 'drastic' break 'from the story line' (Ariel 1994: 15), or when a pronoun introduces 'a shift from the previous unit' or 'an aside' (p. 14-15).

Therefore, to account the pro-drop pattern in Finnish, we could either start from the assumption that Null Subjects are exceptional and need to be motivated (by 18a-d); or that Null Subjects are the normal case, and overt subjects need permission to occur (17a-c). I will adopt the first strategy. Thus, I assume that basically Finnish is a non-pro-drop language, and that all

attested counter-examples result from the conditions for Null Subjects.

3.1. All Languages as Non-Pro-Drop Languages

All languages can be claimed to be non-pro-drop by default. All overt exceptions must then be explained. Hopefully, the explanations form a coherent system by themselves so that the analysis does not reduce to a patchwork.

In Optimality-theoretic terms, non-pro-drop is a constraint that may or may not be violated in a language. If it is never violated, i.e. there must be a subject in all grammatical sentences of the language, the constraint dominates all other relevant constraints. If there are violations, there must be one or more conditions that dominate the non-pro-drop constraint, and conflict with it.

Let us have a simple example: Finnish allows Null Subjects in a clear violation against the non-pro-drop constraint. Thus, according to the theory, there must exist another constraint that prefers Null Subjects and dominates the non-pro-drop constraint. Above, we had a list of candidates for this purpose, (18a) through (18d). We pick up the constraint Avoid Pronoun[7] (hereafter NoPron), and require that it dominate the constraint against Null Subjects (hereafter NoNullSubj):

(19) NoPron > NoNullSubj

What happens is that we get rid of all overt personal pronouns, clearly too dramatic a result. In the following sections, I will show how this elimination can be restricted to smaller groups. Also, we will see how the constraints handle the difference

---

[7] If we take it that a Null Subject is always an option in all contexts in Finnish, Chomsky's principle "Avoid Pronoun" applies everywhere. In the next section, I will return to this point.

between the third and non-third person pronouns shown in Tables 1 and 2.

3.2. The Third and Non-Third Person Null Subjects

In actual fact, we already have a constraint to make a distinction between the third and non-third person pronouns. The Avoid Pronoun principle, NoPron, applies to optional pronouns only. If we interpret all personal pronoun types as optional, in accordance with my data, all overt pronouns are discarded by NoPron. If, however, we consider optionality of tokens instead of types, the third person pronouns probably turn out to be optional less often than the others. The first and second person pronouns are commonly considered unambiguous in all contexts (Vainikka and Levy 1995: 3), so they are always optional in this sense. Nevertheless, if the optionality were the only criterion, as it is in Italian, Finnish should drop even third person pronouns more often than it in fact does. In reality, for any Null Subject sentence resembling (9) or (10) there exist dozens of sentences with overt subjects.

What appears to be a Finnish verb in the third person singular is actually ambiguous between the generic zero reading and the referential reading, and earlier I already suggested that the former is a distinct form in the paradigm. There are some specific sentential and lexical properties which generic zero verbs typically share. These include: (a) these verbs are often used in *if—then* constructions, (b) these verbs are often modals, and (c) a generic zero clause does not begin with the verb itself (see Hakulinen and Karttunen 1973: 160ff for more information). Conversely, let us hypothesize that referential Null Subjects are used in a way that avoids these features. Thus, a third person Null Subject should be allowed whenever the expression does not include any of the features associated with generic zero clauses. Again, the conclusion is that Null Subjects should be more common in Finnish than they actually are. Moreover, if pro-drop

in the third person were the norm, we would not expect it to be stylistically rather marked as it is especially in the groups 2.2.3 and 2.2.4 above.

A third explanation for the asymmetrical Null Subject distribution could be the Politeness Principle: avoid face-threatening acts, such as a direct reference to yourself or to the addressee is in Finnish (Shore 1986: 52-53)[8]. A common strategy adopted by Finnish speakers is to use pseudo-generic, or pseudo-passive, expressions (see 2.2.1 and Shore, ibid.). Pro-drop in the first and second person may also be a variation of the same strategy. At least when in doubt which form to use when addressing another person, it is best to avoid the whole reference, but the second best might be to avoid the use of personal pronouns.[9] Since the evidence for this principle is not conclusive, I will not adopt it in this paper.

I will leave the question about optional third person pronouns (such as pro-drop in subordinate clauses) aside for a moment, and suggest the following constraints as responsible for the difference between the third and non-third persons:

(20) NoPron (Rigau's modification): a full optional pronoun is to be discarded.

---

[8] The most recent analysis (Laitinen, forthcoming) of generic zero sentences emphasizes the positive interpretation of the same strategy: it provides a convenient way for a speaker to generalize what she is saying, and allows the hearer to identify herself with the experiencer-speaker. See for example (2).

[9] A problem with this explanation is that when the pronoun is in the subject position, the verb necessarily agrees with it, and thus reveals whether the speaker dropped *sinä* or *te* (Fr. 'tu' or 'vous'). It is also difficult to say whether the possible impoliteness of repetitive use of (full) first and second person pronouns would result from the fact that they are contrasted (and thus lead to all kinds of pragmatic implications), or whether the mere fact that you talk too much about yourself or pay too close attention to what somebody else is doing, is considered impolite as such, regardless of how you literally express it.

(21) NoAmb: the referential verb form should be disambiguated in the context, for instance with the help of a subject pronoun.

(20) and (21) represent the first two explanations given above. The idea is that their combined force would be enough to predict the difference between the persons: (20) favors pro-drop in the non-third, and (21) disprefers pro-drop in the third person. (21) stands also for (17a). To accommodate reduced (= opposite to full) CF pronouns in my analysis (see 2.4 and below), I will use Rigau's version of NoPron from now on.
It is possible to see NoAmb as a subcondition for NoPron to apply: if the verb form is ambiguous, the pronoun is not optional. That way NoPron would be the only thing we needed.

The ordering of the constraints is:

(22) NoAmb > NoPron > NoNullSubj

In the following table, the constraints are ranked in descending order, some output candidates are judged against them, and all violations are marked with an asterisk. The winner is the analysis that violates the lowest constraint if any. I have marked such surviving candidates with a ✓.

	NoAmb	NoPron	NoNullSubj
menee	*		*
hän menee ✓		*	
menen ✓			*
minä menen		*	

Table 3. The first and third person pro-drop compared.

This constraint table explains why Null Subjects appear to be the unmarked choice for the first and second person, while in the third person just the opposite is the case. The explanation is that overt subjects are always unmarked, but that this effect gets undone by another condition higher in the constraint hierarchy. In passing, notice that the spoken Finnish expression *mä menen*, with a reduced pronoun *mä*, would be the best candidate for the first person under the same constraints:

	NoAmb	NoPron	NoNullSubj
mä menen ✓			

Table 4. The CF expression for 'I go' does not violate any constraint since the verb form is unambiguous and no full pronoun is involved.

Thus the fact that Finnish speakers do not use *mä menen* in SF should be prevented by a genre-specific constraint (the converse of 17c).

3.3. Pragmatic Constraints and Variation

Since third person Null Subjects do occur, there must be a constraint dominating NoAmb, or the conditions for NoAmb are to be extended. I suggest that we do both.

Previously, I only mentioned the option to disambiguate the third person verb with an overt subject (21). To license subordinate pro-drop, other possibilities must be included. Such a legitimate context consists of an antecedent that occurs in the previous main clause as a subject.

(23) NoAmb (modified): the referential verb form should be disambiguated in the context, with the help of a subject pronoun or a DT antecedent in the dominating main clause.

On the other hand, it does not seem reasonable to lump these syntactic conditions with pragmatic ones. Thus I assume that at an upper level, there is a constraint which abandons overt pronominal subjects if used in replies and commands. Notice that this constraint is probably too restrictive to be a universal constraint.

(24) Reply/Command: do not use an overt pronoun in replies and commands.

(25) Reply/Command > NoAmb (modified) > NoPron ...

However, there is still a problem with variation: subordinate pro-drop is possibly genuinely optional, and NoPron should in turn delete all such optional subjects. Nevertheless, (27) is as grammatical as (26).

(26) Liisa vaihtaa aina vaatteet jos **on**
Liisa change+3SG always clothes if is

menossa ulos poikaystävänsä kanssa.
going+INE out boyfriend+ACC+3 with

(27) Liisa vaihtaa aina vaatteet jos hän on menossa ulos poikaystävänsä kanssa.

'Liisa always changes clothes if she **is** going out with her boyfriend.'

What could be the status of such a leaking constraint? In Optimality Theory, one can rank the constraints equally high so that neither dominates. In our case, the essential conflicting constraint seems to be hidden under the conditions of the constraint NoPron itself. Therefore, it should be given an independent status and ranked on its own right. Something along the line of (18c) is needed here.

Admittedly, equal ranking may be just first aid. Perhaps the alternatives turn out to differ in some categorical way which we have not succeeded in finding out. On the other hand,

variation is an inherent part of any natural language, and we need to have tools to express this.[10]

The mere syntactic information is not always enough to specify which reading is the intended one. In the present paper, I have assumed the existence of such situational constraints as Reply/Command. The question is: when does such a constraint apply? (28) might be 'imperative' or 'anaphoric' depending on the situation.

(28) Tulee vaan lähemmäksi.
come+3SG just closer+TRA

'Please come closer.'
or
'(It didn't disappear anywhere.) It just comes closer.'

What we need is more information about the context. In the following example, the previous turn is given. Since it is a question, an answer is what we expect next. Thus, constraints on appropriate replies apply.

(29) Onko vauva jo nukkumassa? - On.
is baby already sleeping+INE is
'Is the baby already in bed? - Yes, it is.'

[10] See Anttila 1995 on how a partially ordered grammar is even able to predict the proportional occurrences of variants. The idea is that a partially ordered constraints produce several competing constraint tableaux.

	Reply	NoAmb	NoPron	Parse
se on	**		*	
<se> on	*	*		*
on se	*		*	
on <se> ✓		*		*

Table 5. An answer to a yes-no question. NoAmb as in (23). < > marks the unparsed (= deleted) part in a candidate.

Notice that there are a number of constraints typically used in Optimality-theoretic constraint tableaux. What I have so far called NoNullSubj is in fact a special case of constraint Parse. It says that everything in the input should be represented in the appropriate analysis. In essence, it is a constraint against any kind of deletion. The assumption is that the subject is there underlyingly, even if it does not surface because of the Parse violation.

The Reply constraint actually combines two aspects: whether the word order is that of an answer to a yes-no question, and whether the answer is of such a form that requires extra presuppositions to be held in the context. This is why the same constraint can be violated twice (see *se on*). We see that higher-ranked Reply kills both candidates with overt subjects. The only candidate to survive the comparison is the one with an unparsed subject and appropriate word order.

Up to this point we have touched on almost all the conditions in (17) and (18). Next we will turn to the remaining two conditions (17b) and (18b).

3.4. Finnish as a Topic-Prominent Language

So far I have ignored the fact that Finnish is not regarded as a subject-prominent language. Finnish avoids verb-initial sentences

(condition 17b above), but almost any nominal constituent can assume this crucial clause-initial position. This is why weather verbs often take an adverbial in front of them (31). The same applies to generic zero sentences (32).

(31) Tänään / Philadelphiassa **satoi** vähän.
today / Philadelphia+INE rained a little
'It **was raining** for a short while today / in Philadelphia.'

(32) Mies tajusi, että autoon **pääsi** sisään helposti.
man realized+3SG that car+ILL got+3SG in easily
'The man realized that one **could get** into the car easily.'

Finnish even uses expletive topics. In priciple, expletive topics are used just like the referential ones. For instance, a topic helps to mark the preceding item as contrastive (17b). These expletives can often be found with weather verbs, and in the passive and generic zero sentences.

(33) No nyt se sataa!
well now it rain+3SG
'Now it's raining!'

Even though expletive topics are mostly used in the constructions that lack a subject themselves, this is not a necessary condition. See (34).

(34) Oltermanni, höylä ja hyvät ruokahalut.
Oltermanni, slicer and good+PL appetite+PL

Mitä sitä suomalainen muuta tarvitsee?
what it+PAR Finn else+PAR need+3SG

'Oltermanni (cheese), slicer and good appetite. What else could a Finn need?' [An advertisement in Helsingin Sanomat 4/26/95]

Expletive topics provide valuable evidence for the constraint in force. Holmberg and Nikanne (1994) even propose that in existential clauses, such an expletive topic (they call it a subject) is obligatory if nothing else fills the position. Their example is repeated here as (35):

(35)	*(Sitä)	leikkii	lapsia	kadulla.
	it+PAR	play+3SG	children+PAR	street+ADE

'There are children playing on the street.'

In essence, they claim that *sitä* corresponds to English *there*, and that the sentence is ungrammatical without it, provided that the predicate verb is not focused and that the adverbial stays where it is (Holmberg and Nikanne 1994: 173, 177). Their example is, however, unfortunate in that most native speakers would regard it as unacceptable[11]. Luckily, Holmberg and Nikanne bring up other ideas as well, two of which I consider worth developing. One is the obligatoriness of subjects in certain constructions, the other is the observation that *sitä* ('it' SG PAR) occurs with verbs that do not have a normal nominative subject, *se* ('it' SG NOM) elsewhere. The former idea I have already dealt with in this paper

---

[11] For some reason, Holmberg and Nikanne ignore the fact that *sitä* is not a pragmatically neutral filler in the way *there* is in English. (The standard analysis of *sitä* can be found in Hakulinen 1975.)

Moreover, Holmberg and Nikanne do not (except for endnote 3 on p. 186) give Right Dislocations their due share as an extremely typical strategy to avoid empty topics in CF. Compare the following:

(i)	Tutti!	Tuos	on	tutti	tuos	hyllyn
	pacifier	that+INE	is	pacifier	that+INE	shelf+GEN

päällä.
over
'Pacifier! There is a pacifier on that shelf.'
[An unrecorded conversation 4/17/95, a 49-year-old woman.]

One would not say:

(ii) Tutti! Sitä on tutti tuos hyllyn päällä.

Sitä is not probably at home in a context in which something new and interesting is introduced.

under the label NoNullSubj. The latter observation conforms nicely with my idea that ambiguity must be avoided (NoAmb). Compare the following examples:

(36) Se väsyy helposti.
it get-tired+3SG easily
'It/she/he gets tired easily.'

(37) Sitä väsyy helposti.
it+PAR get-tired+3SG easily
'One/I get(s) tired easily.'

It is obvious why *sitä* is a better choice than *se* for a dummy topic in (37).

It seems that NoNullSubj (= Parse constraint against Null Subjects) is actually too narrow a constraint in Finnish. Something along the lines of (17b) is needed instead, i.e. a constraint against Null Topics. I will turn to this point in section 3.5. below.

3.5. Unparsing or Epenthesis?

In Optimality Theory, two general repair strategies are made use of: Unparsing (or Deletion) and Epenthesis. In the former, an element in the input[12] is left out of the analysis (= candidate); in the latter, an empty element is introduced into the analysis. An empty epenthetic "box" is filled with default material for the position. Since this sounds exactly what expletive topics (subjects for Holmberg and Nikanne) are about, we are faced with the problem that sometimes we delete, sometimes add topical constituents, and all this would be best handled within one theory of pro-drop.

---

[12] Optimality Theory retains the old distinction between underlying and surface representations, the input corresponding to the former, the output (or the best candidate) to the latter. The candidates preserve the lexical projections of the input (Grimshaw 1993: 40).

In 2.2.1, I mentioned that pro-drop in Finnish affects not only subjects but objects as well. But instead of formulating a specific constraint against NullObjects, it seems reasonable to look for a more general solution. Otherwise, we end up listing constraints against any null elements, one by one. The appropriate generalization appears to be that the verb (or any other head) requires that its arguments be overtly represented. So far, we have only paid attention to the first argument, whether it is called subject in a traditional grammar or not. In the previous section, I preferred the term "topic" for any element that occupies the preverbal position. If there is no default topic (see 2.2.2 fn. 6 for the term), epenthesis can give rise to expletive elements:

(37) **Sitä** ei enää vaa viitti.
it+PAR not+3SG any more just bother+NEG
'One just doesn't bother (to do something/anything) any more.'

In this case, a violation against NoEpenthesis is ranked lower than the constraint against a missing topic. Naturally, to block (38) and to choose (39) instead, a constraint that prefers movement to epenthesis is to be added to the grammar.

(38) * **Se** satoi metsässä.
it rain+3SG woods+INE
'It was raining in the woods.'

(39) Metsässä satoi.

One obvious way of achieving this is to assume a scale of topicality. The first argument of the finite verb heads the list, and last comes the epenthetic expletive. No movement rule is referred to. In other words, movements are to be understood metaphorically in OT. Admittedly, some discourse factors prevent the first arguments from being the best topics in all sentences. Thus Topic=1ARG turns out to be a violable constraint itself (42).

Since the generic zero verb form has its first argument lexically suppressed, an expletive occurring with it must be epenthetic. The following pair of sentences (taken from Hakulinen 1976: 93) shows the clear difference between such generic verb forms and regular verbs with unparsed subjects.

(40) Oppilas tietää, ettei tehtävää
pupil know+3SG that-not+3SG exercise+PAR

pysty ratkaisemaan.
be-able+NEG solve+INF

'The student knows that no one **is able** to solve the problem.'

(41) Oppilas tietää, ettei <hän> **pysty**
pupil know+3SG that-not+3SG be-able+NEG

tehtävää ratkaisemaan.
exercise+PAR solve+INF

'The student knows that he/she **is not able** to solve the problem.'

In (41), the object *tehtävää* occurs after the main verb, while in (40), it has supposedly moved before it. I see this as an argument for an unparsed subject in (41). Since the topic position is filled with <hän> in the best analysis, *tehtävää* does not move to prevent the violation against a null topic. In (40), the input does not contain any topic at all, thus the position is filled out otherwise. Notice that after modifying the assumptions about the input so that all verbs do not have the first (topic or subject) argument, Parse is not a strong enough constraint against Null Subjects. This is why we need a special constraint NoNullTopic to guarantee that expletive topics are favored over Null Topics.

The new additions to our set of constraints, and the hierarchy for the set of constraints considered so far are as follows:

(42) Topic=ARG1: the first argument occupies the topic position.

(43) NoNullTopic: there must be an overt topic in a clause.

(44) Reply/Command
> NoAmb
> NoPron
> Topic=ARG1
> NoNullTopic
> Parse
> NoEpenthesis

Certainly, this hierarchy is still only a partial set of the constraints needed. For instance, the constraint against full first and second pronouns must be dominated by a constraint that requires contrasted material to be present. This takes the form of a special kind of Parse constraint.

(45) *<EMPH>: if there is emphatic material, it should be overtly present.

*<EMPH> dominates not only NoPron, but also the other constraints in (44).

4. Concluding Remarks

I have shown that different factors regulating the Null Subjects in Finnish can be combined into one set of constraints. Syntactic and pragmatic conditions can be meaningfully combined into one grammar.

I have adopted an approach which exploits only negative constraints, i.e. which excludes ungrammatical or otherwise unacceptable strings whenever they violate an important enough constraint. In this respect, it differs from Fred Karlsson's Constraint Grammar in which the contextual constraints may either discard or select a specific analysis (Karlsson 1990).

In Optimality Theory, constraints are merely ordered. One higher-ranked violation is absolutely worse than however many lower-ranked ones. This makes is different from, for instance, Harmonic Grammar (Legendre et al. 1991).
Optimality Theory has been tried in numerous phonological and some morphological problems, but, to my knowledge, syntax and

pragmatics have been rare targets[13]. As exceptions I can name Grimshaw (1993) and Legendre's attempt to cover wh-movements in several languages using ordered GB-based constraints (Legendre 1994).

In the present paper, Null Subjects are considered unparsed segments in the input analyses. Together with some pragmatic constraints, the Parse constraint does its best to deny empty subjects their right to exist. On the other hand, Parse cannot be the source of the expletive subjects. The constraint against clauses without topics produces them via epenthesis. To complete the picture, several minor constraints were introduced, and a tentative hierarchy suggested.

Since pro-drop itself is gradient and divides into smaller phenomena, I do not see any reason to be content with a parameter-type of solution to pro-drop. At first blush, Finnish may look like an exceptional language in that it cannot be claimed to be either pro-drop or non-pro-drop (Vainikka and Levy 1995). However, at closer examination, Finnish appears to be revealing. The apparent pro-drop and non-pro-drop languages seem to be results of different constraint rankings. For example, in languages that are traditionally called non-pro-drop, Parse dominates NoPron, while in Italian, the reverse must be the case.

Some residual problems still call our attention. A case in point is the universality of the constraints. For instance, the constraint NoAmb with its subconditions is clearly only the first step towards a mature part of a constraint grammar. Similar vagueness seems to surround the constraint NoPron. Since these constraints are crucial in a proper analysis of pro-drop in Romance languages, one cannot take them as such black boxes as I have done for the sake of simplicity.

---

[13] For a list of available Optimality-theoretic papers, consult the Rutgers Optimality Archive at ruccs.rutgers.edu/pub/OT/TEXTS.

Appendix 1: Form Glosses

ABL - ablative; ACC - accusative; ADE - adessive; ALL - allative; ELA - elative; ESS - essive; GEN - genitive; ILL - illative; IMP - imperative; INE - inessive; INF - infinitive; NEG - negative; PAR - partitive; PASS - passive; PL - plural; TRA - translative; 1SG - first singular; 2SG - second singular; 3SG - third singular; 1PL - first plural; 2PL - second plural; 3PL - third plural; 3 - third person possessive suffix

Appendix 2: Example Sources

Helsingin Sanomat. A daily newspaper.
Suomen Kuvalehti. A weekly. The year 1987 issues are available on-line at the Department of Linguistics, Helsinki.
A tape-recording. 26.11.1989. A family gathering with six people from South-East Finland.
Notes on unrecorded material.
Personal letters. 18 letters written in 1993-94.

References

Anttila, Arto (1995) *Deriving variation from grammar: A study of Finnish genitives*. Manuscript. Stanford University.

Ariel, Mira (1994) Interpreting anaphoric expressions: a cognitive versus a pragmatic approach. *Journal of Linguistics* 30: 3-42.

Chomsky, Noam (1988) [1981] *Lectures on Government and Binding: the Pisa lectures*. Dordrecht: Foris Publications.

Grimshaw, Jane (1993) *Minimal Projection Heads & Optimality*. Technical Report #4. Rutgers University Center for Cognitive Science.

Hakulinen, Auli (1975) Suomen *sitä*: pragmatiikan heijastuma syntaksissa. *Sananjalka* 17: 25 - 41.

Hakulinen, Auli and Lauri Karttunen (1973) Missing Persons: On Generic Sentences in Finnish. *Papers from the 9th regional meeting of the Chicago Linguistic Society* pp. 157 - 171. Chicago.

Holmberg, Anders and Urpo Nikanne (1994) Expletives and Subject Positions in Finnish. *Proceedings of the North East Linguistic Society* 24: 173 - 187. Amherst: University of Massachusetts.

Karlsson, Fred 1990 Constraint Grammar as a Framework for Parsing Running Text. In H. Karlgren (ed.), *Papers presented to the 13th International Conference on Computational Linguistics, Vol.3* pp. 168-173. Helsinki.

Laitinen, Lea (forthcoming) Nollapersoona. *Virittäjä*.

Laitinen, Lea and Maria Vilkuna (1993) Case-Marking in Necessive Constructions and Split Intransitivity. In A. Holmberg and U. Nikanne (eds.), *Case and Other Functional Categories in Finnish Syntax* pp. 23 - 48. New York: Mouton de Gruyter.

Legendre, Géraldine (1994) *Optimality in Operator Chains*. Handout for a talk given at the University of Delaware.
Legendre, Géraldine, Yoshiro Miyata and Paul Smolensky (1991) Unifying syntactic and semantic approaches to unaccusativity: A connectionist approach. *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*, pp. 156-167. Berkeley.
Parkinson, Stephen (1988) Portuguese. In M. Harris and N. Vincent (eds.), *The Romance Languages*, pp. 131-169. London: Routledge.
Prince, Alan and Paul Smolensky (1993) *Optimality Theory: constraint interaction in generative grammar*. Manuscript. Rutgers University and University of Colorado.
Rigau, Gemma (1986) Some remarks on the nature of strong pronouns in null-subject languages. In I. Bordelois, H. Contreras and K. Zagona (eds.), *Generative Studies in Spanish Syntax* pp. 143 - 163. Dordrecht: Foris Publications.
Rohrbacher, Bernhard (1994) *Explaining the Syntactic Consequences of 'Rich' Agreement Morphology: On the Licensing of V-to-AgrS Raising and pro*. Manuscript. University of Pennsylvania.
Shore, Susanna (1986) *Onko suomessa passiivia*. Suomi 133. Helsinki: Suomalaisen Kirjallisuuden Seura.
Vainikka, Anne (1989) *Deriving Syntactic Representations in Finnish*. Ph.D. Dissertation. University of Massachusetts.
Vainikka, Anne and Yonata Levy (1995) *Empty Subjects in Finnish and Hebrew*. Manuscript. University of Pennsylvania and The Hebrew University.
Vilkuna, Maria (1989) *Free Word Order in Finnish*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Tarja Riitta Heinonen
Research Institute for the Languages of Finland
Sörnäisten rantatie 25
00500 Helsinki.
E-mail: Tarja.Heinonen@domlang.fi