Institute for Research in Cognitive Science

Empty Subjects in Finnish and Hebrew

Anne Vainikka
Yonota Levy

University of Pennsylvania
3401 Walnut Street, Suite 400C
Philadelphia, PA 19104-6228

December 1995

Site of the NSF Science and Technology Center for Research in Cognitive Science

IRCS Report 95-31

Empty Subjects in Finnish and Hebrew*

Anne Vainikka, University of Pennsylvania
Yonata Levy, The Hebrew University

November 1995

1 Introduction

Under the traditional definitions of *pro-drop* and *non-pro-drop* languages, Hebrew and Finnish are mixed languages. In these languages, subject NPs can be omitted in certain persons or tenses, but not in others. The analysis provided in the present paper to account for these unusual null subject patterns is based on the idea that the *syntactic position* of subject-verb agreement features varies cross-linguistically, and even within a language. The contribution of this paper is to provide an analysis of null subjects which (i) covers the complicated mixed systems of Hebrew and Finnish (Sections 3-5), (ii) extends to the systems of null subjects traditionally discussed in the literature (Section 6) and (iii) suggests a way of characterizing the interaction between syntactic and discourse factors in the context of NP subject omission (Sections 3 and 7). The analysis of null subject distribution in Hebrew and Finnish takes into account a morphological generalization, namely the fact that the agreement suffixes in both languages are *phonologically* related to the corresponding pronouns in exactly those persons where subject omission is possible.

In Hebrew past and future tenses – which are inflected for tense, number, person and gender – and in all tenses of Finnish,[1] first and second person subjects are optional; in the third person, however, an overt subject is required, as shown in examples (1a,b) for Finnish (Vainikka 1989) and in examples (2a,b) for Hebrew (Borer 1983). In the Hebrew present tense, where there is no person marking, subject pronouns are obligatory (cf. Table 1). In other words, Finnish and Hebrew (past and future tense) are *pro*-drop languages with respect to some persons (first and second) and non-*pro*-drop elsewhere (third person).

1*a*) $*Nousi$ $junaan.$
$step-PAST-3SG$ $train-ILL$
'(He) stepped on the train'

1*b*) $Nousin$ $junaan.$
$step-PAST-1SG$ $train-ILL$
'I stepped on the train'

*The first author's research was supported by NSF Grant # SBR-8920230. The second author was supported by the Israeli Foundation Trustees, grant 13/92-94, and by the National Institute for Psychobiology in Israel.

[1] Two dialects of Finnish will be considered in this paper: the dialect based on the written language, referred to here as Standard Finnish, and the spoken Tampere dialect (as acquired by the first author in the 1960s), referred to as Colloquial Finnish. "Finnish" will refer to "Standard Finnish" unless otherwise noted.

2a) $*Ala$ al $ha-rakevet.$
$stepped-PAST-3SGM$ on $the-train$
'(He) stepped on the train'

2b) $Aliti$ al $ha-rakevet.$
$step-PAST-1SG$ on $the-train$
'I stepped on the train'

Table 1.
Obligatory S vs. optional (S) subject NPs in Hebrew[2] and Finnish.

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
Hebrew/past & future	(S)	(S)	S	(S)	(S)	S
Hebrew/present	S	S	S	S	S	S
Standard Finnish	(S)	(S)	S	(S)	(S)	S

2 Previous syntactic approaches

2.1 Licensing and identification

Beginning with Rizzi's seminal work on the null subject parameter (1986) and Hyams' analysis of null subjects in child language (1983; 1986), a two-fold distinction between languages has been maintained: languages that allow null subjects, such as Italian and Spanish (*pro*-drop) and those that do not, such as English (non-*pro*-drop). In order to account for the distribution of empty subjects in both types of languages, Rizzi (1986) proposed that null subjects occur when two conditions are satisfied, *licensing* and *identification*. Licensing concerns the syntactic conditions under which NP subjects can be omitted, while identification deals with the recoverability of the referent of the missing NP. An account of Hebrew along these lines has been provided by Borer (1986); this will be discussed in more detail in Section 5.5.

The existence of languages such as Chinese in which there is no overt agreement marking – but yet they exhibit subject omission – motivated a further distinction between languages. That is, we find (i) null subject languages with rich agreement (Italian), (ii) null subject languages with no agreement (Chinese), and (iii) languages with partial agreement which do not allow subject omission (English). To account for this three-way distinction, Jaeggli & Safir (1989) proposed that subject omission is related to the uniformity of the morphological paradigm: A paradigm is *morphologically uniform* if all or none of the entries in the paradigm bear overt inflectional morphology. It was suggested that languages with such uniform paradigms allow subject omission, whereas languages with mixed paradigms require overt NP subjects.

2.2 Recoverability of the referent

The intuitive idea corresponding to Rizzi's notion of identification is that subjects can be omitted if one can tell from the agreement marking who the subject NP would refer to. That is, the subject

[2]The distribution of null subjects in the Hebrew *future* tense although similar to the past tense, is somewhat more complicated; we will return to this in Section 5.2.

can be left out if subject-verb agreement is sufficiently "rich" such that the referent of the subject can be recovered without an overt subject.

A simple functional explanation, according to which the grammatical subject can only be omitted if its referent can be otherwise identified, would indeed explain the pattern of omission in the classical *pro*-drop languages such as Italian and Spanish. It would likewise account for the distribution of null subjects in the Hebrew *present* tense, as it would in most cases in English. Since person agreement features are not overtly marked in English nor in the Hebrew present tense, the agreement morphology does not convey enough information to pick out the referent of the subject.

A further refinement of the functional notion of recoverability might also explain the null subject pattern in Finnish and in the Hebrew past and future tense. One could maintain that the first and second person subjects are optional because the exact referents – the speaker and the hearer in the exchange – can be determined given the agreement affix, at least in the singular; in the third person, however, the referent is much less easily determinable since in the absence of further contextual information, based on the syntactic information contained in the person features, the 3rd person referent could be anyone that is not the speaker or the hearer at the moment (cf. Ariel 1990).

Note, however, that such considerations do not affect the syntax of NP subject omission in Italian-type languages (see discussion in Section 7). Furthermore, the notion of recoverability does not begin to explain the null subject pattern in German, where the first and second person affixes are unique and rich, and yet an overt (thematic) subject is required; a similar pattern obtains in Colloquial Finnish (Vainikka 1989:186-189 and Section 6).[3]

The analysis that we will propose for the mixed patterns of Finnish and Hebrew makes crucial *syntactic* use of the difference between first and second persons and third person in terms of their referential properties.

2.3 Economy of Projection

In an attempt to remove conditions that are specific to null subjects from the grammar, and in order to avoid the arbitrary determination of functional heads that could license null subjects in various languages, Speas (1994; 1995) has proposed a new theory of null subjects based on licensing syntactic positions. The relevant principle of Economy of Projection is given in (3) ((11), Speas 1994:186):

(3) Project XP only if XP has content.

The principle in (3) amounts to saying that either the head position or the specifier position must be filled; Speas further defines "having content" as involving either a distinct phonological matrix or a distinct semantic matrix (1993:187), where "distinct" refers to elements within a projection itself (and not in its complement).

Speas' basic idea with regard to null subjects is the following: null subjects are possible if the Spec(AgrP) position is not required to satisfy any grammatical conditions. For example, in the Italian-type languages AGR contains phonetic material (i.e. an agreement suffix), and therefore the condition in (3) is satisfied without the Spec(AgrP) position. In the English-type languages, no agreement morpheme is base-generated in AGR; subjects cannot therefore be omitted because the Spec(AgrP) position must be filled in order to fulfil condition (3). Finally, if a language has

[3]Hyams & Wexler (1993) show that a recoverability-from-context type explanation does not account for the distribution of null subjects in child language, either.

no agreement marking at all, an AgrP projection is not projected, and therefore nothing requires an overt subject to be realized; this gives rise to the Chinese-type pattern, where null subjects are possible in the absence of agreement morphology. The subject position in such languages is the Spec(TP) position, but since the head of this position (Tense) is filled by semantic (tense) features, the Spec(TP) position may remain empty.

2.4 Strong vs. weak agreement

Speas (1994) points out some clear empirical problems with Morphological Uniformity as defined by Jaeggli & Safir (1989). She notes that neither Swedish nor Russian allow null subjects although both have a uniform paradigm (Swedish with consistently lacking person/number marking, Russian with overt marking of person and number features). Similarly, Rohrbacher (1994) points out that although both European Portuguese and Brazilian Portuguese have a uniform agreement paradigm, only the European variety of Portuguese allows null subjects. Furthermore, a similar observation was made by Vainikka (1989) concerning Colloquial Finnish: although the language has a uniformly rich agreement paradigm, it does not allow null subjects.

Speas (1994; 1995) makes use of the notion of *strong vs. weak agreement* as defined by Rohrbacher (1994), according to whom a language has strong agreement if the features [1st] and [2nd] are distinctively marked in at least one number, and the feature [sg] is distinctively marked in at least one person. Rohrbacher suggests that languages with strong agreement have an AGR head in which the inflectional affixes are base-generated; see type (4a) below. In languages with weak agreement, on the other hand, inflectional morphology is already attached to the stem prior to lexical insertion, and the AGR position is empty (as in 4b). Speas complements these types with (4c), arriving at the following descriptive generalization (10, Speas (1994:186); notes in square brackets added by us):

(4)
a. A language has null subjects if AGR is base-generated with a morpheme in it. [Italian]
b. A language cannot have null subjects if AGR [i.e. the corresponding suffix] is base-generated on the verb. [English]
c. A language has null subjects if it has no AGR. [Chinese]

Speas claims that despite the lack of overt marking on the verb, Swedish is a language with weak agreement where inflection is base-generated on the verb rather than heading its own functional projection, and therefore Swedish falls in category (4b). According to Speas, Swedish should be analyzed as having weak agreement – as opposed to lacking an AgrP altogether – due to the existence of participial agreement and other agreement phenomena.

Speas (1994:198-199) briefly discusses the problem posed for any notion of a uniform paradigm by the mixed null subject system in the Hebrew past and future tense. She tentatively suggests that the affixes in the first and second person might really be incorporated pronouns. Under her suggestion, Hebrew has an AgrP projection which is base-generated empty in both the past and the present tense, due to "weak" agreement. Thus, overt subjects are required in both tenses in order to satisfy (3). The apparent empty subjects in the past tense actually involve overt pronouns that have been incorporated into the verb; the obligatoriness of these elements has not been accounted for.

As mentioned above, Colloquial Finnish (along with Russian) forms a counterexample to Jaeggli & Safir's original criterion of uniformity, since it is a non-*pro*-drop language regardless of its rich

and uniform agreement paradigm. Furthermore, under the Rohrbacher/Speas criterion of strong agreement, both dialects of Finnish exhibit strong agreement and are therefore predicted to allow null subjects throughout, just like Italian and Spanish; in fact, neither dialect is a regular *pro*-drop language.

While we agree with the general thrust of Speas' model – involving the use of syntactic licensing to replace the Null Subject Parameter – we will pursue a related yet distinct approach to the connection between subject-verb agreement and null subjects in Hebrew and Finnish.

2.5 Null subjects in child language

Recent proposals in the context of empty subjects in child language include Hyams & Hoekstra (1994), Roeper & Rohrbacher (1994) and Sano & Hyams (1994). Both Roeper & Rohrbacher (1994) and Sano & Hyams (1994) associate the distribution of null subjects in early child language with finiteness and verb raising. According to Sano & Hyams, the underspecification of INFL features, and hence of verb raising, is responsible for the null subject phenomenon in early child language. Roeper & Rohrbacher (1994), on the other hand, extend Speas' (1994) proposal further to apply to child English and suggest that null subjects are licensed in the Spec(VP) position in the absence of the functional projection IP (cf. also Rohrbacher & Vainikka (1995) for a similar proposal in child German). Hyams & Hoekstra (1994) attempt to explain various properties of child language, including the distribution of null subjects, as being the consequence of the underspecification of *number* features; an underspecified Number Projection, in their view, effects the interpretation of TP (and DP), which in turn influences the distribution of null subjects.

A detailed analysis of the acquisition literature on this topic is beyond the scope of this paper (but see Levy & Vainikka [in preparation]); however, to the extent that these proposals relate to the distribution of NP subject omission in adult language, it will become clear in the following sections that neither finiteness, nor tense, nor number features can account for the patterns found in adult Hebrew or Finnish.

3 The proposal for NP subject omission

Recall that since Rizzi (1986) the intuition behind the optionality of NP subjects in many of the world's languages has involved syntactic licensing along with a requirement concerning the identification of the NP referent. We contend that licensing and identification are not unrelated; rather, identification, which is essentially a pragmatic concern, is in fact reflected in the syntax, affecting the issue of licensing as well.

Table 2 specifies the features adopted here for person and number; features such as gender and animacy are not reflected here.

Table 2.
Person/number features.

1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
[+speaker] [-hearer] [-plural]	[-speaker] [+hearer] [-plural]	[-speaker] [-hearer] [-plural]	[+speaker] [-hearer] [+plural]	[-speaker] [+hearer] [+plural]	[-speaker] [-hearer] [+plural]

The [+/-speaker] and [+/-hearer] features born by the first and second person are *strongly referential* in that they pick out a unique referent in the world for the subject NP. We propose that due to this characteristic, Universal Grammar (UG) allows for the possibility of representing these features in the subject position, along with NP subjects, provided that there is a discernable phonological connection between 1st and 2nd person pronouns and the AGR-related affix. In fact, it is possible that the subject position is the unmarked position for such strongly referential features. Third person forms, on the other hand, lack strongly referential features, and thus occur in AGR in the unmarked case.

Applying this idea to Finnish and Hebrew, it will be argued that since they can guarantee identification of the subject NP, first and second person features occupy the subject position in these languages, as shown in (5) (intermediate functional projections omitted). That is, the [+speaker]/[+hearer] features are base-generated in the Spec(VP) position – where they are assigned a theta-role and in general behave like other NP subjects – and raise to the Spec(AgrP) position. On the other hand, due to lack of strong referentiality, third person features are base-generated in AGR, rather than in the subject position.

```
(5)       AgrP
        /     \
   Spec        Agr'
    |         /   \
[1st/2nd]  Agr      VP
            |      /  \
         [3rd]  Spec   V'
                      /  \
                     V   NP
```

Since overt third person subjects are required in Finnish and Hebrew, theta-role assignment to the features in AGR need not be posited. Theta-role assignment occurs as usual to positions within the VP, assuming that the third person overt subjects are base-generated in the Spec(VP) position, following the by-now standard VP-Internal Subject Hypothesis (cf. e.g. Sportiche 1988, Kitagawa 1986, Diesing 1990).

It appears that the behavior of the 3rd person subjects is most crucial in determining a language type with respect to null subjects, in that a language in which 3rd person subjects are omitted allows subject omission throughout. This situation is presumably derivable from the referential properties of the 3rd person, as compared to the strongly referential 1st and 2nd person. According to our analysis, person/number features are located in the subject position in languages which allow 3rd person subject omission, as exemplified by Italian and Chinese (cf. Sections 6 and 7 for further discussion of languages other than Hebrew and Finnish, such as German, Swedish, Russian, and Chinese). On the other hand, if 3rd person subjects cannot normally be omitted in the language, this indicates that at least some of the relevant features occupy the AGR position, rather than the subject position. An example of a language where all the relevant features occur in AGR is English, where the Spec(AgrP) must thus always be filled by an overt NP. A further distinction is made between the English-type languages and the Hebrew/Finnish-type languages, conditioned upon the presence or absence of a phonological relatedness that holds between inflectional affixes and pronouns, as will be described in Sections 4.4. and 5.3.

The distinction between person features in terms of *syntactic position* will account for the NP subject omission patterns attested in Hebrew and Finnish, given certain assumptions about

syntactic derivation. These involve Chomsky's (1992) Checking Theory and a principle of licensing (6) to be described below, which guarantee the appropriate surface realization of overt and null NP subjects in the present system.

According to Chomsky's Checking Theory, verb forms are inserted into the VP fully inflected, and their features are checked against the features in the functional heads. Subject-verb agreement is controlled by referential *N-features*, which the NP subject, the AGR head, and the inflected verb form bear. Note that all person/number combinations bear Chomsky's N-features, whether strongly referential or not in the sense described above.

Modifying Speas' principle of licensing (cf. Section 2.3) and in line with Vainikka (1989:12-13), we propose the following principle of licensing syntactic positions:

(6) In order to be licensed, both the head and the specifier of a syntactic position must be filled by phonetic or semantic material at some level of representation.

Most crucial for the current proposal, the principle in (6) applies to the AgrP projection; in order to be projected, *both* the head and the specifier position of this projection need to be filled. Note that this contrasts with Speas' proposal according to which *either* the head or the specifier must be filled. However, both Speas' principle and principle (6) allow the Extended Projection Principle – according to which all clauses must have subjects – to be replaced with a more general principle of licencing syntactic positions.

Under our proposal, third person features in Finnish and Hebrew are checked in the usual manner. The verb (with its affixes and features) raises to AGR, where the features of the verb are checked. Furthermore, the N-features on the verb agree with the corresponding features of the subject NP in Spec(AgrP), verifying subject-verb agreement. On the other hand, since the first and second person features are located in Spec(AgrP), the AGR position in this situation contains no person or number features. Yet, the two positions are coindexed due to Spec-head agreement. Given verb raising, the features of the verb are checked against the features in Spec(AgrP) by virtue of the coindexation relationship between AGR and its specifier position.

After a detailed discussion of Finnish and Hebrew in the following two sections, Section 6 provides analyses of null subjects patterns in other languages. The final discussion section will make some suggestions concerning the interaction between syntax and discourse in determining referentiality. These suggestions will be relevant to our understanding of the phenomenon of NP subject omission in the Chinese-type languages which lack overt realization of AGR features and yet allow NP subject omission, given appropriate discourse conditions.

4 Finnish

4.1 Syntactic background

Finnish is an agglutinative language with mixed word order, although head-initial structures predominate (especially in the verbal domain; cf. Vainikka 1995). In simple declarative sentences the verb occurs in the second position, preceding the temporal adverb position and following the subject NP, as exemplified in (7a). However, when a phrase is questioned or topicalized, as in (7b), the verb ends up in the "third" position – making Finnish a non-V2 language. For an overview of various aspects of Finnish syntax, see Vainikka (1989).

7a) *Pekka muistaa yleensä vastaukset.*
NOM remember – PRES – 3SG usually answers – ACC
'Pekka usually remembers the answers'

7b) *Minkä vastauksen Pekka muistaa aina?*
which – ACC question – ACC NOM remember – PRES – 3SG always
'Which answer does Pekka usually remember?'

In general, the finite verb raises to AGR, but not to C. Various versions of a split-INFL tree (Pollock 1989) have been proposed for Finnish (Mitchell (1991), Holmberg et.al. (1993), Vainikka (1994)). The proposals agree on three projections beneath the CP projection, although slightly different names are used. These projections are, from top to bottom: (i) AgrP, (ii) NegP, and (iii) TP. This ordering of functional projections captures the fact that negation is expressed as a verb which bears subject-verb agreement, but no other inflection (cf. (8) and the corresponding structure in (8')); in negative sentences, the main verb bears tense but no agreement morphology.

8) *En muistanut vastausta.*
neg – 1SG remember – PAST answer – PAR
'I did not remember the answer'

```
(8')              AgrP
                 /    \
              Spec     Agr'
               |       /   \
              NP     Agr     NegP
               |      |     /   \
            [pro]    en   Spec   Neg'
                          |    /     \
                         NP  Neg       TP
                          |   |     /     \
                          t   t   Spec       T'
                                   |       /     \
                                  NP     T          VP
                                   |     |         /   \
                                   t   muistanut  Spec  V'
                                                   |   /  \
                                                  NP  V   NP
                                                  |   |    |
                                                  t   t   vastausta
```

An existing proposal for the distribution of null subjects in Finnish is contained in Vainikka (1989). According to this analysis, agreement affixes are anaphors, which in the 3rd person are bound by an overt NP and in the first and second person are bound by the speaker/hearer features.[4] As will become clearer in the sections to follow, the approach developed in the present work – which

[4]Like null subjects, null possessors in Finnish are only possible in the first and second person, whereas an overt NP possessor is required as the binder of the possessive suffix in the third person. Analyzing the possessive construction in Finnish under the present approach is a subject of future research.

is formulated without reference to principles of the Binding Theory – has broader cross-linguistic appeal than the previous approach of Vainikka (1989).

4.2 Subject omission in the first and second person

As already observed in traditional grammars, first and second person subject pronouns are usually omitted in Finnish, as shown in (9), but not third person subjects (Setälä 1960).

9a) *Jään* *kotiin,* *jos* *pyydät* *kauniisti.*
remain – PRES – 1SG *home – ILL* *if* *ask – PRES – 2SG* *nicely*
'I'll stay home if you ask nicely'

9b) *Kun* *soititte,* *olimme* *juuri* *kaupassa.*
when *call – PAST – 2PL* *be – PAST – 1PL* *just* *store – INE*
'When you called, we were just at the store'

9c) *Jukka* *lähtee,* *jos* *hän* *saa* *kutsun.*
NOM *leave – PRES – 3SG* *if* *he – NOM* *get – PRES – 3SG* *invitation – ACC*
'Jukka will go if he gets an invitation'

To account for the possibility of subject omission in the first and second person, our claim is – as outlined in Section 3 – that the relevant speaker/hearer features are base-generated in the Spec(VP) position, and subsequently raised to Spec(AgrP). That is, these strongly referential features occupy the subject position throughout the derivation (in languages with a special type of agreement paradigm, as described below in Section 4.4 for Finnish and in Section 5.3 for Hebrew). Once the feature bundle has been raised to Spec(AgrP) – as a typical subject NP would – the Spec(AgrP) position is licensed, satisfying our licensing principle in (6). The first and second person features in Spec(AgrP) can also be overtly realized as a subject pronoun.[5] The finite verb (with its base-generated affixes; cf. Chomsky 1992) raises to AGR, where its features are checked against Spec(AgrP), and thereby the AGR position is also licensed.

4.3 Overt subjects in the third person

In the third person, an overt NP subject is required, as shown in (10):

10a) *Kun* *hän* *soitti,* *he* *söivat* *juuri* *aamiaista.*
when *he* *call – PAST – 3SG* *they* *eat – PAST – 3PL* *just* *breakfast*
'When he called, they were just eating breakfast'

10b) **Kun* *soitti,* *söivat* *juuri* *aamiaista.*
when *call – PAST – 3SG* *eat – PAST – 3PL* *just* *breakfast*
'When (he) called, (they) were just eating breakfast'

Since the third person N-features are not strongly referential, i.e. they are insufficient for unique identification of the NP referent, they are base-generated in AGR rather than in the subject position. That is, the third person features in Finnish have the the distribution usually assumed for person/number features in general. The Spec(AgrP) position, however, is empty at D-Structure.

[5]It is unclear whether an overt subject pronoun in the first and second person has the same status as in the familiar *pro*-drop languages such as Spanish, or whether – like Hebrew – an overt pronoun is stylistically unmarked. Judgments with regard to this question are mixed, possibly because of interference with the way overt pronouns are used in Colloquial Finnish.

An overt subject NP raised from the Spec(VP) to the Spec(AgrP) position will satisfy (6) by providing a phonetically realized filler, and we have an explanation for why subjects are obligatory in the 3rd person.

To summarize the discussion so far: the person/number features in the first and second person are strongly referential, and occupy the subject position in Finnish, whereas the third person features occupy the usual verbal position. Since the subject position would otherwise be empty (and all positions need to be filled by something, given the licensing principle in (6)), an overt subject NP is required in the third person.

However, in Finnish something other than a subject NP can occur in the Spec(AgrP) position, as exemplified in (11).[6] In this case, the subject NP raises only to Spec(TP), or possibly remains in Spec(VP) (cf. Holmberg 1994). Assuming the Spec(TP) location of the postverbal subject, the corresponding structure is provided in (11').

11) *Nopan* *löysi* *Maija* *lipaston* *alta.*
die − ACC *find − PAST − 3SG* *NOM* *chest − GEN* *under*
'Maija found the die under the chest'

```
(11')              AgrP
                  /    \
               Spec     Agr'
                |       /   \
               NP     Agr     TP
                |      |     /   \
            nopan   loysi  Spec   T'
                            |    /   \
                           NP   T     VP
                            |   |    /   \
                          Maija t  Spec   V'
                                    |    /  \
                                   NP   V    PP
                                    |   |   /   \
                                    t   t  NP    P
                                           |     |
                                        lipaston alta
```

Thus, the licencing requirement in (6) is satisfied with respect to the Spec(AgrP) position even when the subject does not raise to that position, as long as some XP occupies the Spec(AgrP) position. The subject NP is presumably required in structures such as (10) in order to license the Spec(TP) position.[7] In constructions such as (11), feature checking, along with Nominative Case assignment, would have to obtain under a government relationship between the finite verb in AGR and the postverbal subject, or possibly through Spec-head agreement in the TP.

[6] See Vilkuna (1989) for a thorough description of such structures, and Vainikka (1989) for arguments that the subject does not occupy the Spec(CP) position in constructions such as (11), based on the distribution of WH-elements and topicalization.

[7] In examples such as (10a) where the subject NP has raised all the way to Spec(AgrP), it will have occupied the Spec(TP) position at an earlier point in the derivation, thereby licensing both Spec(AgrP) and Spec(TP).

4.4 Position and form of agreement features

The proposed distinction in the syntactic position of the various person/number features is reflected in the morphosyntax of the subject-verb agreement paradigm, as can be seen in Table 3.

Table 3.
Subject-verb agreement and personal pronouns in Standard Finnish.

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
agr.suffix	-n	-t	-V	-mme	-tte	-vAt
pronoun	minä	sinä	hän	me	te	he

As pointed out by L.Hakulinen (1979:87), the agreement suffixes are phonologically related to the corresponding first and second person pronouns (with the 2SG pronoun reconstructed as *tinä, given a a general /ti/–/si/ rule in Finnish phonology), whereas this is not the case in the third person. Since according to our analysis the first and second person features occupy the subject position, it is not surprising to find a phonological similarity between the affixes and the corresponding subject pronouns. No such similarity is expected in the third person, since the agreement features never occupy a subject position.

5 Hebrew

5.1 Syntactic background

Hebrew is an SVO language where the finite verb normally raises to INFL (Borer 1984, in press; Shlonsky 1990; Hazout 1992) – and as exemplified in (12) – although there are some constructions in which the VSO order is found, as a residue of the basic VSO order in Classical Hebrew.[8] Hazout (1992) argues that the VSO order attested in the verbal gerund construction involves verb raising exceptionally high in the tree, to COMP. For the purposes of this paper, only the AgrP projection – the equivalent of IP under Pollock's (1989) Split-INFL Hypothesis – and the usual verb raising to AGR will be considered; thus, intermediate functional projections have been omitted in the structure (12').

12) *Peter* *zoxer* *et* *ha – tshuva*
NOM *remember – PRES – 3SGM* *ACC* *the – answer*
'Peter remembers the answer'

[8] Doron (1983) has proposed that even Modern Hebrew is a VSO language.

```
(12')              AgrP
                  /    \
               Spec     Agr'
                |       /   \
               NP     Agr     VP
                |      |     /   \
            Peter  zoxer  Spec   V'
                           |    /   \
                          NP   V     NP
                               |     |
                               t   et ha-tshuva
```

The distribution of null subjects in Hebrew has been described and analyzed in various articles by Borer, most notably Borer (1986) and (1989). Borer's analysis will be compared to ours in Section 5.5 below. Following the VP-Internal Subject Hypothesis, we take the Hebrew NP subjects to be base-generated in the Spec(VP) position, where they are assigned a theta-role, and from where they raise to the usual surface subject position, the Spec(AgrP).

Each of the Hebrew tenses has a different subject-verb agreement paradigm, due to the *portmanteau* nature of Hebrew inflectional morphology, where tense, person, number and gender are typically conflated into a single affix.

5.2 Past and future tense

As example (13) shows, overt subjects are not required in the first and second person in the Hebrew past and future tense:[9]

13*a*)	*Halaxti*	*itxa*	*ki*	*racita*
	go − PAST − 1SG	*with − you*	*because*	*want − PAST − 2SGM*

'I went with you because you wanted'

13*b*)	*Elex*	*itxa*	*im*	*tirce*
	go − FUT − 1SG	*with − you*	*if*	*want − FUT − 2SGM*

'I will go with you if you want'

Analogously to what has just been proposed for Finnish, first and second person agreement features – being strongly referential speaker/hearer features – are base-generated in the subject position, the Spec(VP), in the Hebrew past and future tense; again, as shown in the next section, there is a discernable phonological connection between the 1st and 2nd person pronouns and the agreement affixes. The feature bundle then raises to the Spec(AgrP) position in order to satisfy the licensing principle in (6), and the inflected verb raises to AGR.[10] Rather than being phonetically empty, the first and second person features can also be phonetically realized as a pronoun; note that overt pronouns are used in Hebrew without special emphasis or contrast.

[9] There is a tendency in Colloquial Hebrew to preserve the first person singular subject pronoun in the future tense; this will be addressed in more detail below.

[10] If it turns out to be the case that the NP subject can remain in some lower specifier position in Modern Hebrew, the analysis outlined for such structures in Finnish in the previous section would apply to these cases, as well.

Unlike the first and second person, an overt subject NP is required in the third person in both the past and future tense, as shown in (14-15).[11] This completely parallels the situation in Finnish.

14*a*) *Hu halax itxa ki racita*
he go − PAST − 3SGM with − you because want − PAST − 2SGM
'He went with you because you wanted'

14*b*) **Halax itxa ki racita*
go − PAST − 3SGM with − you because want − PAST − 2SGM

15*a*) *Hu yelex itxa im tirce*
he go − FUT − 3SGM with − you if want − FUT − 2SGM
'He will go with you if you want'

15*b*) **Yelex itxa im tirce*
go − FUT − 3SGM with − you if want − FUT − 2SGM

Again, due to their failure to be strongly referential, the third person features occupy the AGR position rather than the subject position. In order to satisfy the licensing principle in (6), the subject position must therefore be filled by an overt NP subject. The overt subject is base-generated in the Spec(VP), raising up to the Spec(AgrP) position to phonetically fill that position.

5.3 Position and form of the agreement paradigm

As in Finnish, the syntactic position of person/number features is reflected in the Hebrew morphology. Compare the Hebrew past and future paradigms to the corresponding personal pronouns in Table 4 (the literary feminine plural forms omitted, where the corresponding masculine form is normally used):

Table 4.
Hebrew past and future subject-verb agreement and personal pronouns.

	1sg.	2sg.m.	2sg.f.	3sg.m.	3sg.f.	1pl.	2pl.	3pl.
past	-ti	-ta	-t	0	-a	-nu	-tem	-u
future	e-	te-	te- -i	ye-	te-	ne-	te- -u	ye- -u
pronouns	ani	ata	at	hu	hi	anaxnu	atem	hem

The first and second person agreement affixes are strongly related to the pronoun forms in the past tense. Similarly, there is a partial phonological relationship between these affixes in the future tense (in the plural), while no such relationship is evident in the third person (cf. Berman 1990). This pattern supports the analysis according to which the first and second person features are base-generated in the subject position, which is also the position of the pronouns shown in Table 4. The third person endings, however, are base-generated in AGR and thus would not be expected to be associated with the pronoun forms.

[11] Both Literary Hebrew and Biblical Hebrew allow null subjects in the third person, in specific constructions. The null subjects in Literary Hebrew are arguably instances of Topic Drop, since the phenomenon is rare in embedded clauses. In a narrative construction of Biblical Hebrew, the subject can be omitted in the third person if a particle *va* is adjoined to the verb; this particle makes a past tense out of a future tense. Except for this special construction, third person subjects cannot be omitted in Biblical Hebrew (Goldenberg, personal communication, March 1995).

As already mentioned, in the first person singular of the future tense the subject pronoun is typically present in spoken Hebrew (cf. Borer 1989:95). We take this to be the consequence of the fact that in the future singular, the person distinction is phonetically obliterated (between *e-* and *ye-*; both are pronounced *ye-*), resulting in the same situation as in the Hebrew present tense, where overt subjects are also obligatory. This is what we now turn to.

5.4 Present tense

The Hebrew present tense verb does not mark person distinctions, but it carries gender and number information. Since the features [+/-speaker] and [+/-hearer] are not realized, present tense forms are not strongly referential. Consequently, the features associated with the present tense verb forms in Hebrew cannot occupy the subject position. Rather, the present tense gender and number features are base-generated in the AGR position, similarly to the third person (gender and number) features in the past and future tenses.

Subject omission is not possible in the present tense, as shown in example (16). In order to satisfy the licensing principle in (6), the subject position must be filled by an overtly realized NP subject, given that the agreement features of the verb do not occupy the subject position.

16*a*)	*Amarti*	*lo*	*she – ata*	*merim*	*et*	*ha – kol*
	say – PAST – 1SG	*to – him*	*that – you*	*raise – PRES – SGM*	*ACC*	*the – voice*
16*b*)	*Amarti*	*lo*	**she – merim*		*et*	*ha – kol*
	say – PAST – 1SG	*to – him*	*that – raise – PRES – SGM*		*ACC*	*the – voice*

'I said to him that he is raising his voice'

This example further shows that even when the matrix clause contains a possible referent (*lo* 'to-him') of the omitted subject, subject omission in the embedded clause is not possible (cf. discussion in the next section).

Note that according to the present analysis the syntactic tree is identical for the present and past tense of Hebrew, a desirable consequence. The difference in the subject omission pattern between the tenses is explained solely by presence or location of the person/number/gender features.

5.5 A comparison of the current analysis with Borer (1986;1989)

To account for NP subject omission in Hebrew, Borer (1986, 1989) has proposed that the Hebrew AGR is an anaphor, requiring a binder. This, according to Borer (1989), holds for the past and future tense, whereas in the present tense AGR there is no person slot and thus both binding and so-called *I-identification* are disallowed in the present tense.

The ungrammaticality of subject omission in the third person matrix clauses follows from the lack of I-identification, given impoverished person marking in the third person. However, as pointed out by Borer, when an NP in the matrix clause is present as a binder, a third person subject *can* be omitted in an embedded clause (past or future tense). In (17a) there is a binder (*lo*) in the matrix clause, while in (17b) there is no binder for the embedded AGR, and subject omission is impossible:

17*a*)	*Amarti*	*lo*	*she – yavo*	*iti*
	say – PAST – 1SG	*to – him*	*that – come – FUT – 3SGM*	*with – me*

'I said to him that he will come with me'

17*b*)	$*Amarti$	la	$she-yavo$	iti
	$say-PAST-1SG$	$to-her$	$that-come-FUT-3SGM$	$with-me$

'I said to her that he will come with me'

However, although the pattern for the third person shown in (17) is reminiscent of a binding pattern, it is not an instance of usual anaphoric binding – as also acknowledged by Borer – since (17a) would involve binding into a tensed clause. This not possible with the regular anaphors in Hebrew.

Furthermore, the Finnish data suggest that the null subject distribution pattern shared by Hebrew and Finnish is distinct from the phenomenon illustrated in (17). In Finnish, subject omission in the third person is not possible, even if there is a possible antecedent providing the referent of the null subject in the matrix clause:

18*a*)	*Sanoin*	*hänelle,*	*että*	*hän*	*tulee*	*mukaani.*
	$say-PAST-1SG$	$to-him$	*that*	*he*	$come-PRES-3SG$	$with-me$

'I said to him that he will come with me'

18*b*)	$*Sanoin$	*hänelle,*	*että*	*tulee*	*mukaani.*
	$say-PAST-1SG$	$to-him$	*that*	$come-PRES-3SG$	$with-me$

The overt subject pronoun *hän* is required in (18a), unlike in the Hebrew counterpart (17a).

Under our approach to the Hebrew data, the first and second person features (although not third) are nominal in nature, occurring in the subject position. Rather than being anaphors, they are regular pronouns as far as the Binding Theory is concerned, completely equivalent to their alternative realizations as an overt subject pronoun. However, in Hebrew (but not in Finnish) the referential N-features on the embedded verb can be exceptionally checked against an NP in the matrix clause; we leave the details of this language-specific mechanism open.

6 Null NP subject patterns in other languages

With respect to null NP subjects, languages are divided into the following groups: first, languages in which null NP subjects can occur in all persons, e.g. Spanish, Italian, Tamil and Chinese. In these languages person/number features, if they exist, occur in the subject position. An identifying characteristic of these languages is the possibility of third person subject omission.

Second, languages in which thematic third person subjects are used unemphatically and non-contrastively; these languages fall into two categories:

(a) those in which the agreement paradigm reveals a pronominal connection in both 1st and 2nd person – but not in 3rd person – such as Hebrew and Standard Finnish,

(b) those in which no such pronominal connection can be identified, such as English, Swedish, German, French, Russian, and Colloquial Finnish.

In the (a) type languages, 3rd person features are base-generated in AGR while 1st/2nd features occur in the subject position, resulting in a mixed subject omission pattern. In the (b) type languages, all person/number features occupy a head position, and an overt subject is required to license the subject position.

6.1 Languages which allow subject omission in third person

The claim is that UG does not allow for a pattern where omission of referential subject NPs is possible in the third person, but not in the first and second person. This pattern would require base-generation of the third person features in the subject position, while the first/second features would occur in AGR – such a situation would only arise if the third person features were *more* referential than the first and second person features. Using a discourse-oriented approach which posits differential strength of accessibility of various referential markers, Ariel (1990) has come to a similar conclusion, namely that if a language has *pro*-drop, it cannot be the case that third person subjects are omitted while first and second person subjects are retained. Her reasons, much like ours, concern the difference in accessibility of 1st/2nd person referents vs. 3rd person referents.[12]

Under the approach outlined here, the traditional *pro*-drop languages such as Italian and Spanish represent a situation where all subject-verb agreement features are base-generated in the Spec(VP) position, i.e. the subject position. Since the subject position is already filled by the [+/-speaker, +/-hearer] features, the position is licensed, and need not be phonetically filled. However, the features *can* be overtly realized as a subject pronoun, if emphasized or contrasted. This results in the typical null subject pattern, in all persons.

Languages of this type are the only ones which allow omission of third person subject NPs – i.e. subject omission in the third person uniquely characterizes a 'traditional' *pro*-drop language – as exemplified in (19) for Spanish (from Jaeggli & Safir (1989), ex.(4a)).

19) *Juan/el/0 siempre habla de si – mismo.*
John/he/0 always talks about himself

If the language has agreement affixes, as do Italian and Spanish, the prediction is that although the morphological information seems to be a less reliable indicator of a consistent *pro*-drop language than 3rd person subject omission, such suffixes might tend towards the form of the pronouns, in all persons. An example of a *pro*-drop language where almost all agreement suffixes are pronoun-like is the Dravidian language Tamil, as shown in Table 5.[13]

Table 5.
Subject-verb agreement and personal pronouns in Tamil.

	1sg.	2sg.	3sg.m.	3sg.f.	1pl.	2pl.	3pl.
suffix	-een	-ai	-aan	-aal	-oom	-ingal	angal
pronoun	naan	nii	avan	aval	naangal	niingal	avangal

In addition to the languages with overt person/number agreement marking, there are *pro*-drop languages of the Chinese type in which – although overt AGR features are lacking – null NP subjects are allowed, subject to pragmatic considerations (cf. e.g. Huang 1989). Note that our proposal covers such languages, as well. If agreement features are represented at all in the grammar of these languages, the features occupy the subject position. If no AgrP is posited, following Speas (1994),

[12] Pursuing Ariel's discourse-oriented approach, Gutman (in preparation) also suggests that the morphological relationship that exists between pronouns and AGR in Hebrew and Finnish may account for subject omission.

[13] Thanks to Inigo Thomas for providing us with this information.

only positions in lower projections will need to be licensed.[14] Given the lack of overt morphology, the issue of identification of the missing NP subject, in particular as we consider the language learner, must be allocated to the discourse, as suggested by Huang (1989). We return to this point in the final section of the paper.

Let us now turn to languages which do not allow omission of third person subjects. Two subtypes are found here: (a) the traditional non-*pro*-drop languages (e.g. English), and (b) Finnish and Hebrew.

6.2 Languages which require third person subjects

In languages such as English where subjects are obligatory in all persons, we propose that all person/number features are base-generated in an INFL-level functional head, such as AGR.

It was suggested above that omission of third person subjects is a unique characteristic of *pro*-drop languages. On the other hand, the *presence* of overt third person subjects does not suffice to distinguish English-type languages from Finnish-type languages, since third person subjects are required in both language types. In such cases, it is the agreement paradigm which serves to distinguish English-type languages from Finnish-type languages, as follows. For languages which require overt third person subject NPs:[15]

20)

- if in at least one number
    1. the 1st and 2nd person agreement affixes are identifiable as being derived from the corresponding pronouns,
    2. while the 3rd person affix is identifiable as *not* being derived from the corresponding pronoun,
    3. then the language is a Finnish/Hebrew type language (with *pro*-drop in 1st/2nd, and non-*pro*-drop in 3rd)
- otherwise, the language is an English-type language (traditional non-*pro*-drop)

Recall from Tables 3 and 4 that the agreement suffixes in the plural 1st and 2nd person are clearly related to the corresponding pronouns in both Finnish and Hebrew, while no such connection is found in the 3rd person plural. The connection between pronouns and affixes is less obvious in the singular, but can be demonstrated diachronically. Let us now turn to a comparison between Standard Finnish and Colloquial Finnish.

[14] There is some reason to believe that an IP-level projection with person/number features in the subject position is posited, since according to Huang (1989) a Chinese (embedded) sentence cannot have an overt subject unless it has the possibility of containing auxiliary material. That is, the relevant projection might be altogether lacking in certain embedded clauses, while others (along with matrix clauses) contain a projection which houses both auxiliary elements and an optional subject.

[15] (20) is reminiscent of Rohrbacher's (1994) definition of *strong agreement* – associated with verb raising – according to which first and second person affixes must be distinct in at least one number (and number must be distinctively marked in at least one person).

6.2.1 Colloquial Finnish

The subject omission pattern in Colloquial Finnish is clearly distinct from the Standard Finnish pattern already described. The Colloquial Finnish examples in (21) reveal that omitting the subject NP in any person results in an ungrammatical sentence, a fact not generally observed (although see Vainikka 1989 for some discussion):

21) *Me jännitettiin kauheesti,*
we worry − 1PL − PAST terribly
'We worried terribly...'

21a) *että ehri − n − ks *(mä) viimeseen junaan.*
that reach − 1SG − Q I − NOM last train
'...whether I would reach the last train'

21b) *että ehri − [0] − ks *(sä) viimeseen junaan.*
that reach − 2SG − Q you − SG − NOM last train
'...whether you (sg.) would reach the last train'

21c) *että ehti − i − ks *(se) viimeseen junaan.*
that reach − 3SG/PL − Q he/she/it − NOM last train
'...whether he/she/it would reach the last train'

21d) *että ehri − tään − ks *(me) viimeseen junaan.*
that reach − 1PL − Q we − NOM last train
'...whether we would reach the last train'

21e) *että ehri − tte − ks *(te) viimeseen junaan.*
that reach − 2PL − Q you − PL − NOM last train
'...whether you (pl.) would reach the last train'

21f) *että ehti − i − ks *(ne) viimeseen junaan.*
that reach − 3SG/PL − Q they − NOM last train
'...whether they would reach the last train'

As shown in (21), Colloquial Finnish is a consistently non-*pro*-drop language, suggesting that all person/number features are located in AGR – as opposed to Standard Finnish, where only 3rd person features are in AGR. Consequently, the Standard Finnish equivalents of examples (21a-b) and (21d-e) would be grammatical – and perhaps preferred – without an overt subject.[16]

The differences in NP subject omission between the two dialects of Finnish may be accounted for by the differences in the agreement morphology. Although the paradigms in the two dialects are similar, there are differences, as shown in Table 6. For convenience, Table 6 lists the subject pronouns along with the relevant AGR features.

[16] Various features of the examples in (21) distinguish them from Standard Finnish: the usage of the colloquial 1PL suffix in the matrix clause, the vowel quality in the words *kauheesti* and *viimeseen*, the usage of an overt complementizer *että* in an embedded yes/no question, the realization of the weak grade of the verb *ehtiä* as *ehri-* (rather than *ehdi-*), the form of the question suffix (*-kO* in Standard Finnish), and the form of the 3PL suffix.

Table 6.
Subject-verb agreement and pronouns in Standard and Colloquial Finnish.

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
Standard Finnish agr.	-n	-t	-V	-mme	-tte	-vAt
Standard Finnish pronouns	minä	sinä	hän	me	te	he
Colloquial Finnish agr.	-n	-t	-V	-tVVn	-tte	-V
Colloquial Finnish pronouns	mä	sä	se	me	te	ne

In Colloquial Finnish, the 1PL affix no longer reflects the form of the corresponding pronoun (the colloquial 1PL suffix is also the impersonal passive suffix), and the original 3SG form is now used for both 3SG and 3PL. The clear pronoun/affix connection is thus lacking in the Colloquial Finnish plural paradigm, and – assuming that the reconstructable relationship between 2SG pronoun and agreement (see discussion after Table 3) is not sufficiently strong in the singular – there is thus no number (singular or plural) where where 1st/2nd affixes are related to the corresponding pronouns. Therefore, Colloquial Finnish ends up in the same category as English, where all person/number features are in AGR and where the subject position must be filled by an overt NP.

6.2.2 The Indo-European non-*pro*-drop languages

Other straightforward non-*pro*-drop languages include the Germanic languages English – where there are no overt 1st/2nd affixes – and Swedish, where there are no overt agreement affixes.[17] Similarly, the Romance language French – where the agreement affixes bear no resemblance to the pronouns – belongs to this category.[18] In the present system, these languages reflect a grammar where the [+/- speaker/hearer] features occupy a head position, and therefore the subject position is filled in order to satisfy the licensing condition in (6).

Furthermore, German will be considered to be a non-*pro*-drop language of the English and Swedish type, regardless of the fact that it allows null expletives under certain conditions. Our approach to these null expletives would involve considering the possibility that something other than a subject is occupying the usual subject position in these constructions, along the lines argued by Speas (1995) for Yiddish and German. Referential subject NPs cannot be omitted, however; this again follows if all of the [+/- speaker/hearer] features are base-generated in a functional head. As expected, the agreement suffixes in German bear no resemblance to the subject pronouns.

Russian appears to fall into the category of the Germanic languages and Colloquial Finnish, since subject pronouns are obligatory. Russian agreement affixes might be construed to show some resemblance to the pronouns, at least in 1SG (*-ju* 1SG vs. *ja* 'I'). However, there is no number in which *both* the first *and* the second person affixes resemble the pronouns, and thus Russian – like English and Colloquial Finnish, but unlike Standard Finnish – is a consistently non-*pro*-drop language.

[17] Hermon & Yoon (1989) discuss other languages which do not have overt agreement marking, and yet require overt subjects: Papiamentu, Duka, Guyami and Tagalog. Under the present approach such languages are unproblematic; since they do not allow omission of third person subjects, they are not *pro*-drop languages, and since they do not exhibit the pronoun connection in the agreement paradigm, they will not be treated as the Finnish/Hebrew type. Thus, by the second clause in (20), these languages are consistently non-*pro*-drop, exactly as in the case of Swedish.

[18] If French subject clitics are treated as agreement affixes, then French would be an instance of a *pro*-drop language where the agreement affixes are base-generated in the subject position.

Note that under traditional approaches to the null subject phenomenon, both Russian and Colloquial Finnish are problematic since they are non-*pro*-drop languages with rich subject-verb agreement. Under the present approach, these two languages are straightforward examples of agreement affixes consistently occurring in AGR.

6.3 Acquisition considerations

Under the approach developed here, agreement morphology – although highly relevant in some cases – is no longer the single crucial feature which distinguishes for the learner a consistent *pro*-drop language from a non-*pro*-drop or a mixed *pro*-drop language. Rather, the crucial trigger is omission of thematic NP subjects in the third person. This will suffice to distinguish consistent *pro*-drop languages from other languages. Note that since the trigger is claimed to be omission of *third* person subjects, omission of imperative second person subjects or omission of first person subjects in the English "diary drop" style do not constitute triggers for a *pro*-drop language.

However, the *presence* of third person subject NPs will not as such distinguish language types – since overt subjects are possible even in *pro*-drop languages – whereas the presence of non-constrastive, unemphasized third person subject pronouns will inform the language learner that the language s/he is acquiring is *not* a *pro*-drop language. On the other hand, a further trigger is required in order to distinguish the English-type non-*pro*-drop languages from the Finnish-type languages. The Finnish/Hebrew type of subject-verb agreement paradigm can be taken to act as a trigger for the mixed *pro*-drop language type discussed here, while the absence of such morphological information – along with the absence of the third person omission trigger – leads the language learner to assume that s/he is acquiring a consistently non-*pro*-drop language of the common Indo-European type. (See Levy & Vainikka (in preparation) for an account of the acquisition of NP subject omission in Hebrew and Finnish along the lines suggested here.)

7 Concluding remarks

The motivation for the approach to NP subject omission proposed in the current paper is two-fold. First, our aim was to encode in the syntax the intuition that in languages with rich agreement morphology, agreement marking on the verb in effect functions as a pronominal subject. Such an intuition is evident in much of the previous work on the distribution of null NP subjects.[19] The second, more specific motivation was to offer an analysis which will account for the mixed systems of languages like Finnish and Hebrew in which conditions upon subject NP omission may differ among persons and tenses.

To summarize the analysis that has been proposed here: There are three ways in which person features can be expressed in the syntax.

1. Due to their referential nature, first and second person features may be located in subject position, while third person features are located in AGR (provided the agreement pattern reveals the connection between the pronouns and the affixes in the 1st/2nd person). This situation is found in Finnish and Hebrew.

[19] This was perhaps most strongly expressed in McCloskey & Hale (1984) in their paper on Irish, where – although the classical *pro*-drop analysis is adhered to – the authors in fact express their intuition that for the Irish sub-set of verbs that are fully inflected "the inflectional subject behaves for syntactic purposes as if it were an overt pronoun" (p.493).

2. In the English-type languages, all features are located in AGR.

3. In the Italian-type languages, all features are located in the subject position.

Assuming that both AGR and its specifier position must be filled (cf. the licensing principle in (6)), it follows that overt subjects are required in English, while Italian allows subject omission throughout. The mixed pattern of Finnish and Hebrew – where overt subjects are required only in the third person – also follows from the position of the features, since the subject position must be filled just in the situation in which the person features are not in the subject position, i.e. the third person.

This proposal captures the idea that the fully specified AGR features might in fact occupy the subject position. It further suggests that the pragmatic notion of speaker/hearer and the difference that exists in identifying first and second person referents vs. third person referents within a given discourse may find a formal expression within syntax. Note, however, that not all languages indeed encode this distinction syntactically. An objection may be raised, therefore, concerning the generality which is expected when pragmatic principles are thought to be operating. Aren't discourse related considerations typically expected to apply across languages in uniform ways? For example, if recoverability of the specific NP referent is affecting NP subject omission such that there is a difference between 1st/2nd and 3rd in terms of presence or absence of an explicit NP subject, then – so the argument goes – one would expect to find this difference cross-linguistically, i.e. there should not be languages of the Italian type in which such differences among the persons do not arise.

However, it might be argued that this is not the right way to look at the operation of pragmatic principles. Rather, the generality is to be found in the *need* to resolve pragmatic concerns such as "Who are the referents of the NPs in a sentence? What is their role in the discourse?". Such concerns should indeed apply to all languages, yet the resolution of pragmatic issues need not be uniform cross-linguistically, any more than one might expect syntactic principles to have invariant manifestations cross-linguistically.

For example, in the domain of NP subject omission, the relevant question becomes the following: How and to what extent is the syntactic component operative in determining the referent of the subject NP?

For the languages that have been discussed above, there are two separate yet consonant syntactic devices which are involved in picking out NP referents, namely AGR features and subject NPs. In fact, one may think about the obligatory presence or absence of such devices in terms of a language-specific "division of labor" with regard to the issue of the recoverability of NP referents: Languages may be making use of a single syntactic device to help pick out referents – e.g. either overt AGR or NP subjects are obligatory (Italian or Swedish) – or they may be using both devices (German, Colloquial Finnish), or neither (Chinese); in the latter case contextual support becomes the only way of determining referentiality.

In terms of the "division of labor" metaphor, one might say that in the German type languages, there is a maximal syntactic involvement in the pragmatic issue of determining NP referents since both overt AGR and NP subjects are required for all persons. Similarly, in the case of Finnish and Hebrew, syntax seems to play a central role in the resolution of the problem of referentiality, by making fully explicit, *syntactically*, the differences between 1st/2nd and 3rd, as expressed through their different positions on the tree.[20]

[20]Pronouns are considered to be a special case of full NPs. Although even with the pronoun third person remains

We have thus reached a many-fold characterization of the interaction between syntax and discourse in the case of NP subject omission, rather than a diachotomous one as proposed in Huang (1984). It is most likely that once the problem of NP subject omission is further examined, there might be more ways than contemplated here in which the interaction between syntax and discourse will manifest itself.

---

opaque yet we assume that the redundant mention of AGR + pronoun potentially assists in the recoverability of the referent of 3rd person.

8 References

Ariel, M. (1990) *Accessing noun-phrase antecedents.* London: Routledge.

Berman, Ruth A. (1990) 'On acquiring an (S)VO language: subjectless sentences in children's Hebrew', *Linguistics* 28:1135-1166.

Borer, Hagit (1983) *Parametric Syntax. Case Studies in Semitic and Romance Languages.* Dordrecht: Foris.

Borer, Hagit (1986) 'I-subjects', *Linguistic Inquiry* 17:375-416.

Borer, Hagit (1989) 'Anaphoric AGR' in O.Jaeggli & K.J.Safir (eds.) *The Null Subject Parameter.* Dordrecht: Kluwer.

Borer, Hagit (in press) 'Up and Down of Hebrew Verb Movement' *Natural Language and Linguistic Theory*

Chomsky, Noam (1992) 'A Minimalist Program for Linguistic Theory' *MIT Occasional Papers in Linguistics*, No. 1.

Diesing, Molly (1990) 'Verb Movement and the Subject Position in Yiddish' *Natural Language and Linguistic Theory* 8:41-79.

Doron, Edit (1983) *Verbless predicates in Hebrew.* Unpublished University of Texas (Austin) Ph.D. dissertation.

Hakulinen, Lauri (1979) *Suomen kielen rakenne ja kehitys.* ('The structure and development of the Finnish language') Keuruu, Finland: Otava.

Hazout, Ilan (1992) 'The verbal gerund in Modern Hebrew' *Natural Language and Linguistic Theory* 10:523-553.

Hermon, Gabriela & Yoon, J. (1989) 'The licensing and identification of pro and the typology of AGR', *CLS* 25:174-192.

Holmberg, Anders (1994) 'The OVS construction in Finnish', presentation at the EUROTYP Word Order Workshop, Group 2 Meeting, Durham.

Holmberg, Anders, Urpo Nikanne, Irmeli Oraviita, Hannu Reime & Trond Trosterud (1993) 'The structure of INFL and the Finite Clause in Finnish', in A.Holmberg & U.Nikanne (eds.) *Case and Other Functional Categories in Finnish Syntax.* Berlin: Mouton de Gruyter.

Huang, James (1984) 'On the Distribution and Reference of Empty Pronouns', *Linguistic Inquiry* 15:531-574.

Huang, James (1989) 'Pro-Drop in Chinese: A Generalized Control Theory', in O.Jaeggli & K.J.Safir (eds.) *The Null Subject Parameter.* Dordrecht: Kluwer.

Hyams, Nina (1983) *The Acquisition of Parameterized Grammars*, CUNY Ph.D. dissertation.

Hyams, Nina (1986) *Language Acquisition and the Theory of Parameters.* Dordrecht: Reidel.

Hyams, Nina & Teun Hoekstra (1994) 'The Syntax and Pragmatics of Early Clause Structure', presentation at the Essex Acquisition Workshop.

Hyams, Nina & Kenneth Wexler (1993) 'On the Grammatical Basis of Null Subjects in Child Language', *Linguistic Inquiry* 24:421-459.

Jaeggli, Osvaldo & Ken Safir (1989) 'The Null Subject Parameter and Parametric Theory', in O.Jaeggli & K.J.Safir (eds.) *The Null Subject Parameter.* Dordrecht: Kluwer.

Kitagawa, Yoshi (1986) *Subjects in Japanese and English.* University of Massachusetts/Amherst Ph.D. dissertation.

McCloskey, J. & Ken Hale (1984) 'On the Syntax on Person-Number Inflection in Modern Irish', *Natural Language and Linguistic Theory* 1:442-487.

Mitchell, Erika (1991) 'Evidence from Finnish for Pollock's theory of IP', *Linguistic Inquiry* 22:373-379.

Pollock, Jean-Yves (1989) 'Verb movement, UG and the structure of IP', *Linguistic Inquiry* 20:365-424.

Rizzi, Luigi (1986) 'Null Objects in Italian and the Theory of *pro*', *Linguistic Inquiry* 17:501-557.

Roeper, Tom & Bernhard Rohrbacher (1994) 'Null Subjects in Early Child English and the Theory of Economy of Projection', University of Massachusetts/Amherst and University of Pennsylvania ms.

Rohrbacher, Bernhard (1994) *The Germanic VO Languages and the Full Paradigm: A Theory of V to I Raising.* University of Massachusetts/Amherst Ph.D. dissertation.

Rohrbacher, Bernhard & Anne Vainikka (1995) 'Verbs and Subjects before Age 2: The Earliest Stages in Germanic L1 Acquisition', *Proceedings of NELS* 25(2):55-69.

Sano, Tetsuya and Nina Hyams (1994) 'Agreement, Finiteness, and the Development of Null Arguments', *Proceedings of NELS* 24:543-558.

Setälä, E.N. (1960) *Suomen kielen lauseoppi.* ('The grammar of Finnish language') Helsinki: Otava.

Shlonsky, Ur (1990) '*Pro* in Hebrew Subject Inversion', *Linguistic Inquiry* 21:263-275

Speas, Margaret (1994) 'Null Arguments in a Theory of Economy of Projections', *University of Massachusetts Occasional Papers* 17:179-208.

Speas, Margaret (1995) 'Economy, Agreement and the Representation of Null arguments', University of Massachusetts/Amherst ms.

Sportiche, Dominique (1988) 'A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure', *Linguistic Inquiry* 19:425-450.

Vainikka, Anne (1989) *Deriving Syntactic Representations in Finnish.* University of Massachusetts/Amherst Ph.D. dissertation.

Vainikka, Anne (1994) 'IP-related Projections in Finnish', *Penn Review of Linguistics*, vol. 18.

Vainikka, Anne (1995) 'Head-initial and head-final constructions in Finnish', presentation at the GLOW Workshop on Inflection and Word Order in Finno-Ugrian Languages, Tromso.

Vilkuna, Maria (1989) *Free Word Order in Finnish – Its Syntax and Discourse Functions.* Helsinki:SKS.