Null subjects and polarity focus

Anders Holmberg

University of Newcastle

1. Introduction[1]

Some null-subject languages can't drop the subject pronoun in the second conjunct in sentences which translate as (1), where the pronoun is coreferent with *John*.

(1) They say that John doesn't speak French, but **he** does.

Among the null-subject languages which don't allow dropping the pronoun are Chinese (various dialects), Czech, Finnish and European Portuguese. Among the languages that allow it are Hungarian, Japanese, Persian, and Spanish. The prohibition against dropping the pronoun in this position is quite unexpected, given standard views on null subjects: In the languages which rely on agreement to license a null subject, the agreement is exactly the same in (their counterparts of) the second conjunct in (1) as in other finite clauses. Furthermore, the subject pronoun is old information, not focused, not introducing a new topic, and there is an antecedent for the pronoun in the preceding clause. For some null-subject languages this is sufficient to allow dropping the subject in (1), but for other null-subject languages it isn't.

In this paper I will propose an explanation of this variation. It is based on a correlation which appears to hold between how languages deal with pro-drop in the context (1) and how the languages reply to yes/no-questions (YNQs): The languages which don't allow pro-drop in (1) standardly reply to YNQs by repeating the finite verb of the question, as in (2), from Finnish:

[1] The research for this paper was supported by grant R280168 from the AHRB. The following persons all helped me with the data, either relying on their own intuitions or by collecting data from unnamed informants: Girma Awgichew Demeke, Abdelkader Fassi Fehri, Dimitra Kolliakou, Ur Shlonsky, Aniko Csirmaz , Balazs Suranyi, Anna Cardinaletti, Makiko Mukai, Kook-Hee Gill, Nader Jahangiri, Aniko Liptak, Pascual Masullo, Ricardo Bermudez-Otero, Atakan Ince, Ilhan Cagri, David Willis, Winnie Yi, Peter Biskup, Ove Lorentz, K.A. Jayaseelan, Nianling Yang, Marit Julien, Patrycja Jablonska, K.V. Subbarao, Michelle Sheehan.

(2) – Puhuuko Joni ranskaa?
speaks-Q John French
'Does John speak French?'
– Puhuu.
speaks
'Yes.'

The languages which allow pro-drop in (1) standardly reply by using a special affirmation particle, as in (3), from Hebrew.

(3) – Ha'im John medaber corfatit?
Q John speaks French
'Does John speak French?'
– Ken.
yes

The reason why these two properties might be related is that both involved *polarity focus*. A YNQ is a question about the polarity of a proposition (true or false, yes or no). In the reply, therefore, polarity is focus, while the rest is presupposed. In (1), the second conjunct asserts the falsity of a proposition made in the first conjunct; thus here, too, polarity is focus while the rest is presupposed.

I will call the languages which allow pro-drop in the second conjunct in (1) **A-languages**, and the ones which don't allow it **B-languages**. The proposal is that pro-drop in the second conjunct in (1) is prohibited in B-languages because of a competition of derivations. In languages where YNQs are answered by a bare finite verb, a subjectless sentence has two derivations. One is the null-subject derivation, employing a null pronoun (or deleting a pronoun) in specIP. The other derivation involves movement of the finite verb to C with deletion of IP. In a context forcing polarity focus, the latter derivation wins. However, since the context does not provide a good antecedent for a deleted IP in (1), the result is ill formed. Thus pronouncing the pronoun is a must.

2. A-languages and B-languages

The following are two lists consisting of the languages from which I have so far gathered data, divided according to whether they allow or don't allow pro-drop in (1):

A-languages (null subject OK)	**B-languages** (null subject not OK)
Amharic	Cantonese Chinese
Arabic	Czech
Greek	Finnish
Hebrew	Indonesian
Hungarian	Malayalam
Italian	Mandarin Chinese
Japanese	North Saami
Korean	Polish
Persian	E. Portuguese (some speakers)
Spanish	Telugu
Turkish	
Welsh	

The following are translations of (1), in some cases slightly modified, in some of the B-languages:

(4) **a. Finnish**
Ne sanoo että (minä) en puhu ranskaa, mutta *(minä) puhun.
they say that I not-1SG speak French but I speak-1SG
'They say I don't speak French, but I do.'

b. Cantonese Chinese
Keoidei waa Siuming m-sik gong faatman, daanhai *(keoi) sik.
they say Siuming not know speak French but he know
'They say John doesn't speak French, but he does.'

c. Czech
Říkají, že John nemluví francouzsky, ale *(on) mluvi
say-3PL that John not-speaks French but he speaks

d. European Portuguese
Dizem que o João não fala frances, mas *(ele) fala.
say-3PL that John not speak French but he speaks

e. Malayalam
awar paRayunnu [John French samsaarik'k'-illa enn@], pakSe *(awan)
they say John French speak -not COMP but he
samsaari k'k'-um
speak FUT

The following are translations of (1), or an equivalent sentence, in some of the A-languages:

(5) **a. Amharic**
yohannes färänsayiNNa a -y -nagäri -m yilalu, gin yi -nagär-al
John French neg -3msS-speak-neg say.3PL but 3msS-speak-PRES

b. Hungarian
Azt mondjak, hogy Janos nem beszel franciaul, de beszel.
that say-3PL that John not speak French but speaks

c. Italian
Dicono che Gianni non parli francese ma invece lo parla.
say-3PL that John not speak French, but in-reality it speaks

d. Japanese
Karera wa John ga furansugo hanasanai to itta, dakedo hanasu yo.
they TOP John French speak COMP say but speak EMPH

e. Persian
Miguyand ke John engilisi sohbat nemikonad, amma sohbat mikonad.
say-3PL that John English speak not-does but speak does

It is noteworthy that the A/B division cuts across the agreement pro-drop/discourse pro-drop division. In the A-group, Japanese and Korean are discourse pro-drop languages, entirely devoid of subject-verb agreement morphology (except honorific agreement), while the others are agreement pro-drop languages, relying in some degree on rich subject-verb agreement to license the null subject. In the B-group the Chinese dialects, Indonesian, Malayalam and Telugu are discourse pro-drop languages, while all the others are agreement pro-drop languages.

Neither does the A/B division correlate with defective agreement or partial pro-drop. A comparison of Hebrew and Finnish is particularly instructive in this regard: As discussed in Vainikka & Levy (1999), Hebrew and Finnish are both partial pro-drop languages in very much the same way: Hebrew allows pro-drop of a 1st and 2nd person subject in the past and future tenses, but allows pro-drop of a referential 3rd person subject only when bound by an argument in a higher clause. Finnish allows pro-drop of a1st and 2nd person subject in all tenses and moods, but pro-drop of a 3rd person referential subject only when bound by an argument in a higher clause; see Holmberg (to appear). However, Hebrew is an A-language, Finnish a B-language.

The A/B division mostly follows family lines, but not perfectly: The Semitic languages checked so far, Amharic, Arabic, and Hebrew, are all A-languages. The Slavic languages checked so far (Czech and Polish) are both B-languages. The Dravidian languages checked so far (Malayalam and Telugu) are B-languages. The Chinese languages/dialects checked so far (Cantonese and Mandarin) are B-languages.[2]

Of the Uralic (Finno-Ugric) languages checked so far Hungarian is an A-language while Finnish and North Saami are B-languages. It could be noted, though, that Hungarian is very distantly related to Finnish and North Saami within the Uralic family, while Finnish and North Saami are fairly closely related, and have furthermore been in contact for a very long period of time.

Finally, of the Romance null-subject languages checked so far (Italian, Spanish, European Portuguese) Italian and Spanish are A-languages. In E. Portuguese

[2] I have been informed that Tamil, another Dravidian language, and Taiwanese Chinese are also B-languages.

some speakers accept pro-drop in (1), others don't. This is striking evidence that the A/B division is not that deeply rooted in genetic or 'deep' typological characteristics of languages, but is susceptible to change.

3. Some explanations which don't work

Conceivably the problem for the B-languages is that the antecedent in (1) is inaccessible, being the subject of an embedded clause in the first conjunct. This might be particularly crucial in discourse-pro-drop language, where there is no local licensing of the null subject by agreement. The difference between A and B-language could then be that the A-languages happen to be more tolerant than the B-languages as regards the (hierarchic) distance between the null subject and the antecedent (along the lines of Ariel's (2001) accessibility theory). The following examples from the B-languages Finnish, Cantonese and Telugu show that the structural relation between the subject of the second conjunct in (1) is not the crucial factor. (6) and (7), from Finnish and Cantonese respectively, show that the antecedent of the null subject in a *but*-conjunct may be the subject of an embedded clause in the preceding conjunct, when the *but*-conjunct does not assert the falsity of a proposition in the first conjunct.

(6) On totta että (minä) en puhu ranskaa, mutta (minä) luen sitä mielelläni.
is true that I not speak French but I read it with-pleasure

(7) Mou-cho (ngo) m-sik gong faat man , daanhai (ngo) soeng hok .
No-mistake I not-know speak French, but I want learn
'It's true that I don't speak French, but I want to learn it.'

(8) and (9), from Cantonese and Telugu respectively, show that it makes no difference if the antecedent is the main clause subject in the first conjunct: Pro-drop in the *but*-conjunct is still impossible.

(8) Siuming waa (keoi) m-sik gong faatman daanhai *(keoi) sik.
Siuming say he not-know speak French but he know
'John says he doesn't speak French, but he does.'

(9) Prasaad tanu Hindii maaTlaaDanu ani anTaaDu, kaanii *(tanu) maaTlaaDataaDu.
Prasaad he Hindi not.speaks that says but he speaks
'Prasaad says that he doesn't speak Hindi, but he does.'

(10), from Finnish, shows that the presence of *but* and assertion of the falsity of a preceding proposition are not crucial.

(10) Ne sanoo että (minä) puhun ranskaa, **ja** *(minä) puhun.
they say that I speak-1SG French and I speak-1SG
'They say that I speak French, and I do.'

The second conjunct in (10) does not contradict a proposition in the first conjunct, but instead confirms it. Since confirmation as well as denial of the truth of a preceding

proposition involves focusing polarity, this shows that that is, indeed, the crucial property which excludes pro-drop in the second conjunct.

We can thus formulate the following generalization:

(11) **Generalization 1**: When polarity is focused, the subject doesn't drop, in a class of null-subject languages.

The question is why polarity focus has this effect, and why it has this effect in some languages but not others.

4. A cross-linguistic generalization

The following generalization appears to hold, by and large.

(12) **Generalization 2:** In A-languages a YNQ is standardly answered affirmatively by a special affirmative particle. In B-languages a YNQ is standardly answered affirmatively by repeating the finite verb of the question.

In (13), Y means 'YNQ answered affirmatively by a special particle' (Y suggesting 'yes'), while V means 'YNQ answered affirmatively by repeating the finite verb of the question'.

(13)

A-languages		**B-languages**	
Amharic	Y	Cantonese Chinese	V/Y
Arabic	Y	Czech	V/Y
Greek	Y	Finnish	V/Y
Hebrew	Y	Indonesian	V
Hungarian	Y/V	Malayalam	V/Y
Italian	Y	Mandarin Chinese	V/Y
Japanese	Y	North Saami	V/Y
Korean	Y/V	Polish	V/Y
Persian	Y	Portuguese	V/Y
Spanish	Y	Telugu	V/Y
Turkish	Y/V		
Welsh	V/Y		

This shows a tendency, although the facts are not completely clear. All the B-languages have V as the most prominent alternative (usually given first by informants). However, all of them except Indonesian have Y as another alternative. At least for some of them, possibly all of them, the choice between V and Y is not free. Instead, in certain constructions Y is the only alternative, probably for principled reasons; see Holmberg 2001. Thus the fact that they have both V and Y does not in itself jeopardize Generalization 2.

Furthermore, some of the A-languages can reply by repeating the finite verb, as an alternative to the yes-particle. This is potentially a more serious problem for the theory to be articulated below, and the associated explanation of the division between A and B-languages.

In fact, most or perhaps all of the A-languages may reply to a YNQ by repeating the finite verb or auxiliary in the same way as you can do in English, either in conjunction with *yes* or without.[3]

(14) – Do you speak French?
a. – Yes.
b. – Yes, I do.
c. – I do.

In a pro-drop language the subject pronoun may be dropped. If the language in question then has V-to-I movement, as many, perhaps even all pro-drop languages do, and if it has VP-deletion, then the counterpart to (14c) will consist of just a finite verb.

This is, however, not how replies to YNQs are derived in the B-languages. Instead, as will be demonstrated below taking Finnish as the example, in the B-language YNQ-replies are derived by moving the finite verb to C and deleting IP, thus typically leaving only the finite verb to be spelled out. My conjecture is that the Finnish V-reply and the Korean or Turkish V-replies to a YNQ have quite different derivations: The Finnish reply is derived by IP deletion, the Korean one by VP-deletion and a null subject.

Welsh looks like a genuine exception, though. Welsh is clearly a 'V-language' as regards answers to YNQs; see Jones 1999. Yet the pronoun can be dropped in the Welsh counterpart to (1).

(15) a. – ydy hi wedi gorffen? (Welsh: Jones 1999)
be-PRES she PERF finish
'Has she finished?'
– ydy.
'Yes.'

b. Maen nhw'n dweud nad yw John yn siarad Ffrangeg,
be-3PL they PROG say NEG.COMP be-3SG John PROG speak French
ond mae yn.
but be-3SG PROG
'They say that John doesn't speak French, but he does.'

This indicates that 'A-language' is not a well-defined notion. This will be discussed in the final section, together with a qualification and an explanation of the Welsh exception.

5. A note on Portuguese

That European Portuguese (E. Portuguese[4]) does not allow a null subject in contexts such as (1) has not been observed before, to my knowledge. It is, however, well

[3] The fact that only Hungarian, Korean, and Turkish are marked as Y/V in (13) may thus be due to incomplete or flawed data collection.

[4] Brazilian Portuguese is not a null-subject language, and therefore not relevant in the present connection.

known that Portuguese differs from most, or all, of its Romance relatives except Galician in the way YNQs are answered by repeating the verb rather than using an affirmative particle; see Martins 1994.

(16) – Viste o Joâo? (E.Portuguese)
'Did you see John?'
– Vi. / #Sim.
saw / yes
'Yes.'

(17) – Viste a Juan? (Spanish)
'Did you see John?'
– Sí. / *Vi.
yes / saw
'Yes.'

The fact that E.Portuguese as the only Romance null-subject language checked so far is a B-language is therefore highly suggestive of a connection between these two properties.

I don't know how to account for the split between E.Portuguese informants who accept pro-drop in (1) and those who don't. I don't know whether it corresponds to a regional division, nor do I know whether it correlates with how they reply to YNQs. All speakers consulted can reply by repeating the finite verb; what has not been investigated is whether there are subtle differences indicative of a different grammar of such replies among different speakers.

However, as a number of speakers have independently confirmed that they don't accept pro-drop in (1), I am confident that there is a variety of E.Portuguese, which may or may not be the majority variety, which is a B-language.[5]

6. The syntax of affirmation and negation

In this section I will lay the groundwork for an explanation of Generalization 2, that is the correlation between being a B-language and 'being a V-language', replying affirmatively to YNQs by repeating the finite verb. It is heavily based on the theory of replies in Holmberg 2001.

6.1. English

Consider first English: If you want to contradict a negative statement, you can't do it very well by simply uttering 'yes'. Instead you need the longer form which repeats at least the subject and the finite auxiliary of the statement, in affirmative form.

[5] Galician, which is spoken in Spain just north of Portugal, may present another counterexample to Generalization 2, along with the variety of Portuguese which accepts pro-drop in (1): Apparently the subject can be dropped in the second conjunct of (1), although YNQs are standardly answered by repeating the verb.

(18) – John doesn't speak French.
a. – #Yes.
b. – Yes he does.

The reason why you need the longer form is quite clear, once we acknowledge that a reply 'Yes' or 'No' is an elliptical expression where an entire IP is elided, and recovered from the preceding yes/no-question.

(19) – Does John speak French?
– Yes [$_{IP}$ ~~John speaks French~~]

The long reply (18b), which is an acceptable alternative in this context, is derived by VP-ellipsis, the VP recovered from the preceding question.

(20) – Does John speak French?
– Yes [$_{IP}$ he does [$_{\text{~~VP~~}}$ ~~speak French~~]]

In (21), which is = (18), the antecedent IP is, however, specified for negative polarity.

(21) – John doesn't speak French.
– *Yes [$_{IP}$ ~~John doesn't speak French~~]

The short reply is therefore contradictory: an affirmative focus operator combined with a negative proposition. The long reply *Yes he does* is fine: As in this case all that needs to be recovered from the preceding utterance is the VP, it doesn't matter whether the sentence is negative or affirmative.

(22) – John doesn't speak French.
– Yes he does [$_{VP}$ ~~speak French~~].

6.2. Finnish

In Finnish you reply affirmatively to a yes/no-question by repeating the finite verb of the question.

(23) – Puhuu-ko Joni ranskaa?
speaks-Q John French
'Does John speak French?'

– Puhuu.
speaks
'Yes.'

The reply *Puhuu* 'speaks' is an elliptical expression where the verb, incorporated in Pol(arity) has moved to the C-domain, and IP is deleted and recovered from the the preceding utterance. This is a somewhat simplified analysis; the more precise analysis will be presented below in section 7.

(24) puhuu C [$_{IP}$ ~~Joni *t* ranskaa~~] (IP = PolP)
speaks John French

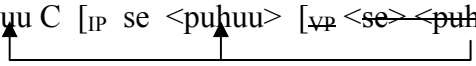
We know that IP is deleted in the reply in (24), because Finnish does not allow pro-drop of 3rd person subjects (see Vainikka & Levy 1999, Holmberg, to appear). Therefore the subject in the reply in (23) must be deleted as part of a larger constituent.

An alternative affirmative reply is

(25) – Puhuu se.
speaks he
'Yes, he does.'

This is derived by moving verb+Pol to the C-domain, and deleting VP, recovered from the preceding utterance (*puhu* is the root form, *puhuu* is Present 3SG). (Copies to be deleted are put within angled brackets.)

(26) puhuu C [$_{IP}$ se <puhuu> [$_{\sout{VP}}$ ~~<se> <puhu> ranskaa~~]]



That is to say, (25) corresponds to the English long reply. The difference is (a) that Finnish but not English has V-movement out of VP, and (b) that Finnish expresses affirmative focus ('reply focus') by moving Pol with the incorporated verb to the C-domain, while English merges an affirmative focus particle *yes* in the C-domain.

As predicted, given the English facts discussed, when contradicting a negative statement in Finnish you can't use the short form (the bare verb).

(27) – Joni ei puhuu ranskaa.
John not speaks French
a. – #Puhuu.
speaks
b. – Puhuu se.
speaks he
'Yes he does.'

The reason is that the short form is derived by IP-ellipsis, but in this case, since the antecedent IP in the preceding utterance is specified for negative polarity, the result is ill formed, just as in the English case (21).

(28) – Joni ei puhu ranskaa.
John not speaks French
– *Puhuu [$_{\sout{IP}}$ ~~Joni ei puhu ranskaa~~]
speaks John not speaks French

The long form, on the other hand, is derived by VP-deletion, so the polarity of the preceding sentence doesn't matter. Consequently it is well formed in this context.

(The middle copy of the verb is obviously deleted as well, by ordinary copy deletion, before the derivation reaches PF.)[6]

(29) Puhuu [$_{IP}$ se <puhuu> [$_{\sout{VP}}$ ~~se <puhu> ranskaa~~]
speaks he speaks he speak French
'Yes he does.'

6.3. Cantonese Chinese

Cantonese Chinese behaves in relevant respects like Finnish and English. There is a short and a long form of the reply to a YNQ:

(30) – Siuming sik msik gong faatman?
Siuming know not-know speak French
'Does Siuming speak French?'

– Sik.
know
'Yes.'
– Sik, keoi sik.
know he know
'Yes, he does.'

Both reply forms are derived by moving the verb to the C-domain. In the short form this is followed by IP-ellipsis, in the long form by VP-ellipsis.

If you want to contradict a negative statement in Cantonese, you cannot use the short form, but you can use the long form.

(31) – Siuming m-sik gong faatman.
Siuming not-know speak French
'John doesn't speak French.'
– #Sik.
– Sik, keoi sik.

As in Finnish and English, the short form is derived by IP-deletion, but as the antecedent IP is specified for negative polarity, the result when the IP is recovered is ill formed, yielding a contradictory reading. The long form, on the other hand, is

[6] Using the long form (29) is not the only way to contradict a statement. Another one is using a special contradiction-focus particle *–pas/päs*, affixed to the fronted verb:
(i) – Joni ei puhu ranskaa.
John not speaks French
– Puhuupas.
speaks-PAS
'He does, too.'
Here, just as in the reply in (28) the IP recovered from the preceding utterance is specified for negative polarity. However, the effect of this focus particle is to neutralize the polarity of the proposition it c-commands (see Holmberg 2001).

derived by VP-deletion, so the polarity of the utterance which provides the antecedent does not matter.

(32) sik C [$_{IP}$ keoi <sik> [$_{VP}$ <~~sik~~> [$_{VP}$ <~~keoi> gong faatman~~]]]
know he know speak French

The only relevant difference between Finnish and Cantonese is that the middle copy, that is the copy of V+Pol, is deleted in Finnish, but spelled out in Cantonese.[7]

7. Null subjects and polarity

As mentioned, the second conjunct of (1), repeated here, asserts the falsity of a negative statement in the first conjunct, which is to say that it focuses affirmative polarity; everything else is presupposed.

(1) They say that John doesn't speak French, but he does.

Being presupposed, the VP can be deleted. On that account the subject could also be deleted, but English is not a null-subject language, and thus requires a pronounced subject, regardless of information-structural content.

The problem that we started out with is why a class of null-subject languages also resist deletion (non-pronunciation) of the subject in this context.

(33) Ne sanoo etten puhu ranskaa, mutta *(minä) puhun. [8]
(Finnish)
they say that-not-1SG speak French but I speak

Taking Finnish as my example, the answer is that a verb pronounced without a subject in a context forcing polarity focus is necessarily analyzed as moved by Pol-movement to C, with IP-deletion. However, much as in the case of the discourses in (18), (27), and (31), where a speaker contradicts a negative statement made by another speaker, this does not yield a reading that can be made sense of, as the only available antecedent, the embedded statement in the first conjunct, is specified for negative polarity. (As will be clarified below, (34) is not just contradictory, but also violates Full Interpretation at LF.)

(34) *Ne sanoo etten puhu ranskaa, mutta puhun [~~$_{IP}$ minä en puhu ranskaa~~]
they say that-not-1SG speak French but speak I not speak French

[7] When comparing Chinese with, for example, Finnish there is the added complication that the affirmative particle (in Cantonese *hai*) can be used to confirm a negative statement or a negative question, which is not possible in Finnish or English. This does not impact on the discussion in the text, however.

[8] In Finnish negated finite sentences subject-verb agreement is marked on the negation; see Holmberg & al. 1993, Holmberg 2003. The negation, in turn, can be cliticized to the complementizer. None of this is of any consequence in the present connection. The reason why I use a 1st person subject in the Finnish example is that Finnish doesn't have pro-drop in the 3rd person, except in some restricted cases.

Pronouncing the pronoun in (33) excludes the analysis where V+Pol is moved to C, and forces the analysis where only VP is deleted. This yields a well formed reading, since the first conjunct provides a good antecedent for the deleted VP.

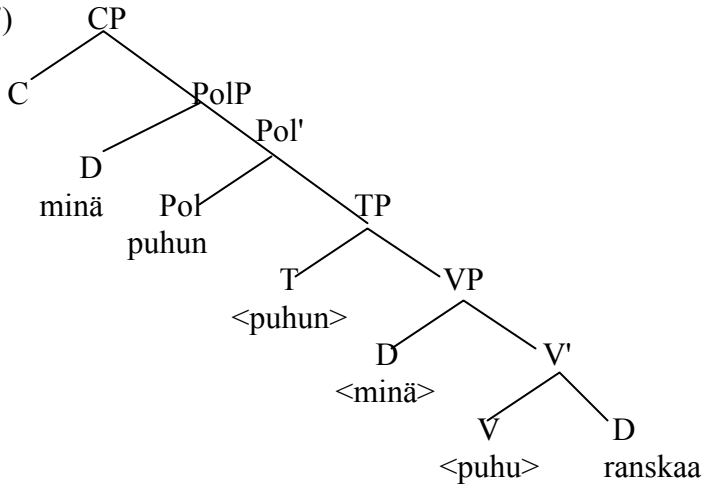
(35) mutta [$_{CP}$ C [$_{IP}$ se puhuu [~~$_{VP}$~~ ~~<se> <puhu> ranskaa~~]]]
but he speaks he speak French

Why is the subjectless second conjunct necessarily analyzed as having V in C, though? This could only be the effect of a **competition of derivations**: There are in principle two derivations leading to (36).

(36) Puhun.
speak-1SG

Starting from the tree (37) where the verb has moved to Pol, via T, and the subject has moved to specPolP, there are two ways to continue the derivation. The first one is (38). (For ease of presentation I have omitted parts of the structure such as little *v* and possible additional copies of the subject; the analysis (37) will in any case be modified in section 8).

(37)



(38) 1. Delete VP;
2. Delete the subject.

As discussed in Holmberg 2001, Finnish has VP-deletion. Since Finnish always does V-raising out of VP, the raised verb is stranded by VP-deletion (as is the case in many languages; see for example Huang 1991). Furthermore, Finnish has 1st and 2nd person null subjects. In Holmberg (to appear) I argue that the null subjects are derived by deletion of a pronoun, not merge of *pro*, but this is (probably) not crucial in the case at hand. Consequently the derivation (38) is predicted to be fine. The result is (39), where, after deletion of the lower copy of raised verb by ordinary copy-deletion, the only word that is actually pronounced is the copy of the raised verb *puhun* in Pol (main clause C has no lexical form).

(39) C [$_{PolP}$ ~~minä~~ [$_{Pol'}$ puhun [$_{TP}$ ~~puhun~~ [$_{VP}$ ~~minä puhu ranskaa~~]]]]
I speak-1sg speak-1sg I speak French

The other derivation is (40):

(40) 1. Move Pol to C;
2. Delete PolP.

The result is (41):

(41) puhun+ C [~~$_{PolP}$ minä~~ [$_{Pol'}$ ~~puhun~~ [$_{TP}$ ~~puhun~~ [$_{VP}$ ~~minä puhu ranskaa~~]]]]

The derivation (40) always wins over (38), for reasons discussed below. This is fine, since the YNQ provides a good antecedent for the deleted PolP. Consequently (42) is a well formed discourse:

(42) –Puhutko ranskaa?
speak-2SG French
'Do you speak French?'

– Puhun.
speak-1SG
'Yes.'

However, the derivation (40) gives an ungrammatical result as the derivation of a sentence asserting the falsity of a proposition, whether it is in a separate utterance, as in (28), or in a preceding conjunct, as in (43).

(43) *Ne sanoo etten puhu ranskaa, mutta puhun.
they say that-not-1SG speak French but speak-1SG

The reason is that the context does not provide a good antecedent for the deleted IP/PolP.

Let us be more precise at this point: The problem in (42), the second conjunct derived as in (40), is not, as implied earlier in connection with (34), that the antecedent of the elided IP is specified for negative polarity, but that it is specified for polarity at all. Note that (44) is as ill formed as (43), without a pronounced pronoun in the second conjunct.

(44) Ne sanoo että puhun ranskaa, ja *(minä) puhun.
they say that speak-1SG French and I speak-1SG

Here the second conjunct is intended to confirm the truth of the embedded proposition in the first conjunct. The antecedent IP is specified for affirmative polarity, so the two conjuncts have the same polarity, still the result is ill formed when derived as in (40), by Pol-raising to C and deletion of IP, the derivation which is forced when the subject is not pronounced.

The reason is the following: C in (40) is a Focus-C, encoding an abstract polarity focus operator. This operator needs a polarity variable to bind, namely an unspecified Pol head, a variable with two possible values, negative or affirmative. Now consider the reply in (42). Mainly for the sake of presentation I will view it from the perspective of the listener/parser deriving the LF of the elliptical reply, and recovering the ellipsis by copying the PolP of the preceding YNQ (45a) into the position of the deleted PolP in the reply:[9] The result is (45b):

(45) a. Puhut C_Q $[_{IP}$ D_i $[_{Pol}$ Aff/Neg] $[_{VP}$ D_i puhu ranskaa]]

b. Puhun C_{POLFOC} $[_{IP}$ D_i $[_{Pol}$ Aff/Neg] $[_{VP}$ D_i puhu ranskaa]]

In the question, C encodes a YNQ-operator, which must be merged with a PolP with unspecified Polarity, a variable with the two possible values Affirmative or Negative.[10] In the reply, C encodes a polarity-focus operator. Both kinds of C attract Pol to specCP (see next section for details of this movement). The PolP of the reply is a copy of the PolP of the question, consequently it contains the required polarity variable. As the Pol head moved to specCP is affirmative-marked, the focus-operator will assign Affirmative as the value of the variable; see Holmberg 2001 for more details.

The subject in (44) is labelled D_i, because the resolution of the ellipsis operates on a representation, call it LF, where the (pronominal) form is not relevant but only its reference, which is the same person in the question and the reply.

Now consider the resolution of the elliptical second conjunct in (43) or (44). Taking (44) as our example, (46a) represents the LF of the embedded PolP *että puhun ranskaa* in the first conjunct, specified for affirmative polarity.

(46) a. C $[_{PolP}$ D_i $[_{Pol}$ Aff] $[_{VP}$ D_i puhu ranskaa]]

b. puhun C_{POLFOC} $[_{PolP}$ D_i $[_{Pol}$ Aff] $[_{VP}$ D_i puhu ranskaa]]

(46b) is the result when this PolP is copied into the place of the ellipsis in the second conjunct. There is a polarity focus operator but no polarity variable to bind. The sentence ends up violating Full Interpretation.

The question that remains to be answered to complete the argument, is: Why does the derivation (40) always win over the derivation (38)? On the face of it, from the point where they diverge, they need the same number of steps to derive the same PF output, the bare main verb. However, we are dealing with a context which induces polarity focus. This means that C_{POLFOC} must be merged. In Finnish C_{POLFOC} triggers movement of Pol to C. If so, then the derivation where VP and the subject are deleted by separate operations also involves movement of Pol to C. The derivations that are

[9] An alternative is to view it from the point of view of the producer of the ellipsis, in which case the IP-ellipsis may be seen as blocked in the absence of an antecedent satisfying the identity conditions on the deletion. See Holmberg (2001) for some discussion of the implications of Finnish IP-deletion for theories of ellipsis.

[10] The relation between the LF and the PF of a YNQ is particularly transparent in Chinese: see (30), where the polarity variable is spelled out as *V-not-V*.

compared are thus (47) and (487), where (487) is less complex as it involves fewer operations.

(47) 1. Move Pol to C.
2. Delete the subject.
3. Delete VP.

(48) 1. Move Pol to C;
2. Delete PolP.

We can thus explain why (43), the Finnish counterpart to (1) with a null subject, repeated here, is ill-formed:

(43) *Ne sanoo etten puhu ranskaa, mutta puhun.
They say that-not-1SG speak French but speak

The second conjunct can only be derived by movement of Pol to C, with deletion of IP (PolP). However, since the sentence is headed by a polarity focus operator, the deleted PolP requires an antecedent which contains a polarity variable. The antecedent here is the embedded sentence in the first conjunct, which has specified polarity. Consequently the sentence ends up violating FI, with a focus operator having no variable to bind.

I conjecture that the explanation for the absence of pro-drop in the second conjunct of (the counterparts of) (1) is the same in all the B-languages.

Why, then, do the A-languages allow pro-drop in their counterparts of (1)? The simple answer is that they don't have Pol-to-C movement, as shown by the fact that they mostly reply affirmatively to YNQs by using an affirmation particle. Therefore, to derive the second conjunct in (1), they can employ derivation (38), deleting VP (if they allow that deletion) and deleting the subject by separate operations.

In the next section I will present in more detail the derivation of polarity focus sentences, along the lines of Holmberg 2001. This will shed light on some outstanding issues, including some of the variation found among the A-languages

8. The derivation of replies in more detail

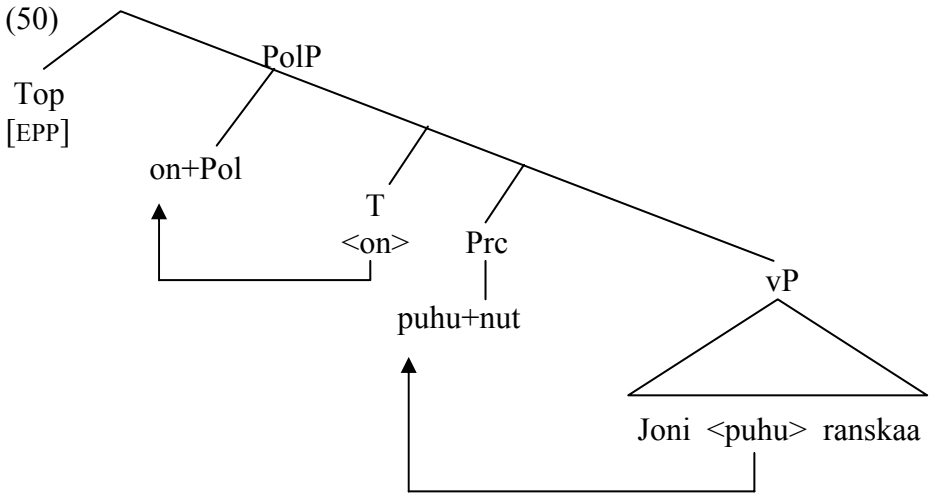
Not just a single verb, but a string of auxiliaries and verb, can move to C in replies.

(49) – Onko Joni puhunut ranskaa?
has John spoken French
– On.
has
– On puhunut (Joni ranskaa).
has spoken (John French)
'Yes.'

As argued in Holmberg (2001), this means that Pol-to-C in a case like this is not head movement, but movement of a constituent including the auxiliary verb and the main

verb, but excluding the subject and the object:[11] a remnant PolP. The derivation is as follows:

(50)

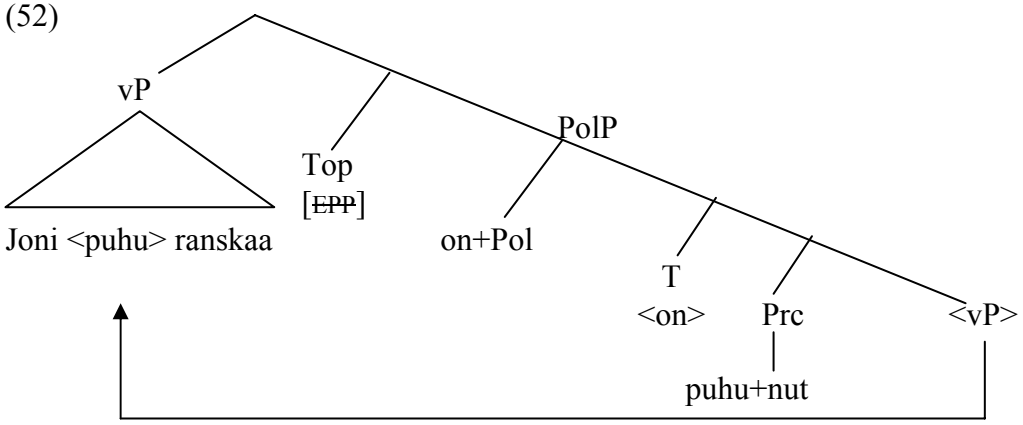


In (50) the verb has moved to the Participle head merged with vP. The auxiliary, an exponent of T, is merged with PrcP. Pol is merged with TP, and T moves to Pol. Finally a head Top(ic) is merged. This head has an EPP-feature. In Finnish, [EPP, Top] usually triggers movement of the subject NP, but may trigger movement of other topic-worthy categories such as an object or a locative adverbial; see Holmberg & Nikanne 2002. It may also trigger movement of the remnant vP. This can be seen in (51) (CON = conditional mood).

(51) Milloin [$_{TopP}$ [$_{vP}$ Joni ranskaa] olisi puhunut?
when John French have-CON spoken
'When would John have spoken French?'

Applied to (50), vP-movement yields (52): [12]

(52)



[11] It may also contain certain adverbs, namely adverbs high in the Cinque hierarchy; see Cinque 1999. This is explained under the theory summarized in the text below. See Holmberg 2001.

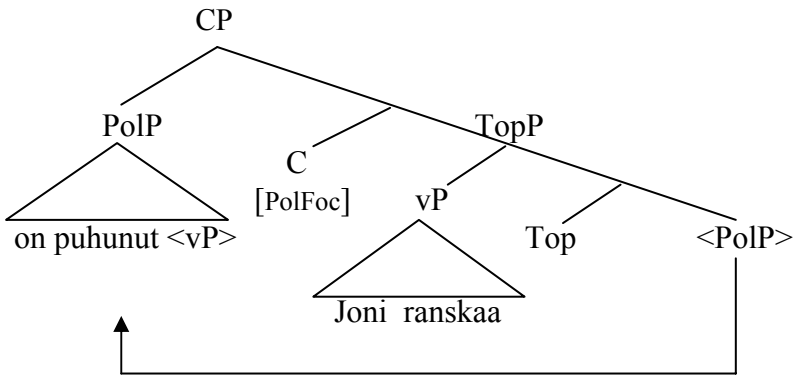
[12] As shown in Holmberg 2001 the vP moved to specTopP also potentially contains adverbials low in the Cinque-hierarchy, such as manner adverbials and aspectual adverbials.

Next, C_{POLFOC} is merged. C_{POLFOC} triggers movement of Pol to C. Finnish has the option of moving just the head Pol or the entire PolP. The former will derive the word order (52):

(52) On Joni ranskaa puhunut.
has John French spoken
'John has (indeed) spoken French.'

PolP-movement yields (53):

(53)



This structure can be spelled out as (54a), or, if TopP is deleted, (54b).

(54) a. On puhunut Joni ranskaa.
has spoken John French

b. On puhunut.
has spoken

Both are fine as replies to the question 'Has John spoken French?'. None of them are acceptable as the second conjunct of (1), for the reasons articulated in the previous section.

What is the structure of the well-formed second conjunct of (1), with an overt subject?

(55) Ne sanoo etten puhu ranskaa, mutta **minä** puhun.
they say that-not-1SG speak French but I speak-1SG

At this point I will diverge from Holmberg 2001. In that paper I assumed, following Vilkuna (1995) and Holmberg & Nikanne (2002), that Finnish is a topic-prominent language in the sense that any category can satisfy the EPP of the finite clause provided it is interpretable as a topic. However, as discussed in Holmberg (to appear), the subject need not be a topic to check the EPP. It can, for example, be an indeterminate pronoun.

(56) Kuka tahansa voi oppia puhumaan ranskaa.
who ever can learn speak-INF French
'Anyone can learn to speak French.'

I suggest that the subject, in the unmarked case, moves to specPolP. This tallies with the fact that Pol is where subject agreement is located in Finnish; see Holmberg 2003. The 'unmarked case' is when the subject is not contrastive focus or a whP (in which case it moves to specCP), or is not topicalized along with the vP.

In the derivation of the second conjunct in (55) the subject pronoun moves to specPolP. Then Top is merged, triggering movement of vP, now containing only the object.

(57) $[_{TopP}\ [_{vP}$ <minä> <puhu> ranskaa] [Top $[_{PolP}$ minä [puhun+Pol [<vP>]]]]]

Then C_{POLFOC} is merged, triggering movement of PolP, now including the subject and the verb. The structure so derived is (58) (omitting some copies for ease of exposition).

(58) $[_{CP}\ [_{PolP}$ minä puhun <vP>] [C $[_{TopP}\ [_{vP}$ ranskaa] [Top [<PolP>]]]]

This structure can be spelled out as (59a) or (59b):[13]

(59) a. minä puhun ranskaa.
I speak French

b. minä puhun.
I speak

(59b) is when vP is deleted. It is important that this should be VP-ellipsis, not TopP-ellipsis or PolP-ellipsis, even if no constituent inside TopP is actually pronounced. It is important, since otherwise our explanation of why a null subject is not well formed in (1) in B-languages is lost. The fact that the (highest copy of) the subject is not included in the ellipsis ensures that it need not be (perhaps cannot be) ellipsis of anything bigger than vP. The category which needs to be recovered from the context is a vP.

Finally, and equally importantly, the reason why the subject can't be deleted in (59) is that it is part of a focused PolP.

This theory makes the prediction that if another category than PolP can check C_{POLFOC}, then the subject can be deleted. Consider (60):

(60) Ne sanoo etten puhu ranskaa, mutta itse asiassa (minä) puhun.
they say that-not-1SG speak French but in fact I speak

Here the affirmative polarity adverbial *itse asiassa*, roughly 'in actual fact', moves to specCP. The subject remains in specPolP, and can be deleted, as can the vP.

[13] In the context we are talking about, (59a,b) have focus on Pol, hence the verb is stressed.

In this light, consider the Italian example (5c). It seems clear enough that the structure is the same as in (60): The adverbial *invece* 'in reality' is in specC_{POLFOC}, hence nothing prevents dropping the subject.

This may also explain the counterexample posed by Welsh: In Welsh, YNQs are standardly replied by repeating the finite auxiliary. Yet, as shown in (15), the subject can be dropped in the Welsh counterpart to (1). However, it can do so only if the sentence includes a focused PROG particle (focus indicated by capitals in (61a)).

(61) Maen nhw'n dweud nad yw John yn siarad Ffrangeg,
'They say that John doesn't speak French,
a. ond mae YN.
but be-3SG PROG
but he does.'
b. *ond mae.

If the focused PROG particle in (61a) is able to check C_{POLFOC}, then the subject may remain in specPolP, and be deleted by usual pro-drop. Thus (61) need not involve deletion of a constituent bigger than VP, and is therefore well formed in this context. In (61b) PolP itself, headed by the finite verb, checks C_{POLFOC}, which leads to a violation of FI in this context. The precise analysis of (61a) is still an open question, though.

References

Ariel, M. 2001. Accessibility theory: an overview. *Text representation: linguistic and psycholinguistic aspects,* ed. Sanders, T., Schilperoord, J. & Spooren, W. Amsterdam: John Benjamins.

Cinque, Guglielmo. 1999. *Adverbs and functional heads – a cross-linguistic perspective*. Oxford: Oxford University Press.

Holmberg, A. 1998. Word order variation in some European SVO languages: a parametric approach. *Constituent order in the languages of Europe*, ed. Anna Siewierska, 553-598. Berlin: Mouton de Gruyter.

Holmberg, A. 2001. The syntax of yes and no in Finnish. *Studia Linguistica* 55: 141-174.

Holmberg, A. 2003. Yes/no questions and the relation between tense and polarity in English and Finnish. *Linguistic variation yearbook* 3, ed. Pierre Pica, 43-68. Amsterdam: John Benjamins.

Holmberg, A. To appear. Is there a little *pro*? Evidence from Finnish. To appear in *Linguistic Inquiry* 36.

Holmberg, Anders, Urpo Nikanne, Irmeli Oraviita, Hannu Reime and Trond Trosterud. 1993. The structure of INFL and the finite clause in Finnish. Anders Holmberg and Urpo Nikanne (eds.) *Case and other functional categories in Finnish syntax*. 175-206. Berlin: Mouton de Gruyter.

Holmberg, Anders and Urpo Nikanne. 2002. Expletives, subjects and topics in Finnish. *Subjects, expletives, and the EPP*, ed. Peter Svenonius, 71-105. Oxford and NewYork: Oxford University Press.

Huang, C.-T. James. 1991. Remarks on the status of the null object. *Principles and parameters in comparative grammar*, ed. R. Freidin, 56-76. Cambridge Mass.: MIT Press.

Jones, Bob M. 1999. *The Welsh answering system*. Berlin: Mouton de Gruyter.

Martins, Ana Maria. 1994. Enclisis, VP-deletion and the nature of Sigma. *Probus* 6: 173-205.

Vilkuna, Maria. 1995. Discourse configurationality in Finnish. *Discourse-configurational languages,* ed. Katalin É. Kiss, 244-268. Oxford: Oxford University Press.

Vainikka, Anne & Yonata Levy. 1999. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory* 17: 613-671.

School of English Literature, Language and Linguistics
University of Newcastle
Newcastle upon Tyne NE1 7RU
UK

anders.holmberg@newcastle.ac.uk