

SUBJECT EXTRACTION IN ENGLISH:
SOME PROBLEMS OF INTERPRETATION

AIMO SEPPÄNEN AND GUNNAR BERGH

University of Gothenburg, Sweden

1. *Introduction*¹

In recent discussion of English syntax, the question of restrictions governing the extractability of the subject has generally been seen as a topic of great theoretical interest, and has therefore received a great deal of attention, especially in theoretically-oriented descriptions of the language. This work has concentrated on present-day English, and left the history of the phenomenon almost totally neglected, apart from the observation that Old English on this point of its structure was clearly different from present-day English (cf. Allen 1980, Pesetsky 1982).

In a full-length corpus study of the development of English extraction patterns from the earliest stages of the language down to the present day (Seppänen – Bergh, forthcoming), we have charted the history of extractions, including subject extractions, and have offered a few preliminary results of the study in a recent article (Bergh – Seppänen 1994). The main findings of our study as far as subject extraction is concerned can be briefly represented as follows.

Subject extraction has always been possible in English, and in Old English could indeed result in the *that*+gap sequence noted by Allen and Pesetsky, as illustrated in (1):

- (1) a. *Sio anlicnes wæs gecueden ðæt Δ sceolde bion on ðæs*
an illustration was said that Δ should be on the

¹ We are grateful to Olga Fischer, Gerald Gazdar, Lilo Moessner, Christer Platzack, Robert Stockwell, Tarald Taraldsen, Johan van der Auwera, Ans van Kemenade, and Wim van der Wurff for comments on an earlier version of the paper, and to Hans Götzsche for help with the Danish material. Needless to say, none of the persons mentioned are responsible for any of the views expressed in our article.

sacerdes hrægle ða readan appla ongemang ðam bellum.
 priest's robe the red apples among the bells
 (Pastoral Care)

'It was said that as an illustration there were to be red apples on the priest's robe among the bells'

- b. *Hu mægþ yfel beonþte ælcas monnes ingeþanc wendþ.*
 how may that evil be that each man's mind thinks
 þte Δ good sie (King Alfred's Boethius)
 that Δ good be

'How can that be evil that each man in his mind thinks is good?'

- c. *Ne bebode nan cristen man oðrum buton þæt he wille*
 not ask any Christian man others but what he wants
 þæt Δ him sylfum geboden si (Wulfstan's Homilies)
 that Δ him self asked be

'Let no Christian man offer others anything but what he himself wants to have offered to him'

Yet, during the whole of the Old English and Middle English period, there was a growing tendency to avoid that sequence and to replace it by one of two alternative forms. The more common of them was the zero+gap pattern, as in (2):

- (2) a. *ðæt me ðincð Δ sie þ hehste good, þætte mon*
 that me seems Δ be the highest good, that man
ne ðurfe nanes oðres godes (King Alfred's Boethius)
 not needs no other good
 'That it seems to me is the highest good that no man needs any other good (besides it)'

- b. *ða getriewan friend þonne ic secgge Δ sie ðæt*
 the true friend again I say Δ be the
deorwyrðeste ðing ealra þissa weoruldgeslða
 dearest thing of all this worldly prosperity
 (King Alfred's Boethius)

'The true friend I say again is the most precious of all these worldly things'

- c. *þus beo zewurðod se man, þe se cyninȝ wile*
 thus be honoured the man, who the king wants
 Δ wurðscipe hæbbe (Assman's Homilies)
 Δ honour have

'Let the man be thus honoured whom the king wants to have honour'

This structure was favoured even at a time when deletion of the complementizer was not normal outside of extractions, and it became the dominant pattern after 1300 and the only living form in the written standard after 1500 (cf. Bergh – Seppänen 1994:134-138).

The second way of avoiding the increasingly awkward *that*+gap structure was

to use a resumptive pronoun in the position of the extracted element, as shown by the examples in (3):

- (3) a. *þa tunglu þu gedest þe gehyrsume, 7 þa sunnan þu*
 the stars you make thee obedient and the sun you
gedest þ heo mid heore beorhtan sciman þa þeostro
 make that she with her bright light the darkness
adwæscð þære sweartan nihte (King Alfred's Boethius)
 dispell of-the black night

'The stars you make obedient to you, and the sun you cause to dispell the darkness of the black night with her bright light'

- b. *Hwæðer þe nu pince þ se mon micelne anwald*
 whether to-you now seems that the man much authority
hæbbe þe him selfum ðincð þ he nænne næbbe, ...?
 have who him self seems that he none not-have
 (King Alfred's Boethius)

'Does it now seem to you that that man has much authority who himself feels that he has none?'

- c. *we mazon cweþan þæt se wæremid him sylfum, þe zesæd is*
 we may say that he werewith him self, who said is
on ðam zodspelle, þæt he fæder forlete (Gregory's Dialogues)
 in the gospel, that he father abandoned
 'We may say that he is alone who is said in the gospel to have left his father'

This device was brought in especially in the case of subject extraction, and less commonly in non-subject position, but it always remained a marginal pattern and fell into disuse in the standard language after 1500.

Taking as our starting-point the early history of the subject extraction patterns as summed up in the above remarks, we wish here to continue the discussion by addressing the more general issues connected with the development of the construction. More specifically, we shall here consider two directly interrelated questions: (i) Why was the *that*+gap sequence of (1) avoided and subsequently totally banned? (ii) What is the proper analysis of the structure of (2) that ultimately took its place?

2. Subject extraction and null subject/non-null subject languages

To view the history of the *that*+gap pattern against a wider background, consider the difference in the status of the following two sentences in English, illustrating the extraction of subjects and non-subjects, and comparing it with the lack of such a contrast in their Italian equivalents.

- (4) a. *What do you think that Mary has written Δ to him?*
 b. **Who do you think that Δ has written this letter?*

- (5) a. *Che credi che Maria gli abbia scritto* Δ?
 what think-you that Maria to-him have written Δ
 'What do you think that Maria has written to him?'
 b. *Chi credi che Δ abbia scritto questa lettera?*
 who think-you that Δ have written this letter
 'Who do you think has written this letter?'

The difference in subject extraction between the two languages reappears in a similar contrast where Italian can dispense with a pronominal subject (an anaphoric pronoun or a dummy) which must be overtly expressed in English, as indicated by (6) and (7) respectively:

- (6) a. *Sandra non a potuto lavorare perchè Ø è stata malata.*
 Sandra not has been-able work because Ø is been ill
 b. *Dove è Gianni? Δ Non siamo dove Ø è.*
 where is Gianni Δ not know-we where Ø is
 c. *Ø È pericoloso sporgersi.*
 Ø is dangerous lean-out
 d. *Ø Piove.*
 Ø is-raining
- (7) a. *Sandra has not been able to work because she has been ill.*
 b. *Where is Gianni? We don't know where he is.*
 c. *It is dangerous to lean out.*
 d. *It is raining.*

Noting this correspondence, Perlmutter (1971) suggested a generalization which in some form or other has figured in much later work (cf. Chomsky – Lasnik 1977, Pesetsky 1982, Platzack 1985, Rizzi 1982, 1990): while languages which generally allow null subjects, like Italian and Spanish, also admit extraction of subjects out of complement clauses, those languages which generally require an overt subject in finite clauses, like English and French, are more restrictive, banning in particular the sequence complementizer+subject gap.

In connecting restrictions on subject extraction in English with the special status of the finite clause subject in the language, it is worth noting that there is nevertheless a genuine difference between the two types of structures compared. The point is that clauses exhibiting subject extraction, like (1) and (4b) above, are of course not strictly speaking subjectless, because the fronted element even in its unorthodox position still functions as the subject of the lower clause, as shown for instance by the subject-verb agreement and the active/passive relationship illustrated respectively in (8) and (9):

- (8) a. **Which boys did you think Δ was guilty?*
 b. *Which boys did you think Δ were guilty?*
- (9) a. *Which of them did you say that Peter had invited Δ?*
 b. *Which of them did you say Δ had been invited by Peter?*

Nevertheless, with the subject moved into the higher clause, the *that*-clause contains a subject gap much like a truly subjectless clause, and therefore the ban on the *that*+gap sequence may well be a reflection of the general similarity of the subject extraction pattern to truly subjectless clauses.

An interesting piece of evidence of the connection between the two cases comes from a diachronic study of the corresponding phenomena in Swedish (Platzack 1985). Studying the disappearance of subjectless sentences and of the extraction pattern *att*-complementizer+subject gap, Platzack found that their development was largely the same, with both becoming gradually less common and disappearing more or less completely during the 17th century. Platzack assumed then that the Null Subject Parameter in Swedish changed its value from plus to minus during that century, and further strengthened this finding by the observation that some related phenomena, assumed to be linked to the minus value of the parameter, gradually came into use while the parameter was changing its value.

3. Null subjects in the history of English

From the facts quoted, it seems reasonable to hypothesize that in English too, the gradual loss of the *that*+gap pattern summarized above is derived from the general status of truly subjectless sentences. There are two kinds of data which can be adduced straightaway as arguments for the correctness of this hypothesis. First, we have the general development of genuinely subjectless sentences from Old English to Early Modern English. As commonly observed, Old English, though generally requiring a subject in finite clauses, admitted a number of exceptions to this main rule (cf. Closs Traugott 1992: 202, 213), but the general trend was for these exceptions to disappear gradually from the language. A few illustrations of this will be in order:

- (10) a. *Wære ðu to-dæg on huntnode? Ic wæs, for-ðæm sunnadæg*
 were you today in hunting? I was, because Sunday
is. (Ælfric's Colloquy, c. 980)
 is.
 'Were you hunting today? I was, because it is Sunday'
 b. *Stodum ... esnæs ... æt gloedum forðon cold wæs.*
 stood ... retainers ... at the fire because cold was
 (Lindisfarne Gospels, c. 950)
 'There stood some retainers around the fire because it was cold'
- (11) a. *ða cwæp he me þyrst. (West Saxon Gospels, c. 1000)*
 then said he I am-thirsty
 'Then he said: I am thirsty'
 b. *Eet this when þe hrengreþ. (Langland, c. 1377)*
 eat this when you are-hungry
 'Eat this when you are hungry'
 c. *Me thynke noghte so as vn-to þe. (Hampole, c. 1340)*
 me seems not so as unto you
 'It does not seem to me the same as to you'

- (12)a. *And as to the tytell that I have to the Lordship of Grettam, schal with in short tyme be knoweyn. (Paston Letters, 1422-1509)*
 b. *As for these townes, wyll neuer tourne frenche. (Berners, 1525)*

In the impersonal constructions of (10), the zero form was “extremely rare” even in Old English (Visser 1963: 4, Ogura 1986: 42-50), and normally replaced by a structure with *hit* as the subject. With the impersonal verbs of (11), the zero form similarly gave way to a construction with a clear NP subject during the Middle English period (cf. Fischer – van der Leek 1987, Denison 1990), and with the special “marked topic” constructions of (12), occurring commonly in late Middle English (cf. Johannesson 1985), the subjectless form survived into Early Modern English and disappeared in the 16th century (cf. Visser 1963: 10).

The picture that emerges from these typical examples is thus one of a situation where the originally subjectless constructions were gradually replaced by constructions with an overt subject, this tendency being clear even in earliest Old English and gradually growing during the following 700-800 years until subjectless finite clauses were generally banned in late Middle English and/or Early Modern English.² This development, it will be noted, is exactly parallel to the history of the *that*+gap structure in extractions, from the growing avoidance of it during the Old English and Middle English periods until its total disappearance, at least in the standard language, in Early Modern English. The parallelism between the two cases, tallying very neatly with the findings of Platzack from Swedish, strongly supports the hypothesis that the history of the *that*-trace pattern can be seen as a reflection of the more general development of subjectless sentences.

The second point to note in this connection is the replacement of the *that*+gap sequence, in Old English and especially Middle English, by the *that*+resumptive structure, as given earlier in (3). To the extent that the introduction of the resumptive pronoun sufficed to remedy the *that*+gap structure, it directly shows that even though the *that*-clause had a displaced subject, the existence of the empty subject position was felt to be disturbing in much the same way as in genuinely subjectless sentences.

Having so far considered subject extractions and their historical development by viewing them as parallel to genuinely subjectless sentences, we must next note that the interpretation we have assumed is confronted with a major problem which, at least at first blush, makes the whole view look questionable. The point is that the form which ultimately replaced the *that*+gap sequence was the structure with a deleted complementizer, the zero+gap sequence, which in an obvious sense seems to be identical with the forbidden structure as far as the status of the subject is concerned. If it was the lack of an overt subject which led to the disappearance of the *that*+gap structure, how was it then possible for the zero+gap construction to survive? With the deletion of the complementizer, the complement clause loses its structural marker, and it might conceivably be assumed that when the clausal

² Many of the instances of subjectless sentences recorded after 1500 seem to reflect a conservative usage that is idiosyncratic with some individual writers (cf. Visser 1963: 35).

status of the complement is not explicitly marked, it can more easily be accepted even in spite of its lack of an overt subject. But if the status of that complement as a clause is not affected by the omission of the complementizer, it is difficult to see how the radical difference in grammaticality between **Who do you think that has done it?* and *Who do you think has done it?* can be reconciled with the view that the former is ungrammatical because it has no overt subject.

The problem confronting us here has for a long time been one of the questions debated within the generative framework, and it will therefore be necessary for us to see how far the analyses offered there can contribute to the solution of our problem.

4. The *that*+gap/zero+gap structures in generative grammar

The discussion of the *that*+gap/zero+gap variation in generative syntax has often included related problems not directly connected with the task facing us. Attempting to disentangle from that discussion those aspects of the complex problem which are most directly relevant to our work, we can see two general approaches emerging, distinguished from each other in their view of the clausal status of the zero+gap pattern. The approaches are associated with different schools of generative grammar and are best discussed separately.

4.1. The GB analysis

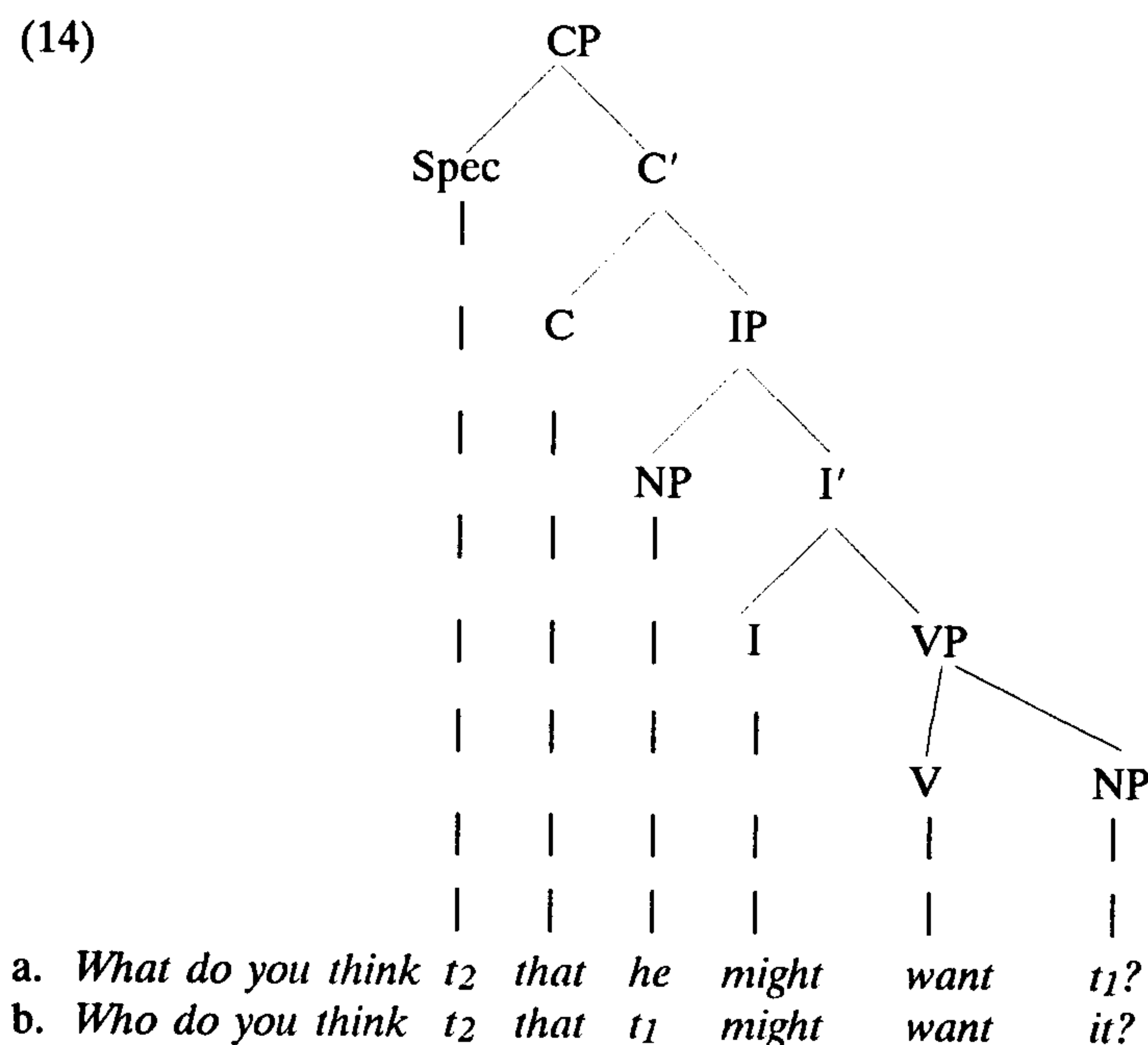
Within the GB tradition, the approach to subjectless sentences proposed by Perlmutter has generally been replaced in later work by a treatment which distinguishes subject extraction from genuinely subjectless sentences. For the latter case, the analysis postulates the empty subject *pro* (“the small *pro*”), which may be an argument (as in Italian), or a mere dummy element (as generally in Old English), or may not at all be allowed (as in Modern English). Attempts to derive the use or non-use of *pro* in individual languages from some universal principle of licensing factors have not been altogether successful³, but the explanation most commonly advanced, the licensing of *pro* by subject/verb agreement in richly inflected languages (cf. Platzack 1985, Rizzi 1986, Platzack 1987), might apply to English and explain the change from Old English to Modern English, as it has been claimed to explain the same change in mainland Scandinavian languages (cf. Platzack 1985, 1987).

Distinct from *pro*, the empty position resulting from subject extraction is occupied by an NP trace exactly as in cases of non-subject positions, and the general trend of GB work has been to show that the occurrence of subject gaps is governed by the same general principles that constrain the occurrence of non-subject gaps: all traces (empty nodes marked as extraction sites of moved elements) must, in the terse formulation of the Empty Category Principle (ECP), be “properly governed”, and once this general requirement is met, there is no basic distinction in the extraction of subjects and non-subjects.

³ For a summary of the attempts to derive the use of *pro* from such a universal principle, see Haegeman (1991: 418-419).

How the crucial term "proper government" is to be defined is a matter of dispute, and in current work different versions of the same basic view are found. We will begin our examination by considering the version proposed by Chomsky (1986: 47-48; cf. Haegeman 1991: 404-407), examining it in the light of the examples in (13) and their partial tree structure representation in (14):

- (13)a. *What do you think (that) he might want Δ?*
 b. *Who do you think (that) Δ might want it?*



In this version of the theory, proper government is either by theta government or by antecedent government. In (14a), the object trace t_1 is theta-governed by the verb which governs it and theta-marks it (assigns a theta-role to it). In the case of the subject trace t_1 in (14b), theta government is excluded, but there are three potential governors of the trace. First, the head of the IP, the I-constituent *might*, m-commands the trace (and assigns the nominative case to it, or to the chain *who* - t_1), but not being coindexed with the trace, *might* is not a proper governor. As the trace is included within the IP-constituent and maximal projections are normally barriers to outside government, the trace can be properly governed only if it is stipulated that IP, though a maximal projection, is in fact not a barrier. With this stipulation, two potential governors are available from outside, the complementizer *that* and the intermediate trace t_2 . By the principle of strict minimality, the trace must then be governed (antecedent-governed) by the nearest governor, i.e. by the complementizer. But not being coreferential with t_1 , the complementizer is not a proper governor; t_1 therefore is not properly governed and for this very reason the sentence is ungrammatical. With the com-

plementizer deleted, the only governor of t_1 that is left is t_2 , which is coreferential with t_1 , thus guaranteeing that t_1 is properly governed and hence that the sentence is grammatical.

The machinery needed for the account is obviously on the heavy side, and contains at least two features which are in principle unattractive. First, as has been pointed out by Horrocks (1987: 141), it is by no means obvious that the two very different processes called theta-government and antecedent government should be thus combined, and secondly, the exceptional treatment of the IP as a non-barrier is evidently an ad hoc stipulation needed to make the system work. Even then, the result is clearly unsatisfactory for our purposes because the whole description is tailored to deal with the facts of present-day English. If the grammar treats *Who do you think that Δ might want it?* as ill-formed, in terms of principles which are meant to be part of the innate universal grammar and hence applicable to all languages, that grammar is simply incapable of dealing with the facts of Old English and Middle English, and uninteresting for our purposes in the form in which it is standardly presented.

In this respect, the reformulation of the analysis proposed by Rizzi (1990) is in principle more promising, although it changes the theory only in a minor way. The crux of the new version of the ECP is that while traces must still be properly governed, this relationship now includes both the antecedent-government assumed on the standard view and a new type of government called head-government. According to this view, (13b) would be treated as follows (cf. (14)):

The intermediate trace t_2 can antecedent-govern the trace t_1 regardless of whether the complementizer position is empty or occupied by *that* (the principle of relativized minimality) and the difference between the grammatical and ungrammatical strings hinges on whether the C, as head of the CP constituent, can head-govern the trace t_1 . Normally not a governor, the complementizer is made one if it agrees with the specifier, and the relationship between the specifier and the head can vary. In present-day English, the agreeing complementizer in complement sentences is zero, which by virtue of this agreement becomes a governor and then head-governs t_1 .

Rizzi does not explicitly deal with Old English and Middle English, but his reference to modern dialects which preserve the old *that*+gap structure shows how the description can be applied to historical material. As the possibility of a subject gap after *that* is dependent on the complementizer being made governor by its specifier-head agreement, the theory requires us to assume that the complementizer *that* did agree with its specifier (the intermediate trace t_2) in Old and Middle English. In standard English, it was then gradually lost by about 1500, but was preserved in some dialects which therefore allow the *that*+gap pattern even in modern usage.

From this account, it is obvious that in one sense the version of ECP presented by Rizzi can deal with the historical data by the introduction of the \pm agreement features for the complementizer. But taking a closer look at the suggested analysis, what does it mean to speak of agreement in this connection? In the case of the structure *Who do you think t₂ - t₁ might want it?* and its Old English and Middle

English predecessors, the description makes the claim that a specifier-head agreement in terms of some unspecified categories (number, gender, etc.) holds between the specifier t_2 and the zero complementizer, i.e. between two empty positions in the assumed tree representation of the sentence, and in the case of *Who do you think t_2 that t_1 might want it?* and its predecessors, such agreement similarly obtained between the empty position t_2 and *that* in the older language but does not today. But is it meaningful to speak of agreement in the case of two empty positions, and in particular with reference to the complementizer position, which is empty even in a purely semantic sense? And in the case of the overt complementizer *that*, how do we know when it agreed with the postulated intermediate trace and when it ceased to do so when there is no sign of any variation of the item itself? As the \pm Agr feature is postulated without being derived from anything else in the language, the only criterion available seems to be the possibility of subject extraction: if *that+gap* is possible, then *that* shows agreement, but not otherwise. But in such an analysis, explicitly suggested for the description of the modern dialectal use of *that+gap* structures (Rizzi 1990: 53), the postulated agreement, as far as we can understand, is nothing more than an ad hoc device for describing extractability: it has no explanatory power in the sense of deriving subject extractability from some other properties of the language. Would it therefore not be simpler, and hence preferable, to state it directly as a brute fact that the *that+gap* pattern was grammatical in Old English and Middle English, but is no longer so?

If Rizzi's account of subject extractions is criticized as relying on a purely technical device of agreement features, then the obvious way to make the account more plausible is to show that these features can be justified in terms of some linguistic patterning which affords independent evidence for the postulated features. This position is the starting-point for a study of the question by Platzack (1993) which is concerned with the equivalent of *that+gap/zero+gap* variation in modern Swedish, but which is in principle similarly applicable to English both in diachronic and synchronic terms. A brief examination of the study is in order here.

The concrete pattern dealt with by Platzack is the variation between the following two sentence types, of which the first is in common use in Swedish and the second is ungrammatical in Sweden but is a natural alternative to (a) in the variety of Swedish spoken in Finland:⁴

- (15)a. *Vem tror du Δ kommer?*
 who think you Δ comes
 'Who do you think is coming?'
 b. *Vem tror du att kommer?*
 who think you that comes
 'Who do you think is coming?'

⁴ In his discussion of the contrast, Platzack describes the zero form as though it was not used in Fenno-Swedish, but in actual fact, the (a)-variant is as natural in Finland as it is in other varieties of Swedish.

Assuming the general approach to subject extraction proposed by Rizzi, the complementizer *att* and zero must be able to identify the extracted subject *vem* by being coreferential with *vem* as indicated by the usual device of coindexing:⁵

- (16)a. *Vem tror du $\Delta_1 t_1$ kommer?*
 b. *Vem tror du att₁ t₁ kommer?*

Being coreferential with *vem*, the complementizer must obviously be nominal, and the acceptance of the (b)-example in Fenno-Swedish must thus imply that in this form of Swedish the complementizer *att* has nominal features which are not found in other varieties of the language. But nominality of the complementizer must in turn mean that the clause headed by it must in Fenno-Swedish have a distribution which is similar to the distribution of other nominal elements. Comparing the syntax of the two varieties of Swedish, Platzack next notes a difference between them in the use of prepositions, as in the following pair:

- (17)a. *Han underrättade henne om beslutet.*
 a'. *He informed her of the decision.*
 b. *Han underrättade henne (om) att han skulle resa.*
 b'. *He informed her (*of) that he would leave.*

Before the NP *beslutet*, *om* is obligatory in both varieties of Swedish, but in front of the clausal complement (where it is obligatorily deleted in English), it can be optionally dropped in Sweden but not in Finland. By means of a test carried out among informants from both countries, Platzack shows that the difference is more generally found in the greater reluctance of Fenno-Swedes to drop a preposition in front of an *att*-clause in contexts where an NP obligatorily requires a preposition.

All in all, Platzack's basic idea of looking for concrete evidence for the assumed nominality of the complementizer *att* is laudable, and his discovery of a Fenno-Swedish reluctance to preposition deletion is of great interest. Nevertheless, before accepting these data as evidence for the subject extraction analysis assumed by Rizzi and Platzack, it seems necessary to view these data from a more general perspective. Above all, it is to be remembered that the most important distinction from the point of view of subject extraction is the contrast between the complementizer *att* and zero in that the zero+gap structure is perfectly acceptable to all speakers of Swedish whereas the *att+gap* structure represents a minority usage which is totally unacceptable to most speakers. If subject extraction is dependent on the nominality of the complementizer, then obviously the zero complementizer must be nominal in all varieties of Swedish, and the distribution of zero+clause must therefore be identical with the distribution of NPs, or at any rate closer to it than is the distribution of the *that+clause* pattern. What the facts of the case are we can examine in the light of the following sentences, illustrating the use of

⁵ Following Platzack's non-technical account, we leave out here the intermediate trace before the complementizer, and replace Platzack's dialectal example by a simpler sentence (cf. Platzack 1993: 162-163).

(*att*)-clauses in the function of subjects, objects and prepositional complements, and in apposition to an NP. The English sentences added are translations of the Swedish, but are at the same time treated as illustrations of the situation in English and are marked for grammaticality according to their status in this language:

- (18)a. *(*Att*) Johan var närvarande på mötet var en överraskning.
 a'. *(*That*) Johan was present at the meeting was a surprise
- (19)a. Hon hänvisade till *(*att*) tåget hade varit försenat.
 a'. She referred to *(*that*) the train had been late
 b. Hon var glad över *(*att*) pojkarna hade kommit tillbaka.
 b'. She was happy about *(*that*) the boys had returned
- (20)a. Hans första förslag, *(*att*) vi skulle stanna en vecka till
 a'. His first suggestion, *(*that*) we should stay one week more,
 är lättare att acceptera
 is easier to accept
- (21)a. Han trodde (*att*) han hade rätt.
 a'. He thought (*that*) he was right
 b. Han trodde tydligen *(*att*) han hade rätt.
 b'. He thought apparently *(*that*) he was right
 c. De kommer att föresl *(*att*) ärendet skall bordläggas.
 c'. They are going to propose *(*that*) the question should be tabled

As shown, the zero+S structure is totally impossible as subject, as prepositional complement and in apposition to an NP, and is acceptable in object function only with certain verbs (cf. (21a) vs. (21c)), and even there only under certain conditions (cf. (21a) vs. (21b)). Compared with the *att*+S pattern, the zero form is thus clearly less nominal in its distribution than the *att*+S structure, which is generally acceptable in all of the functions illustrated. If nominality of the complementizer is a precondition for subject extraction, we must then expect the zero+gap pattern to be generally unacceptable. Furthermore, since the distribution of the zero+S structure, as indicated in our translations, is as limited in English as it is in Swedish, the same conclusion must similarly apply to English. But since the zero+gap pattern is in fact the normal subject extraction structure in both languages, the whole idea of the nominal nature of the complementizer, crucial for the analysis assumed by Rizzi and Platzack, must be rejected at least as far as the zero complementizer is concerned.

Addressing ourselves next to the *att*-complementizer, let us consider Platzack's assumption that the non-deletability of the preposition in front of an *att*-clause is a sign of the nominal nature of the complementizer which then makes subject extraction possible. Since prepositions are (under certain conditions) deletable in front of *att*-clauses in all varieties of Swedish, and often also in English (thus *about* in (19b) above), it would seem that the complementizers *att* and *that* lack nominal features, in spite of the NP-like overall distribution of the clauses introduced by them, and that this lack provides the explanation for the ban on subject extraction. To test the plausibility of this hypothesis, let us for the moment leave

Swedish and English, and consider the situation in closely related languages. One interesting point of comparison is offered by Danish, which on this point differs from Swedish in its general ban on the kind of preposition deletion found in Swedish (and English). The following examples, quoted from a recent discussion of the question (Götzsche 1994: 123-124), illustrate this contrast:

- (22)a. De är villige til at forhandle.
 b. De är villiga att förhandla.
 they are willing to negotiate.
 'They are willing to negotiate'
- (23)a. Lysten til at rejse var stor.
 b. Lusten att resa var stor.
 the desire to travel was great
 'The desire to travel was great'
- (24)a. Lysten til øl er stor.
 b. Lusten efter øl är stor.
 the desire for beer is great
 'The desire for beer is great'

While Swedish and English thus have a preposition before an NP, as in (24), but drop it in front of an *att*-clause, the Danish equivalents retain the preposition in both cases. The examples are from a Danish 19th century text (by H.C. Andersen) and its Swedish translation (from 1875), but the usage described is the same today, as shown in the following examples from Jacobsen (1986: 41-42), who uses them to illustrate the higher degree of nominality of Danish *at*-clauses compared with English *that*-clauses:

- (25)a. De var enige om at det var en trist situation.
 a'. *They agreed on that it was a sad situation
 b. John er bange for at køre bilen.
 b'. *John is afraid of to drive the car
 c. Patienten blev sævkket af at lægerne gav ham
 c'. *The patient was weakened by that the doctors gave him
 den forkerte medicin.
 the wrong medicine.

All in all, it is thus clear that Danish *at*-clauses have a degree of nominality which goes far beyond the nominality of corresponding clauses in either Swedish or English. This leads us therefore to expect that in Danish, subject extraction from such clauses is fully natural in contrast to Swedish and English. However, the expectation turns out to be totally erroneous, as can be seen from the ungrammaticality of the following kind of structures presented by Hansen (1967: 105; cf. Platzack 1993: 154-155):⁶

⁶ The examples from Hansen are here given in a slightly simplified and modernized form.

- (26)a. *Det syntes han (*at)var en drlig present.*
 a'. *That he thought (*that) was a foolish present*
 b. *Hvormange af dem sagde du (*at)havde set udstillingen?*
 b'. *How many of them did you say (*that) had seen the exhibition?*
 c. *Det var Redsted som Jonas vidste (*at)havde vret rejsende*
 c'. *It was Redsted who Jonas knew (*that) had been travelling*
for P.K. Hansen.
for P.K. Hansen.

Since the nominality of Danish *at* is generally not enough to license subject extraction, it is difficult to accept the much weaker nominality of *att* in Fennoswedish as a valid explanation of the extraction of subjects in that language.

Another point of comparison is provided by German, where *dass*-clauses do not normally function as prepositional complements and where non-expression of a preposition in front of a *dass*-clause is common in contexts where a preposition would be obligatory in front of an NP complement. The situation is closely parallel to English, as is clear from the following examples (cf. Seppänen 1989):

- (27)a. *Der Lehrer freut sich (darüber), dass er ihnen die Grammatik*
 the teacher make-happy himself (about) that he themthe grammar
beigebracht hat.
 taught has
 'The teacher is happy that he has taught them the grammar'
 b. *Er ist schuld (daran), dass alles missglückt ist.*
 he is to-blame (for-it) that everything failed has
 'He is to blame for the fact that everything has failed'
 c. *Ich bin mir (dessen) bewusst, dass es sonichtweitergehen kann.*
 I am myself (of-it) aware that it so not continue can
 'I am aware of the fact that it cannot continue like that'

As the degree of nominality of *dass* and *that* appears to be about the same, we might expect them to have the same kind of effect on subject extraction. In reality, however, subject extraction is not always excluded in those varieties of German in which extraction in general is allowed, as in the following examples quoted from Andersson – Kvam (1984: 58-61):

- (28)a. *Welcher Zeitraum würden Sie meinen, dass Δ dafür nötig wäre?*
 whattime would you think that Δ for-it necessary were
 'What time would you think is necessary for it?'
 b. *Was glaubst du, dass Δ hier betrieben wird?*
 whatthink you that Δ here doneis
 'What do you think is being done here?'

Since the lack of nominal features of *dass* does not preclude subject extraction in German, there is again no reason to believe that the nature of *that* can explain the lack of *that*+subject gap structures in English.

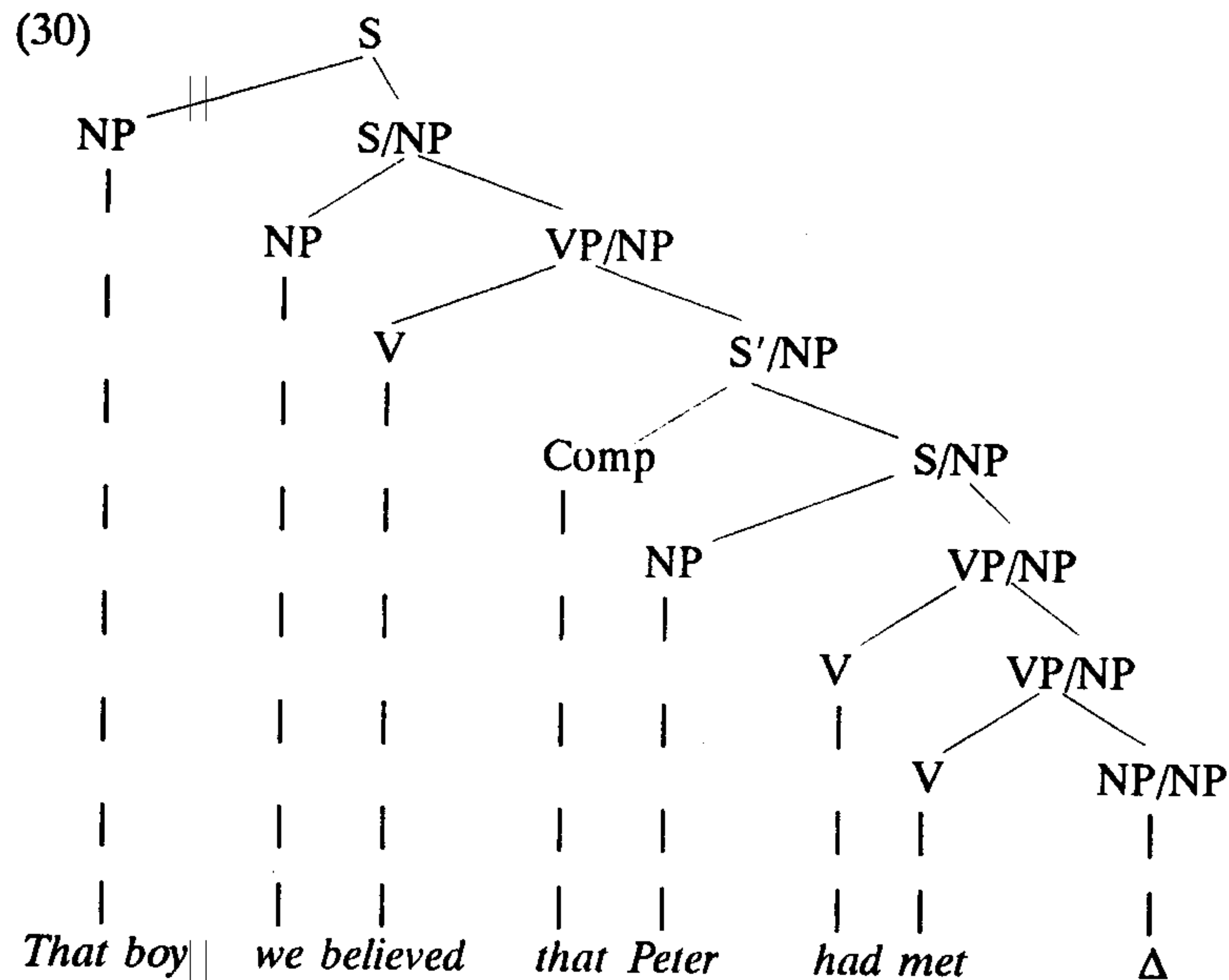
To sum up our discussion, Platzack's attempt to seek independent evidence for the Rizian analysis of subject extraction is in principle much to be recommended. However, when tested in the light of further data not considered by Platzack, the result is very different from that assumed by Platzack himself: whether it is the overt complementizer *that* or the empty complementizer position which is supposed to have nominal features which can identify the extracted subject, the predictions that follow from this assumed nominality are entirely false, and therefore force one to reject this version of the GB description of subject gaps as much as the two earlier ones. For our purposes, the most important aspect of the question is the status of the structure with a zero complementizer plus a subject gap, and its acceptability at a time when zero subjects in finite clauses were no longer allowed.

In the different formulations of the GB theory, it has always been taken for granted that the zero+gap structure has a clausal status with a subject gap, a zero complementizer before it, and an intermediate trace of the extracted subject in the Spec-C position. In view of the unsatisfactory character of the proposed analyses, it is natural to turn to a description where these basic assumptions are not taken as self-evident truths.

4.2. The GPSG account

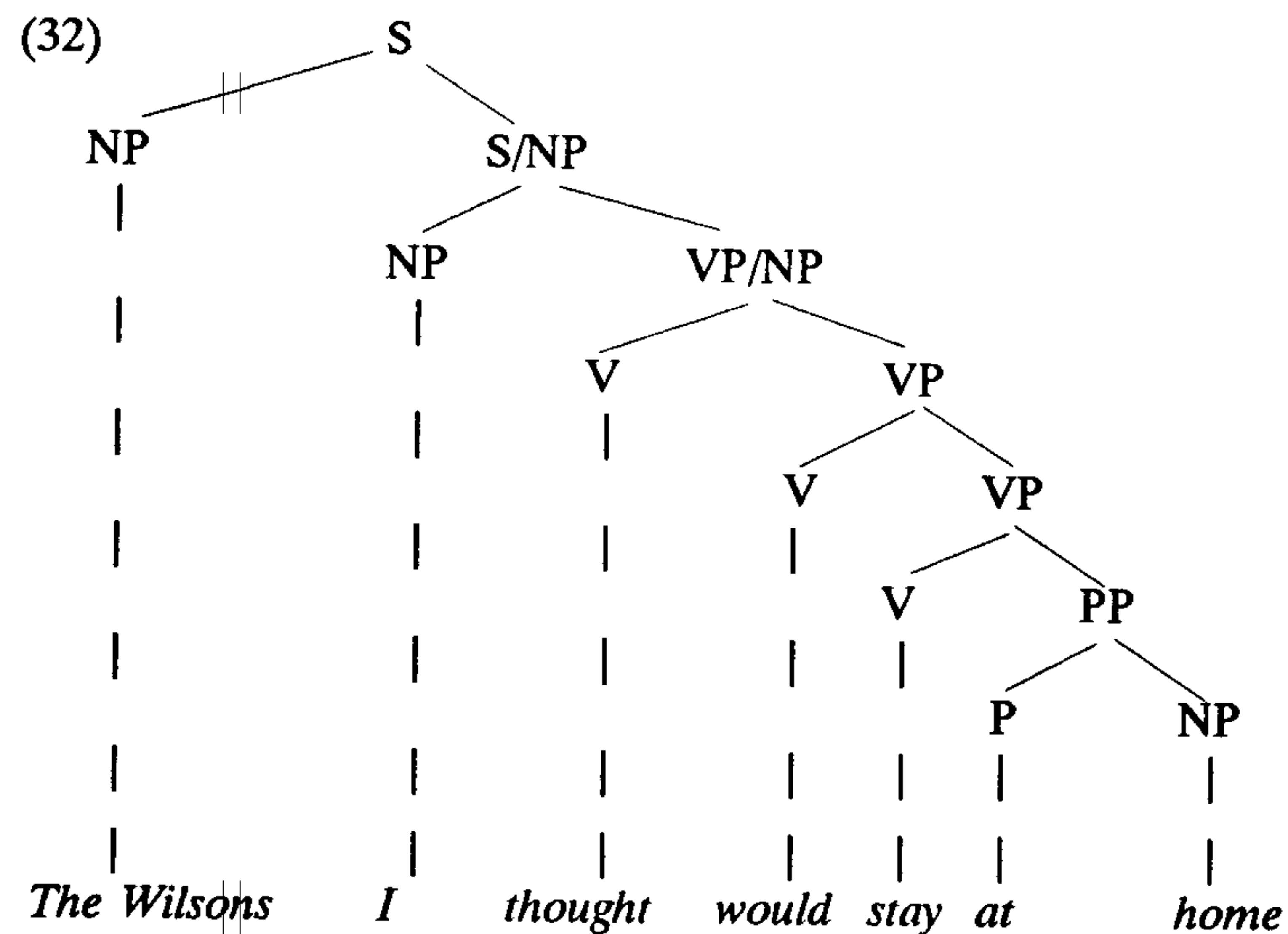
Looking next at the description of null subjects within Generalized Phrase Structure Grammar (GPSG), we may first note that with its total rejection of all transformations, the theory deals with the general ban on null subjects simply by not admitting any PS rule which would generate a finite clause without an overt subject in a language like English. Extraction structures are generated by means of the rule $XP+S/XP$, where XP is equivalent to the fronted constituent and S/XP to the sentence with a missing XP constituent. The sentences in (29) and the tree structure in (30) will illustrate the normal situation where a gap appears in the expected position of the fronted constituent:

- (29)a. *That boy we believed that Peter had met Δ*
 b. *I don't know what they thought that Alex should buy Δ*
 c. *That is not the case that we hoped that you would handle Δ*

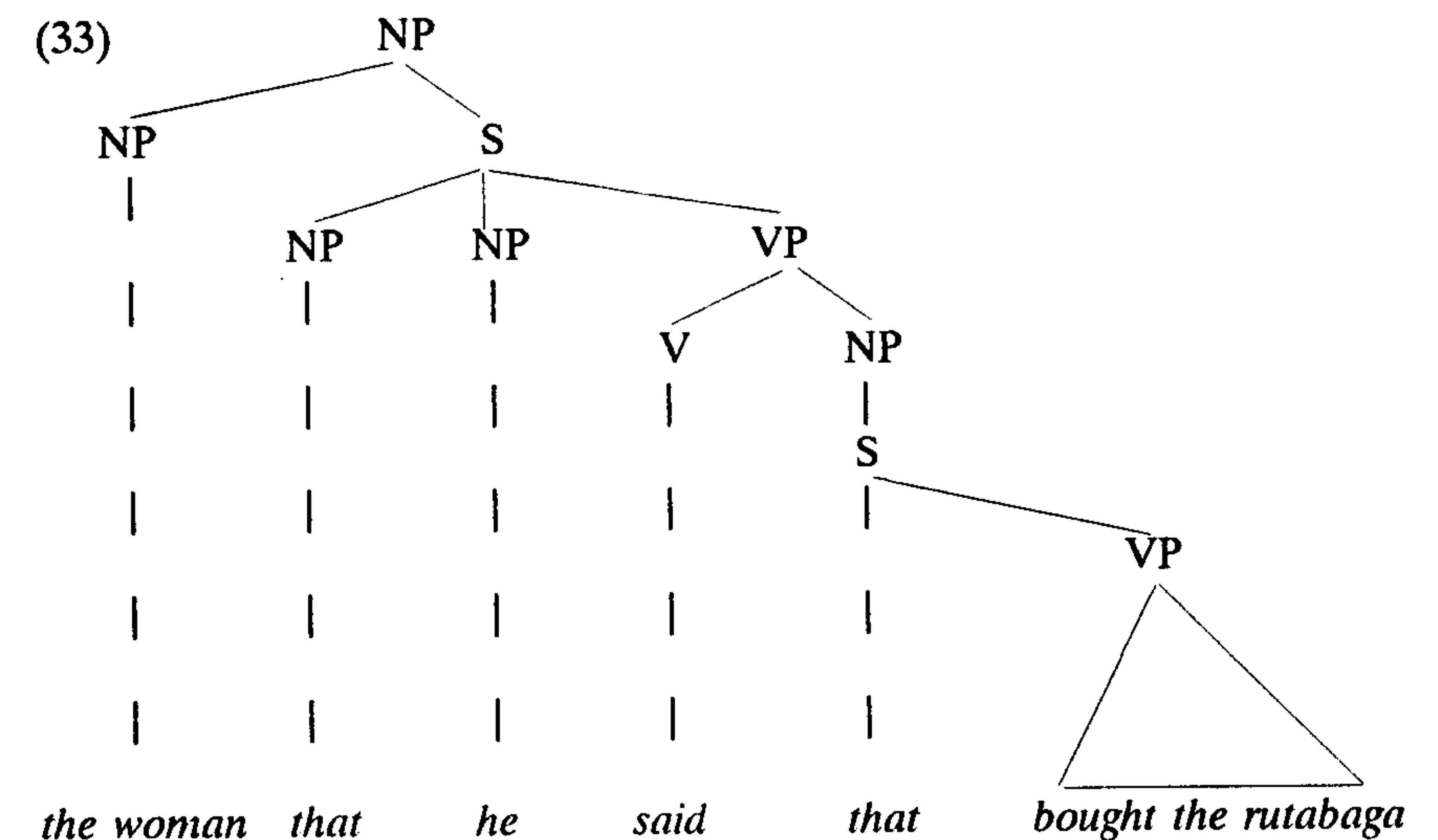


As regards subject extraction, the theory treats it as in one sense radically different from non-subject extraction. When a non-subject XP is extracted, the VP constituent is left but contains a gap (VP/XP, where the XP can be an object NP, predicative NP or AdjP, etc.); but when the subject XP is extracted, the subject constituent as such disappears and the S/XP is rewritten as a mere VP. The examples in (31) and the tree structure in (32) will illustrate the situation:

- (31)a. *The Wilsons I thought [VP would stay at home.]*
 b. *Which of them did you think [VP was the best of the lot?]*
 c. *I didn't much care for those books which Martha said [VP were so good.]*



Compared with the GB analysis, this description obviously offers a very different structural interpretation of the data. On the other hand, the structure assumed is not altogether novel, being identical with the surface structure postulated by Perlmutter (1971) on the basis of a tree-pruning rule originally proposed by Ross (1967: 24-64). The essence of the rule is the pruning of any non-branching S-node. Thus, in the structure *the woman that he said that bought the rutabaga*, the tree after the relativization is as follows (cf. Perlmutter 1971: 112):



Here the lower S is a branching node by virtue of the retained complementizer, and thus ungrammatical as a finite clause lacking a subject. If the complementizer itself is zero, then S does not branch, and must be pruned so that *bought the rutabaga* comes to be a mere VP rather than a clause, and thus does not violate the subject requirement.

This earlier description of the mere-VP analysis of the subject extraction pattern is of interest, but the crucial question to ask is of course whether this analysis can be justified in terms of concrete data which are not dependent on any particular view of transformational operations. One argument for its correctness, derived from coordination, is illustrated by the following examples (Gazdar et al. 1985: 178):

- (34)a. *I know a man who Bill saw and Mary liked.*
 b. *I know a man who saw Bill and liked Mary.*
 c. **I know a man who Bill saw and liked Mary.*
 d. *I know a man who Mary likes and hopes will win.*

In (34a-b), the sentences contain a straightforward coordination of like categories, two S/NPs and two VPs, respectively. Different from this, (34c) fails because it attempts to coordinate two different categories, S/NP and VP, whereas (34d)

again is grammatical because the coordinated elements are both VP/NPs. The interesting case here is (34c), showing the critical role of the distinction between S/NP and VP in making correct predictions about grammatical and ungrammatical sentences.

Another argument to the same effect comes from the facts of auxiliary reduction, discussed by Schachter (1984). Consider the following set of examples, with the structures analyzed in terms of GPSG:

- (35)a. *How much* $\left\{ \begin{array}{l} \text{milk has} \\ \text{milk's} \end{array} \right\}$ *been spilt?*
 b. *How much* $\left\{ \begin{array}{l} \text{milk has} \\ \text{*milk's} \end{array} \right\}$ *Ben spilt?*
 c. *How much* $\left\{ \begin{array}{l} \text{milk has} \\ \text{*milk's} \end{array} \right\}$ *John said's been spilt?*

- (36)a. [*S* [*NP* *How much milk*] [*VP* *has been spilt?*]]
 b. [*S* [*NP* *How much milk*] [*S/NP* *has Ben spilt* Δ?]]
 c. [*S* [*NP* *How much milk*] [*S/NP* *has John said* [*has been spilt?*]]]

It seems thus that auxiliary contraction is blocked by a sentence boundary, as in (35b-c), but is possible when no sentence boundary intervenes, as in (35a).

Next, consider the contractions in the following sentences:

- (37)a. *What do you think's happening?*
 b. *What do you think's been happening?*
 c. *How much milk has John said's been spilt?*

In each case, the subject of the lower clause is extracted, and the possibility of contraction shows that the residue is treated as a mere VP rather than an S/NP.⁷

One further indication of the mere VP status of the residue is the following contrast obtaining in some varieties of modern colloquial English, as noted by Kroch (1981: 127). The examples are quoted from his account, with his judgements of acceptability:

- (38)a. *There was one guy who I didn't even think that he would come.*
 b. **There was one guy who I didn't even think he would come.*

If, after the extraction of the subject, the *that*-clause has the full clausal structure (*that*) - Δ - VP, then there is no reason to expect the resumptive pronoun to be any better in the *that*+gap structure than in the zero+gap structure, whereas the bare VP, not containing any empty subject position, will automatically exclude the resumptive pronoun. The contrast is not all that clear in many, perhaps most, varieties of English, but in Swedish, where the same VP analysis seems to apply, the contrast is very clear because the complementizer+resumptive pattern is fully grammatical, as shown by Engdahl (1985: 40):

⁷ It should be added that the force of this argument is weakened by the fact that individual judgements are not altogether uniform on these and equivalent structures.

- (39)a. *Kalle kan jag sl vad om att han kommer att klara sig.*
 a'. **Kalle can I bet about that he is going to succeed*
 b. **Kalle kan jag sl vad om han kommer att klara sig.*
 b'. *Kalle can I bet about he is going to succeed.*

Thus, there is some concrete evidence for the view that in cases of subject extraction, the defective structure that remains is in fact not a clause in present-day English but a mere VP. Accepting that analysis, we must note, however, that the GPSG description of the subject extraction pattern is not generally adequate for our historical study. The crux of the matter is that the "missing subject" construction of (32) is generated by means of a metarule, which as such can apply only to lexical rules, and hence not to the rules which expand S[COMP *that*], which are not lexical rules. The effect of all this is that the grammar can generate a sentence like *The man who I think chased Fido returned*, but not **The man who I think that chased Fido returned* (Gazdar et al. 1985: 161-162). But while this works for present-day English, it is obviously inadequate as a general approach to subject extraction, and even as a framework for describing the history of that structure. To remedy this shortcoming, the system would obviously have to be modified, but although different alternatives have been considered as possible avenues of approach (Gazdar, personal communication), so far at any rate none of them has been worked out.

To sum up, it seems then that the current version of GPSG is inadequate as an overall framework for a description of the history of extractions in English. In our discussion of this history (in the next section), we shall assume that the GPSG view of subject extraction structures as bare VP complements is correct, but will not make any specific assumptions about the way this structure is generated: in terms of the GB theory, by developing a way of generating the bare VP structure by means of tree-pruning or trace-deletion⁸, or in terms of GPSG, by modifying the description so as to enable it to handle the *that*+gap sequence.

5. Syntactic reanalysis in subject extraction structures

After our long discussion of subject extraction patterns in present-day English, it is now time to return to the history of these structures. Sketching the development of this area of English syntax, we saw how the admission of null subjects was severely limited even in earliest Old English, and came to be more and more restricted during the Old English and Middle English periods till their total demise in standard usage in Early Modern English. This development may be described in terms of a general rule of obligatory subjects and of a gradual elimination of exceptions to the rule. In his discussion of null subjects, Perlmutter (1971: 100) formulated the rule as a surface constraint, but we can reformulate it without his

⁸ Platzack (personal communication), accepting the bare VP analysis as a possible interpretation of the data, applicable even to the description of Modern Swedish, offers some remarks on how the resulting structure could be dealt with within the GB theory.

transformational implications and ignoring details (imperatives, etc.) not directly relevant here:

(40) *A finite clause must have an overt subject.*

The rule is clearly similar to the Extended Projection Principle (EPP) of more recent GB syntax which requires every sentence to have a subject but allows for the possibility of non-overt subjects (cf. Haegeman 1991: 54, 240-243), and may indeed be taken to be a special case of the EPP. The imposition of the subject rule was bound to have an effect on the subject extraction patterns, and was directly reflected in two developments which we have noted above: the decline of the *that*+gap pattern and its total abandonment around 1500, and the use of the *that*+resumptive structure as one way of replacing the disfavoured sequence. The question that this situation raises is the status of what we have been calling the zero+gap pattern: if the development on this point was generally a consequence of the overt subject rule, how was it possible for this pattern to survive and ultimately become the only acceptable form in standard English? After our discussion of the structure of this sentence pattern in present-day English, we can now suggest that the pattern survived because it was reanalyzed so that it was no longer an exception to the usual subject requirement. In its original form, the subject extraction structure was a sentence with a zero complementizer and a subject gap, to be represented in the GB and GPSG systems as follows:

(41)a. *John_i we thought [CP t_i -[IP t_i [VP was the best candidate]]]*
 b. *John_i we thought [S' [S Δ_i [VP was the best candidate]]]*

After the reanalysis, the structure is to be represented as a mere VP, as in (42):

(42) *John we thought [VP was the best candidate]*

The reanalysis involves changes which are recognized from other syntactic developments in English: the loss of clause boundary (cf. the rise of *I believe John to be guilty*, Lightfoot 1981, van Kemenade 1987: 227, Fischer 1988: 69-72, 1989: 161) and the loss of an empty category (for several examples and a general discussion, cf. van der Wurff 1990: 35-58). The special combination of these component changes, we believe, is due to the similarity between the subject extraction structure and a structure involving a parenthetical comment clause, as illustrated in (43):

(43)a. *John_i we thought Δ_i was the best candidate.*
 b. *John, we thought, was the best candidate.*

While the two structures may be clearly distinguished from each other, for example by verbal inflection (subjunctive vs. indicative), by the presence of the complementizer, by a different placement of the comment clause (cf. *John was, we thought, the best candidate* or *John was the best candidate, we thought*), or by a

difference in the word order (main clauses vs. dependent clauses), often no unambiguous indicator of this kind is present and the sentences are in fact quite commonly confused in written English (cf. Seppänen – Bergh, forthcoming). It is this influence from the parenthetical construction that we suggest lies first at the root of the early omission of the complementizer in cases of subject extraction, and then in the reanalysis of the S/NP structure with a subject gap as a mere VP. Considering the very slow replacement of *that*+gap by zero+gap, the reanalysis may also have taken a long time, but was probably completed by the beginning of the Early Modern English period, when subjectless sentences in general disappeared from the language.

One indicator of this syntactic change, we believe, is the demise of the zero+resumptive pattern noted above: as long as the structure was interpreted as involving a zero complementizer and a subject gap, the gap could also be filled by a resumptive pronoun, but in its new structural interpretation it did not have a subject node to be filled by a resumptive pronoun.

Finally, the reanalysis seems to provide an explanation for the increase in the percentage of subject extractions among all extraction patterns noted after 1500 (27 per cent before 1500, 38 per cent after this date). As is well known, extractions of elements out of a finite clause are often subject to restrictions, and are indeed in many languages only very marginal, for example in German (cf. Andersson – Kvam 1984), or virtually non-existent in the written language, for example in Russian (cf. Pesetsky 1982: 298-299, 315-321) and Finnish (cf. Seppänen et al. 1988: 96-99). Even in languages which are relatively free in this respect, like English, they seem to be subject to stylistic restrictions which often vary between different individual writers (cf. Seppänen – Bergh, forthcoming). It seems that subject extraction in its reanalyzed form, as it does not leave behind a truncated clause S/NP, is less bound by such restrictions, and being in this sense more readily available in different contexts, could increase its frequency in relation to other types of extraction.

REFERENCES

- Adamson, Sylvia – Vivien Law – Nigel Vincent – Susan Wright (eds.)
 1990 *Papers from the 5th International Conference on English Historical Linguistics*. Amsterdam: Benjamins.
- Allen, Cynthia L.
 1980 "Movement and deletion in Old English", *Linguistic Inquiry* 11: 261-324.
- Allén, Sture – Lars-Gunnar Andersson – Jonas Löfström – Kerstin Nordenstam – Bo Ralph (eds.)
 1985 *Svenskans Beskrivning 15*. Göteborg: Göteborgs Universitet
- Andersson, Sven-Gunnar – Sigmund Kvam
 1984 *Satzverschränkung im heutigen Deutsch*. Tübingen: Gunter Narr.
- Baker, Carl Lee – John McCarthy (eds.)
 1981 *The logical problem of language acquisition*. Cambridge, Mass.: MIT Press.
- Bergh, Gunnar – Aimo Seppänen
 1994 "Subject extraction in English: The use of the *that*-complementizer", in: Francisco Fernández – Miguel Fuster – Juan José Calvo (eds.), 131-143.

- Chomsky, Noam
1986 *Barriers*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam – Howard Lasnik
1977 "Filters and control", *Linguistic Inquiry* 8: 425-504.
- Closs Traugott, Elisabeth
1992 "Syntax", in: Richard M. Hogg (ed.), 168-289.
- Denison, David
1990 "The Old English impersonal reviewed", in: Sylvia Adamson – Vivien Law – Nigel Vincent – Susan Wright (eds.), 111-140.
- Engdahl, Elisabeth
1985 "Parasitic gaps, resumptive pronouns, and subject extractions", *Linguistics* 23: 3-44.
- Fernández, Francisco – Miguel Fuster – Juan José Calvo (eds.)
1994 *English historical linguistics 1992: Papers from the 7th International Conference on English Historical Linguistics*. Amsterdam: Benjamins.
- Fischer, Olga
1988 "The rise of the for NP to V construction", in: Graham Nixon – John Honey (eds.), 67-88.
1989 "The origin and spread of the accusative and infinitive construction in English", *Folia Linguistica Historica* 8: 143-217.
- Fischer, Olga – Frederike van der Leek
1987 "A 'case' for the Old English impersonal", in: Willem Koopman – Frederike van der Leek – Olga Fischer – Roger Eaton (eds.), 79-120.
- Gazdar, Gerald – Ewan Klein – Geoffrey Pullum – Ivan Sag
1985 *Generalized phrase structure grammar*. Oxford: Blackwell.
- Göttsche, Hans
1994 *Derivational syntactic structures: A contrastive linguistic study in the syntax of Danish and Swedish*. Göteborgs Universitet: Institutionen för Svenska Språket.
- Haegeman, Liliane
1991 *Introduction to government and binding theory*. Oxford: Blackwell.
- Hall, Robert A. (ed.)
1985 *The Eleventh Lacus Forum 1984*. Columbia, S.C.: Hornbeam Press.
- Hansen, Aage
1967 *Moderne Dansk. Bind I: Analyse*. København: Grafisk Forlag.
- Hogg, Richard M. (ed.)
1992 *The Cambridge history of the English language*. Vol. I. Cambridge: Cambridge University Press
- Horrocks, Geoffrey
1987 *Generative grammar*. London: Longman.
- Ivars, Ann-Marie – et al. (eds.)
1993 *Språk och Social Kontext*. Helsingfors: Helsingfors Universitet.
- Jacobsen, Bent
1986 *Modern transformational grammar*. Amsterdam: North Holland.
- Johannesson, Nils-Lennart
1985 "Oh gracious god my harte it burstes – On the historical development of unmarked and marked topics in English", in: Robert A. Hall (ed.), 432-446.
- Koopman Willem – Frederike van der Leek – Olga Fischer – Roger Eaton (eds.)
1987 *Explanation and linguistic change*. Amsterdam: Benjamins.
- Kroch, Anthony
1981 "On the role of resumptive pronouns in amnestying island constraint violations", *Chicago Linguistic Society Papers* 17: 125-135.
- Lightfoot, David
1981 "The history of NP movement", in: Carl Lee Baker – John McCarthy (eds.), 86-119.

- Nixon, Graham – John Honey (eds.)
1988 *An historic tongue: Studies in English linguistics in memory of Barbara Strang*. London: Routledge.
- Ogura, Michiko
1986 *Old English 'impersonal' verbs and expressions*. Copenhagen: Rosenkilde and Bagger.
- Perlmutter, David
1971 *Deep and surface constraints in syntax*. New York: Holt, Rinehart and Winston.
- Pesetsky, David
1982 "Complementizer-trace phenomena and the Nominative Island Condition", *The Linguistic Review* 1: 297-343.
- Platzack, Christer
1985 "Syntaktiska förändringar i svenskan under 1600-talet", in: Sture Allén – Lars-Gunnar Andersson – Jonas Löfström – Kerstin Nordenstam – Bo Ralph (eds.), 401-417.
1987 "The Scandinavian languages and the Null-Subject Parameter", *Natural Languages and Linguistic Theory* 5: 373-402.
1993 "Satsfläta med att-sats-subjekt", in: Ann-Marie Ivars – Hanna Lehti-Ekland – Pirkko Lilius – Anne-Marie Londen – Helena Solstrand-Pipping (eds.), 153-167.
- Rizzi, Luigi
1982 *Issues in Italian syntax*. Dordrecht: Foris.
1986 "Null objects in Italian and the theory of pro", *Linguistic Inquiry* 17: 501-558.
1990 *Relativized minimality*. Cambridge, Mass.: MIT Press.
- Ross, John R.
1967 *Constraints on variables in syntax*. Bloomington: Indiana University Linguistics Club.
- Schachter, Paul
1984 "Auxiliary reduction: An argument for GPSG", *Linguistic Inquiry* 15: 514-523.
- Seppänen, Aimo
1989 Inflexional and prepositional marking in complementation. [unpublished MS.]
- Seppänen, Aimo – Claes Göran Engström – Ruth Seppänen
1988 "On the Complex Noun Phrase Constraint", *Studia Linguistica* 42: 93-114.
- Seppänen, Aimo – Gunnar Bergh
forthcom. *Extractions in English syntax: A synchronic and diachronic study*.
- van der Wurff, Wim
1990 Diffusion and Reanalysis in Syntax. [Unpublished Ph.D. dissertation, University of Amsterdam.]
- van Kemenade, Ans
1987 *Syntactic case and morphological case in the history of English*. Dordrecht: Foris.
- Visser, Frans Theodor
1963 *A historical syntax of the English language Vol. I*. Leiden: Brill.