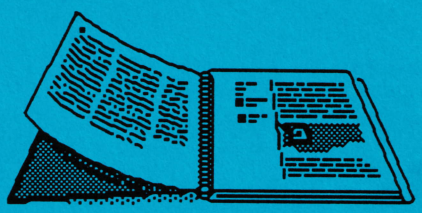
Forschungsschwerpunkt Allgemeine Sprachwissenschaft,
Sprachtypologie und Universalienforschung
der Förderungsgesellschaft Wissenschaftliche Neuvorhaben mbH

FAS Papers in Linguistics

Volume 4
December 1995



Edited by

Artemis Alexiadou
Nanna Fuhrhop
Paul Law
Sylvia Löhken

ISSN 0947-7055

Table of Contents

SVO and EPP in Null Subject Languages and Germanic*

Artemis Alexiadou & Elena Anagnostopoulou
FAS, Berlin & University of Tilburg/UCLA
artemis@fas.ag-berlin.mpg.de & e.anagnostopoulou@kub.nl

Outline

In this paper we will examine SVO/VS(O) alternations across languages (Greek/Spanish, English, Icelandic[1]) within the framework of the Minimalist Program (cf. Chomsky 1995) that restricts optional operations. We will argue that SVO in Greek/Spanish involves Left Dislocation and that inverted orders lack an expletive unlike English/Icelandic. We will propose that although the above holds, Greek/Spanish are strong EPP languages. Crucially, we will propose that there exists a parametrization of EPP checking: UG provides two options; the strong features of I° can be checked either by Moving or Merging an XP to Spec,IP or by moving the verbal head to I° provided that this has a set of specific properties. English and Icelandic opt for the first possibility whereas Null Subject Languages (NSLs) for the Move X^0 one. In our discussion we will limit ourselves to Greek and Spanish as the most representative NSLs, since other languages that are included in this group, for instance Italian, present further complications. Under this proposal crosslinguistic differences are attributed to irreducible morphological variation and to universal rules, such as Move and Merge which are also governed by the particular morphological properties of the languages in question.

1. The Problem

A. SVO/VS(O) Alternations Crosslinguistically

As is well known languages like Greek or Spanish show a certain flexibility in their word order. (1) shows that both SVO and VSO orders are acceptable: [2]

---

*An earlier version of this paper was presented at the 11th Comparative Germanic Syntax Workshop hosted at the University of Rutgers. We would like to thank Viviane Deprez, David Embick, Anastasia Giannakidou, Eric Haeberli, Riny Huybregts, Sabine Iatridou, Giusseppe Longobardi, Jaklin Kornfilt, Francisco Ordoñez, Jamal Ouhalla, Ian Roberts, Ken Safir, Ur Shlonsky, Tim Stowell, Peter Svenonius and Jan-Wouter Zwart for comments and discussion. All errors remain our own.

[1] We assume that the Icelandic facts represent the situation that holds in German and Dutch.

[2] VOS orders are also possible:

i. pandreftike tin Ilektra o Petros

ii. leyo el libro Juan

As it has been argued in Ordoñez (1994) for Spanish and Alexiadou (1994) for Greek, VOS orders involve leftward object movement over the subject. However, this type of scrambling is different from the one we find in Germanic (cf. Alexiadou 1994, Anagnostopoulou 1994): i) the object always follows the participle in periphrastic constructions, so it is moved to a relatively low position and ii) in Greek, weak NPs can occur in VOS orders unlike Germanic where only strong DPs can 'scramble':

iii. ehi agorasi vivlia o Janis
has bought books the-John-NOM
'John has bought books'

(1) a. O Petros pandreftike tin Ilektra *Greek*
the-Peter-NOM married the-Ilektra-ACC
'Peter married Ilekrta'

b. pandreftike o Petros tin Ilekrta
married the-Peter-NOM the-Ilektra-ACC
'Peter married Ilekrta'

c. Juan leyo el Libro *Spanish*
'Juan read the book'

d. leyo Juan el Libro

As known, inverted constructions occur also in the germanic languages, their properties being somehow different from the properties these orders have in NSLs. An overt expletive is present and the Definiteness Restriction Effect holds (cf. 2-3). Icelandic but not English has Transitive expletive constructions (TECs cf. 3a):

(2) a. There arrived a man *English*
b. A man arrived

(3) a. pad lasu einhverjir studentar bokina *Icelandic*
there read some students the book
'Some students read the book'

b. Einhverjir studentar lasu bokina
'Some students read the book'

B. Facts

VSO orders in Greek/Spanish have a specific set of properties absent in Icelandic/English expletive constructions. These are:

i) the SVO/VSO alternation in Greek/Spanish is not restricted to root clauses but also occurs in embedded contexts (non CP-recursion contexts, cf. Iatridou & Kroch 1992 a.o.). (4) indicates that both orders are equally possible with a complex NP and an 'if clause':

(4) a. i idisi oti (o Janis) episkeftike (o Janis) tin Ilektra
the news that the-John-NOM visited the-John-NOM the-Ilektra-ACC
'The news that John visited Ilektra' *complex NP*
b. an (o Janis) episkefti (o Janis) tin Ilektra
if the-John-NOM visits the-John-NOM the-Ilekrta-ACC
'If John visits Ilektra' *if-clause*

For reasons why this scrambling takes place see Alexiadou (1994, 1995) for Greek, Zubizarreta (1994) for Spanish. Crucially, VOS orders in NSLs cannot be analysed as similar to object shift constructions in Icelandic. Note that the availability of object shift is argued to correlate with the availability of Spec,TP as a subject position in Germanic (cf. Jonas & Bobaljik 1993).

ii) postverbal subjects in NSLs (Greek/Spanish) occur with all eventive predicates (transitives/intransitives) as indicated in (5):

(5) a. efige o Janis — *unaccusative*
left-3S the-John-NOM
'John left'

b. epekse o Janis — *unergative*
played-3S the-John-NOM
'John played'

c. ektise i Maria to spiti — *accomplishment*
built the-Mary-NOM the-house-ACC
'Mary built the house'

d. kerdise i Maria ton agona — *achievement*
won the-Mary-NOM the-race-ACC
'Mary won the race'

e. egrafe i Maria to grama olo to proi — *process*
wrote-IMP the-Mary-NOM the-letter-ACC all the morning
'Mary was writing the letter the whole morning'

In English on the other hand, inverted subject constructions display an intransitivity constraint (cf. Levin and Rappoport 1995 a.o. for a recent discussion).

iii) VS orders in NSLs do not display any Definitiness Restriction (DR) effects unlike English/Icelandic/Dutch/French e.t.c.. Thus, (6) with a postverbal strong DP (a universal quantifier in this case) is grammatical in Greek but not in English:

(6) a. irthe kathe pedi — *Greek*
arrived every child
'Every child arrived'

b. *There arrived every child — *English*

iv) in VSO orders in NSLs the subject is VP internal, unlike Irish (cf. McCloskey 1994, Carnie 1993) and Icelandic (cf. Jonas & Bobaljik 1993). Evidence for this claim is provided by the following set of constructions:

a) periphrastic tenses (cf. Alexiadou 1994, Anagnostopoulou 1994) show that subjects remain VP internal; in (7a-b) we see that the subject must follow the participle in Greek, but not in Icelandic (7c) or Irish (7d). Alexiadou (1994) and Anagnostopoulou (1994) have independently provided evidence that the participle moves out of the VP in Greek. Note that it is not the case that strict adjacency is required between the auxiliary and the participle as adverbs may intervene:

(7) a. an ehi idi figi o Janis *Greek*
if has already left the-John-NOM
'If John has already left...'

b. *an ehi idi o Janis figi

c. pad hafa sennilega margir studentar lesi_ bokina *Icelandic*
there have probably many students read the book
'Many students have probably read the book'

d. Ta an teangeolai ag ol na beorach *Irish*
bePRES the linguist PROG drink.DVM the beer
'The linguist is drinking the beer'

b) aspectual adverbs, which as it has been argued in Alexiadou 1994 are situated in Spec, AspectP, precede subjects in Greek but follow subjects in Irish. Moreover, VS sequences in VSO orders may be interrupted by adverbials in Greek unlike Irish (cf. 9):

(8) a. an diavaze sinithos o Janis *Greek*
if read usually the-John-NOM
'if John usually read'

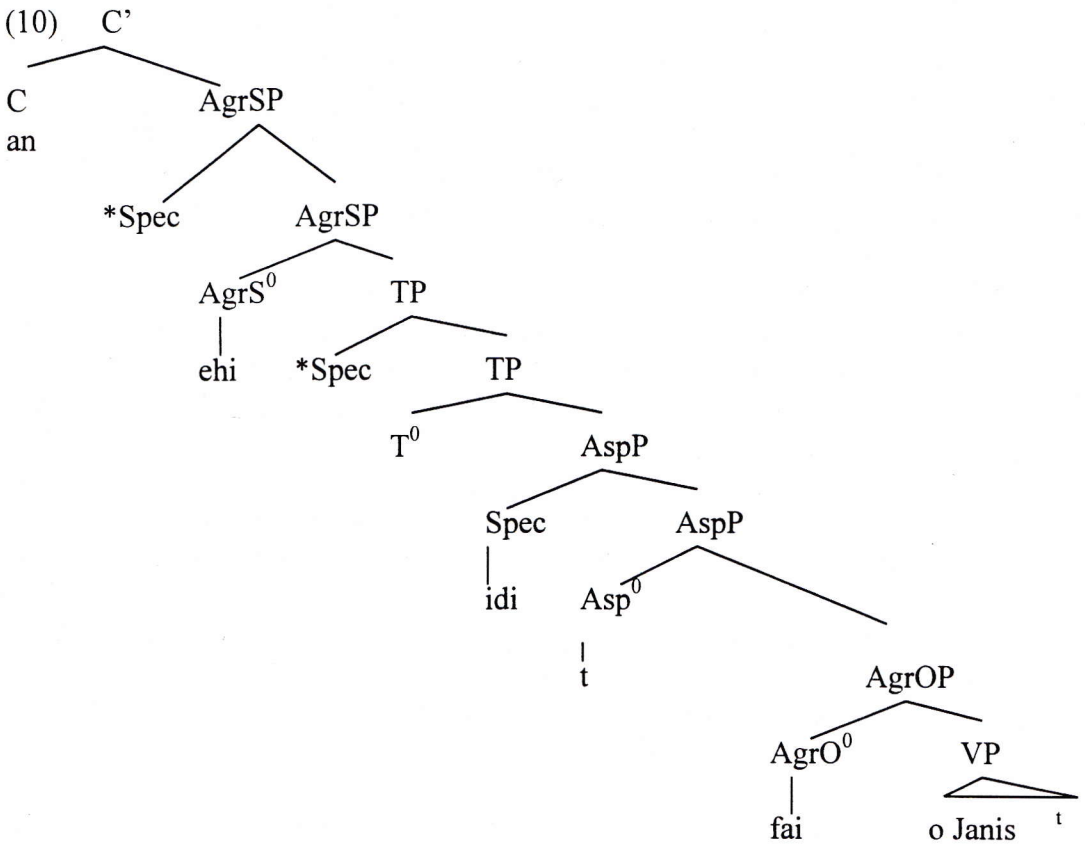
b. *an diavaze o Janis sinithos
if read the-John-NOM usually

c. an pandreftike ktes i Maria ton Petro
if married yesterday the-Mary-NOM the-Peter-ACC
'Yesterday Mary married Peter'

(9) deireann(*i geona) siad (i geona) o paidir roimh am lui *Irish*
say always they always a prayer before time lie
'they alwas say a prayer before bed-time'

The above facts can be straightforwadly accounted for if we assume a phrase structure as the one in (10). The auxiliary is generated in Asp° and is moved to AgrS°, the participle is raised from within the VP to AgrO° (as it has been argued on the basis of participial agreement facts for French), the adverb is located in Spec,AspP and the subject remains in VP internal position:

(10)



From the discussion so far we conclude: VSO orders cannot be analysed as involving I-to-C movement, since there is an absence of root vs. embedded asymmetries. Additionally, we have shown that subjects in these constructions are VP internal. Given that no overt subject movement is triggered, the obvious conclusion would be that the N features of T/Agr are weak in Greek/Spanish.

C. Two analyses within MPLT

The existence of word order alternations is problematic for a framework that does not permit optional operations. Crucially, if (1a) is analysed as involving subject movement to Spec,IP, then why is the alternative in (1b) altogether possible? Depending on whether one assumes the proposal in Chomsky (1993) or the most recent one in Chomsky (1995), the above alternations can be dealt within the Minimalist Program in two ways:

I. Within the framework of Chomsky (1993), where i) movement is regulated by the strong version of the Principle of *Greed* and ii) there is no clear way to define reference set which determines more economical derivations, there are the following potential solutions to the SVO/VSO puzzle:

a) One option would be to assume that the N-features of T/Agr are optionally strong, thus deriving SVO and weak, thus deriving VSO. Similar proposals have been made in Chomsky 1993 for Arabic which exhibits a similar alternation with different agreement patterns in each case, and in Branigan 1992, Branigan & Collins 1993 for Object Shift. However, there seems

to be no independent evidence that this is true at least in Greek/Spanish. Additionally, the possibility of developing a more restrictive theory should be preferred.

b) The other option would be to assume that VSO actually reflects the strength of the N features of Agr/T. In other words, in VSO orders there is no subject movement involved. If this is so, then SVO is best analysed as involving Topicalization (in the sense of **Left Dislocation**).

II. Within the framework of (1995) which we assume throughout, the *reference set* which determines optimal derivations is clearly defined: in evaluating derivations for economy only alternatives with the same *Numeration* are considered. An additional claim in Chomsky (1995) is that the EPP can be seen as involving checking of a categorial nominal feature. SVO in languages like English/Icelandic is related to EPP. Given that the nominal features of I° are strong in English/Icelandic, overt subject movement has to take place. Expletives are also inserted to check the strong feature of I°. Assuming Chomsky's definition of the reference set,[3] we would like to point out that the two derivations, the one with the expletive and the one without cannot be compared: SVO is derived from a numeration without an expletive. Expletive Constructions on the other hand, are derived from a numeration with an expletive. Expletive Merge is less costly than overt Movement of the subject when both are part of the same numeration.

In this framework there are again two potential solutions to the SVO/VSO puzzle:

a) One is to assume that VSO in Greek involves an expletive, i.e. it is really pro_{expl} V S O as traditionally assumed (cf. Rizzi 1982). This would lead us to conclude that EPP is strong in Greek and thus one has to analyse SVO orders as involving EPP driven movement, Case and Agreement being checked as free riders. The implication of this solution would be that SVO in Greek and English should behave alike. The crucial assumption behind this is that VSO and SVO cannot be compared as they involve different numerations.

b) An alternative would be to propose that VSO does not involve an expletive. That would lead us to assume that Greek does not respect the EPP (Case and Agreement being checked covertly). In that case, SVO should be analysed as involving Left Dislocation. Under this proposal, SVO in Greek and English are expected to behave differently. The crucial assumption this analysis relies on is that SVO as involving subject Movement to Spec,IP will always be ruled out as a Procrastinate violation given that VSO/SVO have the same numeration. In other words, VSO should be always preferred by the computational system.

In this paper we will argue that (b) is on the right track but crucially we will propose that EPP *is* strong. We will proceed as follows: comparing Greek/Spanish to Germanic, first we will provide independent evidence that SVO involves Left Dislocation. Then we will show that VSO orders can be analysed as lacking an expletive. We will propose that although both facts about SVO/VSO orders are true, Greek/Spanish respect the EPP. They differ from the germanic languages in that they choose another mode to satisfy it.

[3] But see Fox (1994), Reinhart (1995) for alternative proposals.

2. Evidence for (b)

A. SVO = Left Dislocation (LD)

If SVO in Greek/Spanish involves a dislocated subject,[4] then we would expect these orders to behave differently in the two language groups. As a matter of fact, this is correct:

i) First of all note that in Greek, for which it has been argued that it involves V-raising to AgrS°, SVO does not involve a Spec-head configuration. It is important to notice that Greek allows multiple dislocations. As we can see in (11), adverbs intervene between subjects and verbs in Greek but not in English:

(11) a. O Janis xtes meta apo poles prospathies sinandise ti Maria
the-John-NOM yesterday after from many efforts met the-Mary-ACC
'John finally met Mary yesterday'

b. *John yesterday has met Mary

Moreover, subjects can precede complementizers in Greek, but not in English:

c. o Janis an erthi
the-John-NOM if comes
'John if he comes'

d. *John if comes

ii) In Spanish, where multiple topicalizations are not allowed, other elements compete for the preverbal position (cf. Zubizarreta 1992, Ordoñez and Treviño 1995). Thus (b) is ungrammatical as both the adverb and the subject compete for the topic position:

(12) a. Temprano salia Julia de casa
early left Julia the house

b. *Temprano Julia salia de casa

iii) The interpretation of QPs/Indefinite preverbal subjects is different in the two language groups (cf. Sola 1992 for Catalan and Barbosa 1994 for Romance):

(13) a. Enas heretise ti Maria
one greeted the-Mary-ACC *strong(partitive/specific)*
'A certain person/one of the people greeted Mary'

b. heretise enas ti Maria
'Someone greeted Mary'

c. ?Enan ton heretise i Maria
one-ACC cl-ACC greeted the-Mary-NOM
'Mary greeted one of the people'

[4] See also Philippaki-Warburton (1985), Tsimpli (1990) a.o.

The preverbal subject has strong (partitive/specific) interpretation in (13a), but weak in (13b). This is not the case in English where the QP is ambiguous. The subject in (13a) behaves like the clitic left dislocated (CLLDed) object in (13c).

Similar observations can be made concerning the scope of existential quantifiers:

(14) a. kapjos fititis arhiothetise kathe arthro
some student-NOM filed every article

b. arhiothetise kapjos fititis kathe arthro

c. kapjo pedi to eksetase kathe kathigitis
some child -ACC Cl-ACC examined every professor-NOM

Wide scope of *kapjos* over the universal QP is strongly preferred in preverbal position, whereas narrow scope is preferred in postverbal position. The subject in (14a) behaves like the CLLDed object in (14c) with respect to scope possibilities.

One potential objection against the claim that preverbal subjects are left dislocated in (13, 14) might be that that quantifiers are generally not assumed to be able to occur in left dislocated positions, because these positions mark 'topichood' and quantifiers/indefinites are generally not capable of functioning as topics. However, note that at least in Greek quantifiers/indefinites are clearly permitted in positions involving LD; in (15) the bare quantifier *kapjos* precedes a CLLed object; thus, the subject is unequivocally left dislocated.

(15) Kapjos ton Petro ton sinelave
someone-NOM the-Peter-ACC Cl-ACC arrested-3S
'Someone arrested Peter'

iv) Another piece of evidence pointing at the same direction as the previous examples comes from Relative Clause Extraposition. As observed in Cinque 1982, (and see Barbosa 1994, Kayne 1994 for a more recent discussion) Relative Clauses do not undergo extraposition in NSLs, as opposed to Germanic (cf. 16a vs. 16c). Extraposition, as known, is blocked when the 'head' of the relative clause is a definite (cf. 16b):

(16) a. A man came that wanted to talk to you
b. *The man came that wanted to talk to you

c. *Enas andras irthe pu ithele na su milisi
a man came that wanted SUBJ you-GEN talk-2SG

For Kayne (1994), who analyses relative clauses as CPs complements of a determiner, *a man* forms a unit which raises further, stranding its clause. However, *the* and *man* do not form a unit as *the* is located under D° and *man* in Spec,CP. Hence, further raising is not possible. Kayne attempts to account for the contrast between NSLs and non-NSLs by argueing that the preverbal position in NSLs cannot tolerate QP subjects. This line of explanation can be restated in theory neutral terms in terms of the *Specificity Constraint* of Fiengo & Higginbotham 1991. Preverbal indefinites in NSLs are, as we saw, specific, and extraposition is expected to be ungrammatical.

v) The following asymmetry between pre- and post verbal subject pronouns indicates that preverbal subject positions have a different status from postverbal ones. As pointed out in Sola (1992) and Barbosa (1994) pronouns can be bound only in the latter. This is shown in (17) with a catalan example. This test cannot be reproduced for Greek because Greek doesn't have third person personal pronouns making use of demonstratives instead: demonstratives cannot be construed as bound variables (they are subject to principle C):

(17) a. *Tots els estudiants$_i$ es pensen que ells$_i$ aprovaran *Catalan*
all the students think that they will-pass
'All the students think that they will pass'

b. tots els jugadors$_i$ estan convencus que guanyaran ells$_i$
all the players are persuaded that will-win they'
'All the players are persuaded that they are the ones who will win'

Sola (1992) and Barbosa (1994) account for these facts on the basis of the assumption that only post verbal subjects occupy an A-position, thus being able to be construed as bound variables.

vi) The behavior of preverbal subjects in 'Triggered' inversion constructions (cf. Torrego 1984, Canac Marquis 1991, Anagnostopoulou 1994) indicates that they behave similarly to topics:

(18) a. Pjon (*o Petros) ide (o Petros)?
whom (the-Peter-NOM) saw (the-Peter-NOM)

b. Pote (o laos) apofasise (o laos) na andidrasi?
when (the-people-NOM) decided (the-people-NOM) SUBJ react
'When did the people decide to react?'

c. Pjon apo tus filus tu (o Petros) agapai (o Petros)
whom from the friends his (the-Peter-NOM) loves (the-Peter-NOM)
perisotero?
more
'Which one of his friends does Peter like most?'

(19) a. *Pjos ton Petro ton ide?
who the-Peter-ACC Cl-ACC saw
'Who saw Peter?'

b. Pote tin tenia tin provalan ja proti fora?
when the-movie-ACC Cl-ACC showed-3PL for first time
'When did they show the movie for the first time?'

c. Pjos apo tus fitites tin askisi tin elise
who from the students the excersice-ACC Cl-ACC solved-3SG
amesos?
immediately?
`Which one of the students solved the excersice immediately?'

As (18) and (19) show subjects/CLLDed objects are not allowed to interfere between the wh-phrase and the Verb when the fronted element is a non D-linked argument. Torrego (1984) and Canac Marquis (1991) analyse this as a Subjacency effect which Anagnostopoulou (1994) attributes to the status of preverbal subjects as LDs.

From the above discussion we conclude that SVO involves LD.

Before concluding the present section, we would like to discuss a number of aspectual restrictions on word order (stative vs. eventive, cf. Vendler 1967) that occur in NSLs. These facts can be viewed as evidence for the A' character of preverbal subjects in Greek/Spanish, and they also provide support for the claim that VSO orders can be analysed as lacking an expletive. Consider the sentences in (20):

(20) a. *misi/agapai/fovate/kseri i Maria ton Petro
hates/loves/fears/knows the-Mary-NOM the-Peter-ACC

vs. (1a)=b. pandreftike o Petros tin Ilekrta
married the-Peter-NOM the-Ilektra-ACC
'Peter married Ilekrta'

c. i Maria misi/agapai/fovate/kseri ton Petro
the-Mary-NOM hates/loves/fears/knows the-Peter-ACC

but: (21) misise/agapise I Maria ton Petro
hated-PERF-3S/loved-PERF-3S the-Mary-NOM the-Peter-ACC

Stative predicates do not permit VSO orders (20a), unlike eventive ones (20b). Whenever perfective aspect is present on the verbal morphology, VS-orders with stative predicates become possible (21). Notice, though, that the meaning of the verb in (21) changes: "loved" is understood as "fell in love" (episodic reading).

It has been observed that VSO orders in NSLs are understood as answers to the question "what happened" (cf. Philippaki 1985, Comorovski 1991, Anagnostopoulou 1994, Zubizarreta 1994). SVO orders are unacceptable in these contexts (cf. 22).

- what happened?

(22) a. molis espase o Janis tin kristalini lamba
just broke the-John-NOM the crystal lamp
'John just broke the crystal lamp'

b. *molis o Janis espase tin kristalini lamba

The generalization appears to be that only non-stative stage level-predicates can appear in VSO orders in NSLs. According to Zubizarreta (1994), this has to do with the fact that VSO-orders in NSLs correspond to all-focused sentences because of the way focus propagates from the object to the subject when both are VP-internal. Statives cannot appear as answers to the question "what happened" , as they are inherently incompatible with these contexts. Generic sentences are also expected to be excluded: they correspond to categorical judgements, they

are non-stage level (cf. Kuroda 1972, Ladusaw 1993). In fact, this prediction is borne out, as the following examples show. Generic readings are suppressed under VSO (cf. 23a vs. 23b):

(23) a. I gata kinigai pondikia *generic*
the-cat-NOM chases mice-ACC
'Cats chase mice' or 'The cat chases mice'

b. kinigai i gata pondikia *cannot be generic*
chases the-cat-NOM mice-ACC
'The cat chase mice'

The Greek examples in (23) are strongly reminiscent of Japanese generic sentences which always have the topic marker *wa* as shown in (24). The presence of a different marker (i.e. the nominative marker *ga*) forces a non-generic interpretation. Greek differs from Japanese in that it expresses the same distinction with the choice of a specific word order:

(24) a. Inu **wa** hasiru *Japanese*
Dogs TOP run
'Dogs run'

b. Inu **wa** neko o oikakeru
Dogs TOP cats chase
'Dogs chase cats'

c. Inu ga neko o oikakete iru
'The dog is chasing a cat'

We propose that (20c) should be analysed as Left Dislocation which is an obligatory process with statives due to the special discourse function associated with VSO orders and the inherent incompatibility of statives with this function. This instance of LD is a process of de-focusing in the sense of Reinhart (1995), necessary to avoid the clash that is produced from the movements that take (V-movement) or do not (DP-movement) take place for reasons of feature checking and the discourse function of the structure. Under this analysis, whenever morphologically trigerred movements give rise to "inappropriate" information structures, LD of the subject or the object are expected to apply. Witness the following example:

(25) ton Petro ton misi/agapai/fovate i Maria
the-Peter-ACC cl-ACC hates/loves/fears the-Mary-NOM
'Peter Mary hates/loves/fears'

In (25), left dislocation of the object has applied. The structure is as acceptable as (20c) and they both contrast with (20a).

The same facts could be accounted for under the *pro*VSO hypothesis by appealing to the function of the expletive *pro*: *pro* can be viewed as an expletive included in a numeration associated only with certain readings (and only with certain predicates, namely eventive ones). However, there are two main objections to this analysis:

(a) The postulation of *pro* is not really necessary under the theory of the interaction between phrase structure and information-structure developed in Zubizarreta (1994), which is needed for independent reasons.
(b) If we take this line of reasoning, we will be in trouble with example (25) where an expletive *pro* would have to be assumed (since NSLs would qualify as strong EPP languages): this *pro* would be associated with stative predicates and its role would be completely vacuous.

So far we have shown not only that there is evidence that SVO involves LD but also that VSO can be analysed as lacking an expletive. Let us now see whether we have reasons to assume the presence of an expletive in inverted constructions.

B. VSO = VSO

Is there independent evidence for assuming that VSO orders do (not) involve expletive Merge? Note that according to the tradition (cf. Rizzi 1982 and related literature) inverted (VOS) constructions involve an expletive *pro*. Chomsky (1995) adopts this analysis. For NSLs it is difficult to decide how to analyse VSO orders since the expletive is not overt. A potential argument for the presence of *pro* in VSO orders would be the existence of DR effects.
As known, DR effects show up with 'there' type expletives and 'il' type expletives across languages (cf. 26):

(26) a. There arrived a man/*the man/*every man *English*

b. il est arrive un homme/*l'homme *French*
c. er heeft iemand/*Jan een huis gebouwd *Dutch*
there has someone/Jan a house built

However, it has been observed (cf. Jaeggli 1980, Rizzi 1980, Burzio 1981, Chomsky 1981, Safir 1985, Calabrese 1990 and see (28)) contra Belletti (1988) that DR effects are absent in NSLs. The fact that in unaccusative (and some unergative) constructions DR effects do not show up in all languages shows that in these constructions, DR effects are syntactically triggerred, but not in existential and 'donkey anaphora' contexts. For this reason, we will adopt Chomsky's (1995) and Frampton's (1995) analysis of DR effects in unaccusative and transitive expletive constructions, namely that they arise because 'there' is a Determiner which takes an NP complement, hence the DR effects (cf. Chomsky 1995, Frampton 1995). Consider now the sentences in (27):

(27) a. eftase ena pedi/ o Jorgos/ kathe filos mu *Greek*
arrived a child-NOM/the-George-NOM/every friend mine
'A child/John/every friend of mine arrived'

b. diavase ena pedi/kathe pedi to vivlio
read-3S a child/every child the-book-ACC
'A/every child read the book'

c. pad lasu einhverjir studentar bokina *Icelandic*
there read some students the book
'Some students read the book'

There is a sharp contrast between (27a,b) and (27c): the Icelandic TEC in (27c) shows obligatory DR effects. The systematic absence of DR effects in NSLs, on the other hand seems to suggest that there is no expletive in inverted constructions in these languages.

Of course, it could be claimed that the presence vs. absence of DR effects is related to the nature of the expletive: overt vs. covert. Greek has a covert expletive, and as a result it lacks DR effects, Icelandic has an overt expletive and, therefore, it shows DR effects. However,

(i) there are Arabic dialects which do not display any DR effects with covert expletives and there are languages like Dutch which display DR effects with covert expletives (Riny Huybregts p.c.). Hence, the correlation between the (c)overtness of the expletive and the presence of DR-effects seems to break down.

(ii) Moreover, Greek and Spanish (cf. Suñer 1982, Safir 1985) have constructions for which we have reason to assume that they are null expletive constructions: these are "impersonal-have" constructions which show default agreement and DR effects. The reason why it can be argued that an expletive is present in these constructions is that the overt argument has accusative Case, so there must be some covert NP in the structure receiving Nominative Case:

(28) ehi anthropus/*kathe anthropo edo
has peopleACC/every person-ACC here

If an expletive is present in (27a, 28) then we must conclude that Greek/Spanish have two covert expletives with completely different properties: one associated with DR effects and one which does not. It is not clear why Greek and Spanish are like that while, for instance, French has DR effects in both types of constructions (cf. *il est arrive un homme* vs. *il y a des enfants)*.

Notice that whenever we have reasons to postulate a null expletive in Greek we always have default agreement (cf. 29):

(29) a. fenete oti tha erthun
seems that FUT come-3PL
'It seems that they will come'

b. prepi/ bori/ na erthun ta pedja
must-3S/might-3S SUBJ come-3PL the-children-NOM

Naturally, it is possible to postulate several types of expletives, as has been done for English (*it* vs. *there*). The question is whether it is necessary to analyse VSO orders as a transitive expletive construction.

McCloskey (1994) has reached a similar conclucions for Irish; the lack of DR with unaccusative constructions and with transitives seems to indicate that no expletive is present:

(30) deireann siad i geona o paidir roimh am lui
say they always a prayer before time lie
'they alwas say a prayer before bed-time'

We believe that VSO orders are not TECs, i.e. Greek VSO orders should not be assimilated to Icelandic VSO orders. Jonas & Bobaljik (1993) have established a correlation between the availability of subject inversion with transitive predicates and the availability of Spec,TP. In English inverted orders the subject must remain VP-internal and there is always an intransitivity constraint on inverted constructions. We have shown that Greek and Spanish are like English in that they always have VP-internal subjects (cf. the position of subjects in participial constructions (7)). Moreover, Greek and Spanish, like English, do not display scrambling/object shift of the Germanic type. Hence, they are not expected to license TECs, if Jonas and Bobaljik are right.

Jonas & Bobaljik divide languages into two types: those that license Spec, TP and those that do not.[5] In Greek/Spanish/English, Spec, TP as a subject position is not licensed, thus there is only one (if any for NSLs) external specifier position for subjects. This position is related to the EPP, Case being checked as free ride in English. In Icelandic, Spec, TP is licensed thus, there are 2 VP external Spec positions for subjects, 1 devoted to Case and 1 to the EPP. Moreover, Celtic is assumed to be a language that licenses Spec,TP as a subject position, but the licensing is devoted to Case not to EPP (cf. Carnie 1993).

So far we have shown that a) SVO = LD and b) VSO = VSO. That SVO involves LD does not necessarily imply that VSO lacks an expletive, because one might assume that there is always an expletive. That VSO lacks an expletive implies that SVO involves LD, since lacking an expletive the language would qualify as a no-EPP language and SVO could not be analysed as EPP-driven movement.

Note that if McCloskey is right in his claims about Celtic VSO orders, it follows that Celtic SVO structures are necessarily Topicalization structures. As a matter of fact this is the case; evidence is provided from the obligatory presence of a special topicalization marker in Celtic SVO orders (examples from McAulay 1992). Thus, SVO is never related to EPP:

---

[5] Thrainsson (1995) and Bobaljik (1995) propose a similar but not identical division: they claim that languages vary with respect to the functional projections they instantiate. Some languages have both AgrSP and TP (Icelandic), whereas others (English) have an 'unsplit Infl' (cf. Iatridou 1990). It is argued that overt morphology determines the number of projections in the structure: a 'fused' Agr and Tense morphology indicates that the language in question has only one functional projection, whereas evidence for separate tense and agreement morphology is taken as evidence for separate agreement and tense projections. We would like to point out that Greek provides arguments against this correlation: in Greek we have overt morphological evidence for separate tense and agreement morphology, but syntactic behavior (in the relevant aspects) similar to languages where agreement and tense are not separated:

i.	Icelandic: kasta 'throw'		English: tremble		Greek: rihno 'throw'	
	Present	Past	Present	Past	Present	Past
2sg	kasta-**r**	kasta-**ði-r**	3sg tremble-**s**	tremble-**d**	3sg rihn-**i**	**e**-riks-**e**
	TWO morphemes		ONE morpheme: *tremble-d-s		TWO morphemes	

For this reason we would like to suggest that at least for Greek, the solution outlined in Jonas & Bobaljik (1993) for English is preferred: they argue that in English functional head movement of Tense to Agr takes place prior to verb movement, thus Spec,TP is not licensed. We would like to suggest that it is presumably this movement that leads to the fusion in Greek, movement that takes place before the verb raises overtly to Agr° (see Kissock 1995).

(31) a. carr **a** bhuail an coisi *Irish*
car REL.PART hit ART pedestrian
'It was a car that hit a pedestrian'

b. Mair **a** fwrodd ef *Welsh*
Mair REL.PART hit-she him
'Mair hit him'

3. A Proposal: GB meets the Minimalist Program

Given the results of our discussion so far we may be led to two different conclusions: a) NSLs are no-EPP languages or, perhaps not so obvious, b) they are strong EPP languages where the EPP feature is not checked by Move/Merge XP but by a different mode. We believe that the latter option is preferred and we propose that this different mode is **V-movement**. To make this proposal work we need to capitalize on the nature that verbal agreement morphology in NSLs as traditionally assumed and combine it with the ideas about the nature of EPP in Chomsky (1995).

The basic intuition in the GB literature about NSLs is that they have **(pro)-nominal agreement** (cf. Taraldsen 1978, Rizzi 1982, Chomsky 1981, Safir 1985 a.o.). In Chomsky 1995, EPP is seen as checking of a nominal feature in I°. Thus, it can be claimed that *pro-drop* languages have agreement properties that permit them to satisfy the EPP via verb-raising. Crucially, then V-movement is sufficient to check the nominal feature of I°, since the verb itself contains the nominal feature needed. In our proposal EPP checking is reduced to Agr checking in the sense of [nominal] feature checking (not Case).

That the verbal agreement morphology in NSLs includes a nominal element is claimed on the basis of the following example. In (32), the agreement affixes play exactly the same role as the pronouns in the English paradigm:[6] [7]

---

[6] The implication of our proposal is that referential *pro* does not exist either. Ouhalla (1994) has reached a similar conclusion to the one presented here. He argues that the EPP is related to the relative richness/impoverishment of agreement morphology. English has to insert an overt expletive in subject positions not filled with a noun phrase argument to check the EPP feature. NSLs do not have to insert an expletive because the features of AgrS are 'identified' in terms of agreement morphology. See also Philippaki- Warburton (1989).

[7] A question that arises is how our proposal can be implemented to account for control and raising structures. One potential answer would be to assume mechanisms of feature percolation common in computational frameworks. ECM constructions might raise potential problems for the analysis outlined in the text. As Lasnik (1995) points out, ECM constructions show EPP effects, even though the subject of an ECM infinitive is not a case position. For Spanish, it is not clear what the prediction would be, since the embedded verb is infinitival and thus the agreement is of a different type. However, Greek lacks embedded infinitival complements; thus in ECM constructions the embedded verb is inflected for person and number (cf. Iatridou 1993). These constructions should be analysed in terms of Case, because crucially the subject does not have to raise in ECM contexts requiring a clitic (cf. Schneider-Zioga 1994). It is not an EPP problem, otherwise we would not expect this alternation:

i. a. perimeno ton Petro na erthi
expect-1S the-Peter-ACC SUBJ come-3S
'I expect Peter to come'

b. ton perimeno na erthi ton Petro
cl-ACC expect-1S SUBJ come-3S the-Peter-ACC

Similar effects obtain in double object constructions where when the dative is doubled by a clitic, it cannot move, while when it is not it must be adjacent to the verb:

(32)	a.	I love	b.	agapo
		you love		agapas
		he loves		agapa
		we love		agapame
		you love		agapate
		they love		agapane

What we are suggesting is that verbal agreement morphology has the status of a clitic which is part of the lexical verb and not of I° (cf. Safir 1985). An alternative possibility would be to derive the nature of verbal agreement from pronoun incorporation (cf. Hale 1987, Taraldsen 1993 a.o.). We believe that the basic problem for the incorporation analysis is the status of the subject argument. One would have to assume that the subject is somehow peripheral to the VP. Under our proposal, what we actually have is a doubling effect.[8] It would be interesting to examine the differences and similarities between this type of doubling and object doubling which occurs in Greek and Spanish.[9]

However, it seems that examining the agreement affixes only is not enough (cf. Trentino/Fiorentino, Brandi & Cordin 1989):

	Fiorentino	Trentino	
(33)	(e) parlo	parlo	I speak
	tu parli	te parli	you speak
	e parla	el parla	he speaks
	la parla	la parla	she speaks
	si parla	parlem	we speak
	vu parlate	parle	you speak
	e parlano	i parla	they (masc.) speak
	le parlano	le parla	they (fem.) speak

(34) a. *(tu) parli *Fiorentino*

b. Mario e parla
"Mario speaks'

c. *Mario parla

d. gli e venuto la Maria

---

ii. a. edosa tu Petru to vivlio
gave-1S the-Peter-GEN the-book-ACC
'I gave Peter the book'

b. *edosa to vivlio tu Petru

c. tu edosa to vivlio tu Petru
cl-GEN gave-1S the-book-ACC the-Peter-GEN

We believe that whatever the solution for the doubling constructions is, it will also apply to the ECM examples.

[8] Ordoñez (p.c.) informed us that he has reached a similar conclusion about Spanish.

[9] Crucially, Sportiche's Filter (1993) would apply to both Agr Projections. Note that Greek and Spanish which are the 'most well-behaved' wrt to VSO also have object clitic doubling. Italian and Catalan on the other hand which do not fit the classification, lack object clitic doubling and show VOS orders, VSO ones being marginal. VOS orders in Italian/Catalan seem not to be similar to the VOS orders in Greek/Spanish, but to the Object Shift constructions in Icelandic. If this is really so, then it correlates with the fact that they seem to show some 'strong EPP' properties.

is come the Maria

Trentino/Fiorentino are similar to Greek with respect to the lack of DR effects, but they differ in that subject clitics are present. The verbal agreement in Trentino/Fiorentino is similar to Italian, yet the clitics are obligatory.The above paradigm shows that it is not enough to look at the richness of the paradigm to conclude that Agreement is pronominal or not, but one has to look at the syntactic properties of the language.

A potential counterargument to the correlation between agreement and subject clitics we are trying to establish might be French, a language that has subject clitics but is similar to English in all other relevant aspects (Word Order, DR effects, e.t.c):

(35) a. Il mange
b. Jean, il mange
c. Jean mange
d. *il est arrive Jean

However, we believe that this case is not problematic: we assume that there is a distinction between two types of subject clitics (cf. Cardinaletti & Starke 1994 a.o.), the French type with **XP properties** and the Fiorentino/Trentino type with **X^0 properties** (cf. Brandi & Cordin 1989 a.o.). It follows that these two types behave differently and crucially, only the Fiorentino/Trentino clitics have the same effects as agreement affixes.

The implications of the proposal outined here are: a) EPP is universally strong because it is a formal property of sentences, b) the mode of EPP satisfaction is parametrized, c) EPP triggers V-raising in NSLs and d) linguistic variation reduces to morphological variation.

4. Typology

From the above discussion it became clear that there are two parameters which regulate word order variation in the IP domain:

a) **the Spec, TP parameter** (cf. Jonas & Bobaljik 1993, Carnie 1993)
b) **the EPP parameter** (XP vs. X^0)

Tense is linked to Case and is EPP linked to the Agr, all movement being triggered by [-interpretable] features of I°. Combining these two parameters we arrive at 4 possible language types:

(36)	EPP(XP)	Spec,TP
a.	+	-
b.	+	+
c.	-	-
d.	-	+

These combinations have the following properties: (a) will have i) ECs with intransitivity and VP-internal subjects ii) DR effects. (b) will have i) TECs with external subjects ii) DR effects. (c) will have i) VSO with internal subjects, ii) no DR effects, (d) will have i) VSO with external subjects ii) no DR effects. We believe that (a) is English, (b) is Icelandic, (c) is Greek/ Spanish and (d) is presumably Celtic.

If (d) is Celtic, then we can explain why it is never SVO (unlike Icelandic): the subject will never have to move beyond Spec,TP to satisfy the EPP. Additionally, we expect (c) to differ from (d) in that only in (c) word order would be highly sensitive to aspectual restrictions. This is in fact true. Moreover, we expect that different language types will emerge as soon as interactions between the complementizer and the INFL system come into play. A potential candidate would be Arabic.

In Celtic and the Italian dialects of Trentino and Fiorentino,[10] different agreement patterns occur depending on the presence or position of an overt subject respectively. In the Italian dialects, fully referential agreement is used in SVO orders, whereas default, i.e. 3sg agreement is used in VSO orders (cf. 38a-b). In Celtic, default, 'analytic', agreement is used in VSO orders and fully referential one, 'synthetic', in structures lacking an overt lexical subject (cf. 38c-d), i.e. subject/verb agreement and overt subjects are in complementary distribution:

(38) a. Mario e parla *Fiorentino*
"Mario speaks'

b. gli e venuto la Maria
is come the Maria

c. canodd/* canasant y bechgyn yn yr eglwys *Welsh*
sang-3S sang-3PL the boys in the church

d. canasant/*canodd yn yr eglwys
cang-3PL / sang-3S in the church

We believe that these agreement patterns do not affect our proposal concerning the nature of the EPP checking. The data from the italian dialects can in fact be seen as supporting the view that preverbal subjects are topics; as Barbosa (1994) points out in a language where there is a distinction between referential and default agreement it is expected that the former will always occur with dislocated subjects. Agreement is 'inherently' nominal. In NSLs, where fully inflected paradigms exist, it is expected that there is going to be a default form used in specific contexts.[11]

The Celtic paradigm, however, demands a different explanation, which is related to the ability of verbal morphology to check case: we would like to suggect (cf. Roberts and Shlonsky 1994 for similar ideas) that in (c) the verbal head does not check case (it checks the EPP only), thus the subject can overtly occur. On the other hand in (d) the verbal head checks case (and the EPP) thus the presence of an overt DP is not licensed. Case-checking is mediated via incorporation: it has been argued that in Celtic synthetic forms are derived from pronoun incorporation (cf. Taralsden 1993, Roberst & Shlonsky 1994 a.o.) and it has been

[10] Many thanks to Ian Roberts for bringing this 'asymmetry' to our attention.

[11] In a way this is what happens with the greek examples that show 3sg verbal morphology. A potential and rather speculative explanation for the use of the alternation in these dialects might be the following: the dialects lack object clitic doubling; we have hinted that VSO orders and object clitic doubling are parallel. Now if a language lacks doubling altogether it will not be able to use the doubling forms for the subject either.

claimed, independently, that incorporation is an alternative (to spec-head) case checking mechanism provided by UG (cf. Baker 1988, Rizzi & Roberts 1989, Sportiche 1993).[12] This suggestion does not contradict our claim that agreement/EPP checking is not linked to Case checking; the nominal properties which are responsible for the EPP checking are not derived via incorporation. The only thing that additionally happens is that once the pronoun incorporates into the verbal head, its case is licensed (see footnote 12). Notice that an incorporation approach is more favorable for Celtic since, as expected, -given that Spec-head and incorporation are two alternative mechanisms-, when it takes place, no overt DP in Spec,TP is acceptable.

The question that immediately arises is what happens with the case of the subject in Greek/Spanish/Italian dialects, where, as we claimed, the nominal properties of agreement are not derived by incorporation either. Crucially, NSLs seem to lack both mechanisms for the licensing of Nominative Case, at least overtly. We have shown that Case is weak in NSLs and that there is no case specifier available. With respect to the case of the overt DP, one would have to say one of the following two things: either case is checked covertly (as it happens in English inverted constructions cf. Chomsky 1995) or the GB literature is right in assuming that there is no Case to be checked, since nominal agr absorbs case. If we follow the latter option we are led to assume that Nominative Case is a default case in Greek/Spanish and as such it need not be checked (cf. Chomsky 1995).[13] Nominative Case is not a default case in Celtic though, thus the various patterns.

5. Conclusion

In this paper we have examined word order alternations across languages, comparing Greek/Spanish to Germanic. We have shown that SVO orders in the former involve LD, whereas they involve EPP driven movement in the latter. Moreover, VSO orders lack an expletive in NSLs. We argued that in spite of these facts, NSLs respect the EPP and check it via V-raising due to the nominal properties of their verbal agreement morphology. Combining the EPP parameter with the Spec,TP parameter proposed in Jonas & Bobaljik (1993), we arrived at four language types with a specific set of properties, which we were able to identify.

References

Alexiadou, A. (1994) *Issues in the Syntax of Adverbs*. Ph. D. Universität Potsdam.

Alexiadou, A. (1995) 'Word Order Alternations in Modern Greek'. ms. FAS, Berlin.

Anagnostopoulou, E. (1994) Clitic Dependencies in Modern Greek. Ph.D. University of Salzburg.

Baker, M. (1988) *Incorporation*. Chicago University Press.

---

[12] Presumably the incorporation alternative is chosen as the pronominal element has a clitic like status, hence it must actually incorporate (cf. Chomsky 1995).

[13] This claim can be reinforced by the fact that nominative Case is the case used in dislocation:

i. o Janis, den ton boro me tipota
the-Jonh-NOM NEG cl-ACC stand-1S with anything

Moreover, the agreement features of the DP do not have to be checked, as they are [+interpretable] and as such are visible at LF (cf. Chomsky 1995). Crucially, there is no overt or covert DP movement for reasons of feature checking.

Barbosa, P. (1994) 'A new Look at the Null Subject Parameter'. Paper presented at CONSOLE III, Venice.
Belletti, A. (1988) 'The Case of Unaccusatives' in Linguistic Inquiry 19.
Bobaljik, J.D. (1995) 'Free AGR! Parametrization in Germanic Morphosyntax' Paper presented at the 11th Comparative Germanic Workshop.
Bobaljik, J.D. & A. Carnie (1992) 'A Minimalist Approach to Some Problems of Irish Word Order' ms MIT.
Brandi, L. & P. Cordin (1989) 'Two Italian Dialects and the Null Subject Paremeter' in O. Jaeggli & K. Safir (eds.) *The Null Subject Parameter* 111-142.
Branigan, P. (1992) *Subjects and Complementizers*. Ph.D. MIT.
Branigan, P. & C. Collins (1993) 'Verb Movement and the Quotative Construction in English' *MIT Working Papers* 18: 1-14.
Burgio, L. (1981) 'Intransitive Verbs and Italian Auxiliaries' Ph.D. MIT.
Calabrese, A. (1990) 'Some remarks on Focus and Logical Structures in Italian' in Harvard Working Papers in Linguistics 91-127.
Canaq-Marquis, R. (1991) 'On the Obligatory Character of Inversion in Spanish' in *Proceedings of WCFL* 10.
Carnie, A. (1994) 'Nominal Predicates and Absolutive Case Marking in Irish' in *MIT Working Papers* 19: 131-174
Cardinaletti, A. & M. Starke (1994) "The Typology of structural deficiency: on the three grammatical classes' ms. University of Venice andUniversity of Geneva/ASG.
Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht, Foris
Chomsky, N. (1993) 'A Minimalist Program for Linguistic Theory' in *The View from Building 20* MIT Press.
Chomsky (1995) *The Minimalist Program*. MIT Press.
Cinque, G. (1983) 'On the Theory of Relative Clauses and Markedness' in *The Linguistic Review* 1 247-294.
Comorovski, I. (1991) 'Partitives and the Definiteness Effect' in *Proceedings of WCCFL* 10: 91- 102.
Fiengo, R. & J. Higginbotham. (1991) 'Opacity in NP' in Linguistic Analysis 395-421.
Frampton, J. (1995) 'Expletive Inserion', Paper presented at the Role Economy Principes in Linguistic Theory Workshop in Berlin, 9-12.2.95.
Fox. D. (1994) 'Quantifier Scope in VP Ellipsis'. ms MIT.
Hale, K. (1987) 'Incorporation and the Irish synthetic verb forms. ms
Jaeggli, O. (1980) On some phonologically null elements in Spanish. Ph.D, MIT
Jaeggli, O. & K. Safir (1989) The Null Subject Parameter. Kluwer Academic Publishers.
Jonas, D & J. Bobaljik (1993) 'Specs for Subjects' in *MIT Working Papers* 18, 59-98.
Iatridou, S. (1990) 'About AgrP' in Linguistic Inquiry 21 551-577.
Iatridou, S. (1993) 'On Nominative Case Assignement and a Few Related Things' in MIT Working Papers 19: 175-196.
Iatridou, S. & A. Kroch (1992) 'The Licensing of CP-recursion and its relevance to the Germanic Verb-second Phenomena in *Working Papers in Scandinavian Syntax*. 50.
Kayne, R. (1994) 'The Antisymmetry of Syntax'. MIT Press.
Kissock, M. (1995) 'Evidence for T to Raising to AgrS' Paper presented at the 5th Colloquiun on Generative Grammar la Coruna.
Kuroda, S.Y. (1972) 'The categorical and the thetic judgement' in *Foundations of Language* **9**: 153-185.
Ladusaw, W. (1994) 'Thetic and Categorial, Stage and Individual, Weak and Strong' in *Proceedings of SALT IV* 220-229.

Lasnik, H. (1995) Lecture Notes.
Levin, B. & M. Rappaport. (1995) *Unaccusativity: at the Syntax-Lexical Semantics Interface*. MIT Press.
McAulay, D. (1992) *The Celtic Languages*. CUP.
McCloskey, J. (1994) 'Subjects and Subject Positions in Irish' ms. Santa Cruz.
Ordoñez, F. (1994) 'Post-verbal Asymmetries in Spanish'. Paper presented at the Vienna GLOW Colloquium.
Ordoñez, F. & E. Treviño (1995) 'Los sujetos y objetos preverbalos en español.' Paper presented at the 5th Colloquium on Generative Grammar, Coruña, Spain.
Ouhalla, J. (1994) 'The Syntactic Representation of Arguments'. ms Max Planck Berlin.
Philippaki- Warburton, I. (1985) 'Word Order in Modern Greek'. in *Transactions of the Philological Society.*
Philippaki-Warburton, I. (1989) 'Subjects in English and in Greek' ms, University of Reading.
Reinhart, T. (1994) 'Wh-in situ in the Framework of the Minimalist Program' Ms, University of Utrecht.
Rizzi, L. (1980) 'Negation, wh-movement and the Null Subject Parameter' paper read at the 5th GLOW Colloquium.
Rizzi, L. (1982) *Issues in Italian Syntax*. Foris.
Rizzi, L. & I. Roberts (1989) 'Complex Inversion in French' in *Probus* 1 1-30.
Roberts, I. & U. Shlonsky (1994) 'Pronominal Enclisis and VSO Languages' *GLOW Newsletter* 32: 50-51.
Safir, K. (1985) *Syntactic Chains*. Cambridge University Press.
Schneider-Zioga, P. (1994) Ph.D USC.
Sola, J. (1992) Agreement and Subjects. Ph.D Universitat Autonoma de Barcelona.
Sportiche, D. (1993) 'Sketch of a Reductionist Aprroach to Syntacti Variation and Dependencies' ms UCLA.
Suñer, M. (1982) 'Impersonal Haber' chapter 1 of Ph.D.
Taraldsen, K. (1978) 'On the NIC, Vacuous Application and the that-trace FIlter' ms. MIT
Taraldsen, K. (1993) 'Subject/verb agreement and word order in Celtic and Romance' in *Proceedings of NELS* 23 495-504.
Thrainsson, H. (1995) 'On the non-universality of functional projections' ms. Harvard.
Torrego, E. (1984) 'On Inversion in Spanish and Some of its Effects' in Linguistic Inquiry 15.
Tsimpli, I-M. (1990) 'The Clause Structure and Word Order of Modern Greek' in *UCL Working Papers in Linguistics.*
Vendler, Z. (1967) 'Verbs and Times' in *Linguistics and Philosophy*. Ithaca,Cornell University Press.
Zubizarreta, M.-L. (1992) 'Word Order in Spanish and the nature pf Nominative case'. ms USC.
Zubizaretta, M-L. (1994) 'Grammatical Representation of Topic and Focus; Implications for the Structure of the Clause' in *Cuademos de Linguistica del I.U. Ortega y Gasse*. **2**, 181-208.

HAS BARE PHRASE STRUCTURE THEORY SUPERSEDED X-BAR THEORY?[1]

(Hans-Martin Gaertner)

Chomsky (1995, p.233f) points out that

". . . the minimalist program, right or wrong, has a certain therapeutic value. It is all too easy to succumb to the temptation to offer a purported explanation for some phenomenon on the basis of assumptions that are of roughly the order of complexity of what is to be explained. If the assumptions have broader scope, that may be a step forward in understanding. But sometimes they do not. Minimalist demands at least have the merit of highlighting such moves, thus sharpening the question of whether we have a genuine explanation or a restatement of a problem in other terms."

Clearly, assessing the order of complexity and the scope of assumptions is not an easy matter. Competing assumptions, being as a rule embedded in widely diverging theories, require some careful analysis and rephrasing before one can take a stand on which of them might be preferred. Secondly, transferring parts of an explanation into an area of theorizing that is not well-understood often has - although it broadens the scope of assumptions - the effect of immunizing a theory against serious evaluation.
In this paper, I set myself the task of asking whether the assumptions related to the "Structure Preserving Hypothesis" (SPH) are best captured in a system that has access to something like inherent X-bar status or in the bare phrase structure model of minimalist syntax, that construes the X-bar status of elements as a relational property.[2]
Let's assume that the pre-minimalist system could have stated the intended generalization in terms of the X-bar format for syntactic structures in (1):

(1)	X^{α}	$\rightarrow$	$X^{\alpha}\ Y^{\alpha}$	[$\alpha = X^{\circ}$ or X^{max}] [adjunction]
	X'	$\rightarrow$	X° (YP)	[YP = "complement"]
	XP	$\rightarrow$	ZP X'	[ZP = "specifier"]
	XP	$\rightarrow$	X° (YP)	[no specifier]

The grammar then rules out the unwanted cases on the basis of (1) interpreted as a filter on representations (cf. Chomsky 1986)[3] . We'll see below that as far

---

[1] Thanks to Bob Frank for discussing some of the issues with me. The usual disclaimers apply.

[2] There seem to be quite different opinions on the usefulness of exercises as the one undertaken here. See Zwart (1994) and Gaertner & Steinbach (1994). I would count it as a success of the short remarks to follow if the questions surrounding current work on phrase structure in generative grammar became a little more tractable.

[3] Chomsky (1986, p.4) states the following principles for substitution:

a There is no movement to complement position
b Only X° can move to the head position
c Only a maximal projection can move to the specifier position

as representations are concerned, X' can be an adjunction site in the minimalist theory.
Now, bare phrase structure theory does away with explicit statements such as (1) on principled grounds and sets out to derive the desired constraints from independently motivated assumptions.
At the core of revising X-bar theory lies the idea ". . . that bare output conditions determine the items that are "visible" for computations" (Chomsky 1995, p.242). Roughly, this means that whatever need not be available at the interfaces (PF/LF) should not play a role in the computations of C_{HL}. For the case at hand, it suffices to note that "bare output conditions make the concepts "minimal and maximal projection" available to C_{HL}. But C_{HL} should be able to access no other projections." (ibid.) Crucially, bar-levels or markings of minimal and maximal status in the form of features should not figure in syntax either. Instead Chomsky adopts the strategy of "taking these to be relational properties of categories, not properties inherent to them. [. . .] There are no such entities as XP (X^{max}) or X^{min} in the structures formed by C_{HL}, though I continue to use the informal notations for expository purposes, along with X' (X-bar) for any other category." (ibid.)
The most important definition is then given as follows:

(2) "A category that does not project any further is a maximal projection XP, and one that is not a projection at all is a minimal projection X^{min}; any other is an X', invisible at the interface and for computation." (p.242f.)

It is immediately obvious that (2) achieves something that also follows from (1). X' can neither be moved nor can it be adjoined to. As mentioned before, in the strictly bottom-up generation of pre spell-out structures of Chomsky (1995), the latter point is not correct if looked at representationally. Outputs of the form in (3) are licensed.

(3) $[_{XP}$ YP $[_{X'}$ ADJ $[_{X'}$ X° ZP]]]

Looked at derivationally no problem arises. The adjunction operation is an application of Merge that takes ADJ and $[_{XP}$ X° ZP] as its input. [X° ZP], being a projection of X° and not having projected any further at that stage, is of the status [+max, - min] (= XP). Thus [X° ZP] is visible and the adjunction can occur. It would lead us far afield if we were to go into an empirical debate on "second-effects" and possible adjunction sites relevant to whether the operation in question is desirable. "Second-effects" had been a reason for disallowing it in earlier models (cf. Chomsky 1986, p.6). The revised line on this

---

d Only minimal and maximal projections (X° and X") are "visible" for the rule Move-α. He then assumes that b and c would "follow from an appropriate form of Emonds's Structure-Preserving Hypothesis . . . ".
Restrictions on XP and X° adjunction are given on pages 6 and 73, respectively, the latter amounting to ". . . a kind of generalization of Emonds's Structure-Preserving Hypothesis . . .".

is that such effects "may belong to the phonological component" (Chomsky 1995, p.368).
Here, I can only explore the theory-internal consequences of the invisibility of X'- categories. It is interesting, to begin with, to reflect on how C_{HL} computes the c-command relation at the PF-branch in order to construct the precedence relation for terminal elements. In keeping with Kayne's LCA, this can be done according to the following principle.

(4) α precedes β iff some node which dominates α (perhaps α itself) asymmetrically c-commands some node which dominates β (perhaps β itself) (Frank & Vijay-Shanker, 1995)

Crucially, X' because of its invisibility does not c-command the specifier and consequently does not asymmetrically c-command anything dominated by YP in (5).

(5) $[_{XP}\ [_{YP}\ Y°\ WP\ ]\ [_{X'}\ X°\ ZP\ ]\ ]$

At the same time, however, X° (and ZP) must be prevented from c-commanding YP and asymmetrically c-commanding Y° and WP. Otherwise, no precedence relation is defined for the relevant terminals. Thus, Chomsky (1995, p.391, fn.110) adds the assumption that "L [= X'] is part of the structure, however; otherwise we would have a new and inadmissible syntactic object. Thus, the branching structure remains, and *m, p* [= X°, ZP] do not c-command out of L." The result is that X° and ZP do not c-command YP, etc. Now, this account seems to me to work by fiat only. Intuitively speaking, being part of the structure is not the same as being part of the *visible* structure. Conversely, not being part of the visible structure does not imply not being part of the structure.
To clarify the issue we must recapitulate what admissible syntactic objects are. For the purpose at hand we can take (6) to be a sufficient list. (ibid. p.243)

(6) a. lexical items
b. K = { γ, { α, β } }, where α, β are objects and γ is the label of K.

The object K is the result of an application of Merge (potentially as a suboperation of Move) to α and β. Chomsky further defines the functioning elements of phrase markers in the following way. (ibid., p.247)

(7) For any structure K,
a. K is a term of K.
b. If L is a term of K, then members of the members of L are terms of K.

As a matter of explication Chomsky adds that "For the case of substitution, terms correspond to nodes of the informal representations, where each node is understood to stand for the subtree of which it is the root." (ibid.)[4]
We can infer that every term is a syntactic object. (The expression "structure" in (7) might be construed to refer to syntactic objects in (6)). We need a further step to fully understand the remarks on what it means to be part of the structure. (cf. ibid., p.339)

(8) The relations of dominance and c-command are restricted to terms.

For (5) we can take the set of terms to be

(9) T = {XP, YP, Y°, WP, X', X°, ZP }

The resulting dominance(D)- and c-command(C)-relation would be the following:

(10) D = {<XP,YP>,<XP,Y°>,<XP,WP>,<XP,X'>, <XP,X°>,<XP,ZP>,
<YP,Y°>,<YP,WP>,<X',X°>,<X',ZP> }
C = {<YP,X'>,<YP,X°>,<YP,ZP>,<Y°,WP>,<WP,Y°>,<X',YP>,<X',Y°>,
<X',WP>, <X°,ZP>, <ZP,X°> }

If one now eliminates all the pairs containing X', the desired result is obtained for C. YP asymmetrically c-commands the elements dominated by X', but nothing dominated by YP is asymmetrically c-commanded by anything reflexively dominated by X'.

(11) D = {<XP,YP>,<XP,Y°>,<XP,WP>,<XP,X°>,<XP,ZP>, <YP,Y°>,
<YP,WP>}
C = {<YP,X°>,<YP,ZP>,<Y°,WP>,<WP,Y°>,<X°,ZP>, <ZP,X°> }

There is one problem though. For C to be constructed from D, C_{HL} had to be able to see the pairs containing X' in D. The strict invisibility of X', however, would require C to be constructed from D in (11). Yet, application of c-command as defined in (12) to D in (11) yields C' in (13).

---

[4] It is not entirely clear why the term-node correspondence should not hold for adjunction. "Adjunction differs from substitution, then, only in that it forms a two-segment category rather than a new category." (Chomsky 1995, p.248) As long as the definition of terms is as in (7), both segments of an adjunction structure count as terms. Differences affect the label. Thus no such thing as a term <α,α> will result. As long as (8) below holds, both segments will also - unless invisibility interferes - figure in the computation of dominance and c-command. Pages 338-340 contain a discussion of c-command relations for adjunction structures. The intention is to let adjuncts c-command outside of their adjunction site. This might be welcome for X°-elements. In structures like (3), however, no precedence order will result for ADJ. and the specifier. The subsequent discussion terminologically mixes terms and categories for the dominance relation and refrains from making a number of crucial decisions. It is therefore hard to draw any conclusions from it.

(12) X c-commands Y if (i) every Z that dominates X dominates Y and (ii) X and Y are disconnected. (ibid., p.339)

(13) C' = {<YP,X°>,<YP,ZP>,<Y°,WP>,<WP,Y°>,<X°,ZP>, <ZP,X°>, <X°,YP>,<X°,Y°>,<X°,WP>,<ZP,YP>,<ZP,Y°>,<ZP,WP>}

C' will not be totally ordered for precedence as soon as the head of ZP is taken into account - an unwelcome result.
Still, it is arbitrary to say that C_{HL} can "see" X' while it is constructing D and C but that it cannot "see" it when precedence is computed.
Note, that a dynamic solution[5] of the problem does not improve the situation. It could be argued that at the point where X' is operated on it is still of the status XP. Dynamic computation of D would give us a sequence of results:

(14) a. Merge X°,ZP D_a = {<XP,X°>,<XP,ZP>}
b. Merge Y°,WP D_b = {<XP,X°>,<XP,ZP>,<YP,Y°>,<YP,WP>}
c. Merge YP,XP D_c = { ?? }

A number of factors are involved in constructing D_c. Should we update the pairs that we carry over from D_b to read X' instead of XP or should these pairs not be carried over at all because of invisibility:

(15) D_c1 = { <X',X°>,<X',ZP> . . . } (?)

Should D_c contain <XP,XP>, <XP,X'>, or none of these? Whatever the answer is. D will either contain all or none of the pairs mentioning X'. The same reasoning that applies to the "static" computation above will then apply for the dynamic computation and no progress has been made.[6]
The invisibility of X' will have its welcome results and avoid the unwelcome ones by fiat only. I take it therefore that in that respect no advance has been made over the rival assumptions in (1). Although the scope of the proposal has been broadened by bringing to bear considerations of output conditions on a subpart of X-bar theory, that broadening doesn't appear to be felicitous.[7]

Let's turn to the principle that specifiers and complements have to be maximal projections. This follows directly from (2). Specifiers and complements being the categories that do not project when paired with another must be [+ max].

---

[5] For some intriguing empirical results see Steinbach & Vogel (this volume, footnote 5).
[6] Epstein (1995) comes closer to giving a way of deriving the intended results for the LCA. His system, however, creates problems of a different kind to do with among other things X°-movement and adjunction. I therefore skip a detailed discussion of his proposals.
[7] Although it may not directly be obvious, nothing of the above reasoning has to be changed if we adopt Chomsky's notation instead. [$_{XP}$ [$_{YP}$ Y° WP] [$_{X'}$ X° ZP]] is translatable into { X, { { Y, { Y, WP }}, { X, { X, ZP }} } }. The question would be whether or not D and C can integrate the term { X, { X, ZP }}.

For movement theory, things are more complicated. We are left with explaining why it is the case that

(16) a XP is not allowed to adjoin to X°,
b X° is not allowed to move to specifier position, and
c X° is not allowed to adjoin to XP.[8]

(16) b and c are ruled out by an additional assumption. (ibid., p.253)

(17) A chain is uniform with regard to phrase structure status

X° being of the status [- max / + min] in its base position would change into [+ max / + min] in the target position given (2) since it would not project any further there. Thus, (17) is violated and (16) b and c cannot arise. The ungrammaticality of (16)a, which does not lend itself to the same kind of treatment, is - not implausibly - attributed to morphological requirements checked on the PF-branch of C_{HL}, i.e. after spell-out. (ibid., p.319)

(18) Morphology deals only with X° categories and their features

"On this natural assumption, the largest phrases entering Morphology are X°s, and if some larger unit appears within an X°, the derivation crashes." (ibid.)[9]

---

[8] Adjunction to XP by movement might actually be ruled out by checking theory if elements not included in the maximal projection of α are not in the checking domain of α. Since movement is licensed for reasons of feature checking alone, no movement to an XP-adjoined position can take place. Chomsky (1995, p.319/326) seems to be inclined to define checking domains in the required way.

[9] Although broadening the scope of explanation potentially deepens the interest and understanding created by syntactic theory, this move might run into factual problems of the following kind. There are certain nominal constructions that appear to require a CP to occur inside N°.

(i) diese $[_{N°}\ [_{CP}$ jeder-sorgt-für-sich-selbst] Einstellung]
<this everyone-provides-for-oneself attitude>

One argument for taking CP to be inside N° is the fact that the resulting structure has exactly the distribution of N°. Thus, adjectival modifiers precede CP but cannot intervene between CP and N.

(ii) diese egoistische jeder-sorgt-für-sich-selbst Einstellung
(iii) * diese jeder-sorgt-für-sich-selbst egoistische Einstellung

Secondly, the CP induces the same stress pattern that compounding results in.

(iv) diese jeder-sorgt-für-sich-SELBST Einstellung
(v) diese EgoISTeneinstellung

The main stress - marked by capitals - falls into CP in (iv) as well as into the incorporated nominal element in (v). Standard syntactic modifiers, like adjectives and prenominal genitives, however, are less prominent than the head N°, unless a contrastive reading is intended.

(vi) diese egoistische EINstellung
(vii) Peters EINstellung

The question, of course, arises how X°-movement fares under that proposal. In fact, given (17) X°-movement should be blocked in principle. Whatever the landing site turns out to be is irrelevant since the moved X° will not project in its target position and thus - relationally - count as [+ max/ + min], in violation of (17). Clearly, condition (17) is imposed on chains at LF where C_{HL} filters out illegitimate objects, well-formed chains being the only legitimate objects. Thus, another principle is required to allow X°-movement to escape (17).

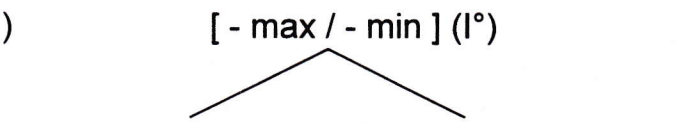
(19) At LF, X° is submitted to independent word interpretation processes WI "where WI ignores principles of C_{HL}, within X°." (ibid., p.322)

(17), thus, appears to be harmless for X°-chains. Morphology, at the same time seems to be able to distinguish an element marked [+max / + min] from a "real" maximal category, i.e. [+ max / - min]. This may be achieved by the simple principle (20).

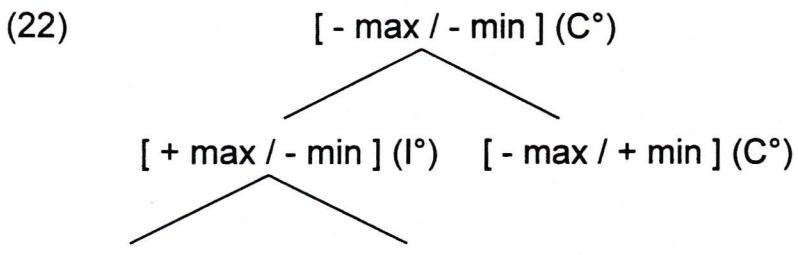
(20) [+ min] must not dominate [- min] (in morphology)

Of course, not much is known about WI to be able to assess the scope of that additional proposal. It is possible, however, to inquire into the technical execution of (18)/(19), since it is not directly evident how the components morphology and WI actually recognize X°-elements in the first place. Call this the "recognition problem". Surprisingly, it is not sufficient to appeal to (2) above. According to (2), only categories that are not projections at all should be [+ min]. Adjunction to X°, however, requires X° to *project*. (ibid., p.249/ p.260/ p.321) Consequently, the object created by adjunction relationally acquires the status [- min]. If we only go by phrase structure status, we want morphology/WI to accept configurations like (21) and (22) for e.g. V°-to-I° and V°-to-I°-to-C° respectively.

(21)



(22)



It turns out that even if one takes into account the shape of the label in adjunction structures ($<\alpha,\alpha>$, α = head) one cannot distinguish the unwanted structures from licensed ones on purely configurational grounds. (I refer readers to the appendix for an illustration of why I think this is the case). Moreover, the recognition problem is not the only complication we run into within the system set up so far. Recall that categories of the status [- max / - min] were supposed to be invisible to C_{HL}. (This - by the way - makes the recognition problem even harder to overcome.) Thus, successive-cyclic X°-movement should not be able to take place. The complex I° in (23) is relationally assigned the phrase structure status of (21).

(23) $[_{I'}\ [_{I^\circ}\ V^\circ\ I^\circ\ ]\ [_{VP}\ \ldots\ t_v\ ]\ ]$

Consequently, it is frozen in place. A further step to adjoin the complex I° to C° is not an option for C_{HL}, given the invisibility of [- max / - min]. There surely is a way out. Assuming the recognition problem can be solved, successive-cyclic X°-movement might be attributed to PF properties entirely.
For the LF branch of the computation, however, even the latter strategy is not sufficient. One of the innovations of Chomsky (1995) is allowing features to move on their own at LF via "Move-F". For subjects, objects, or verbs that do not overtly leave their base positions, feature checking is assumed to take place in I°-adjoined positions. (ibid., p.370f) (T = I° / Vb = verbal complex)

(24) $[_T$ FF(Obj.) $[_T$ FF (Subj.) $[_T$ Vb T]]]

This time, as soon as the first adjunction has taken place, the resulting structure should count as (21) in terms of phrase structure status and thus be rendered invisible. Any further adjunction to it must be blocked by C_{HL}, on a par with adjunction to X' nodes. To avoid this problem, the elements in (24) could adjoin to each other in a successive-cyclic way.

(25) $[_T\ [_{Vb}\ [_{FF(Subj.)}$ FF(Obj.) FF (Subj.)] Vb] T]

Although it is conceptually quite unattractive to assume that the formal features of the subject attract the formal features of the object, this seems to be required if both phrase structure status is computed relationally according to (2) and at the same time X'-categories are assumed to be invisible.

Indeed, Chomsky (1995) already contains a caveat concerning X° categories. Thus (2) has to be qualified for X° categories. (ibid.,p.243) The adjustment is made in the following passage (ibid.,p.245):

"To review notations, we understand a *terminal element* LI to be an item selected from the numeration, with no parts (other than features) relevant to C_{HL}. A category X^{min} is a terminal element, with no categorial parts. We restrict

the term *head* to terminal elements. An X° (zero-level) category is a head or a category formed by adjunction to the head X, which projects. The head of the projection K is H(K). If H = H(K) and K is maximal, then K = HP. We are also commonly interested in the maximal zero-level projection of the head H (say, the T head of TP with V and perhaps more adjoined). We refer to this object as $H^{\circ max}$."

The introduction of $H^{\circ max}$ is a departure from the relational concept of phrase structure status. As far as I can see (cf. appendix) $H^{\circ max}$ must be an inherent property (perhaps assigned as an (affixal) feature) of certain structural configurations. In that respect, no progress has been made over the traditional view of X-bar status as expressed in (1).
One might still argue that resorting to certain features in a single case is less costly than a full-fledged assignment of inherent X-bar status to syntactic categories, the latter disregarding the minimalist warning that

". . . with sufficiently rich formal devices (say, set theory), counterparts to any object (nodes, bars, indices, etc) can readily be constructed from features. There is no essential difference, then, between admitting new kinds of objects and allowing richer use of formal devices; we assume that these (basically equivalent) options are permitted only when forced by empirical properties of language." (ibid., p.381, fn.7)

Since the use of such concepts as ordered pairs (p.248/ p.252) (and even numerical indices (or multisets) in the case of numerations(ibid., p.227f) is allowed - forced by empirical properties of language, supposedly - it is not entirely clear to me why one cannot interpret syntactic categories as ordered pairs in the first place. One element of the pair being a numeral constructed in the classical fashion from the empty set and interpreted as the phrase structure index of the respective structure.

(26) $< X, \varnothing > = X^{\circ}$
$< X, \{ \varnothing \} > = X'$
$< X, \{ \varnothing, \{ \varnothing \} \} > = X'' = XP$

Summing up, I consider the following points to argue against a departure from systems like (1) in the way proposed in Chomsky (1995):

- The relational concept of phrase structure status as formulated in (2) cannot be fully adhered to when it comes to X°-adjunction. The system has to partially retain inherent phrase structure status in any case.
- The invisibility of X'-categories leads into contradictions when it comes to the computation of the precedence relation and it produces a "recognition problem" for the components of morphology and word interpretation.

- No invisibility argument has been advanced for other elements operated on by C_{HL} that play no role at the interfaces (e.g. uninterpretable formal features). The argument for the invisibility of X'-categories, thus, appears to be less than compelling.
- The uniformity condition on chains wrt phrase structure status (17) requires the components of morphology and word interpretation to act as filters on phrase structural outputs, which additionally produces a "recognition problem". The uniformity condition on chains would follow directly from (1).

Note, finally, that I do not object to any of the principles (2), (17), (18), and (19) if taken in isolation.

APPENDIX:

I) Take the following to be an abbreviation of [+/- max, +/- min] ("max" being the value on top, and "min" the value below)

+/-
+/-

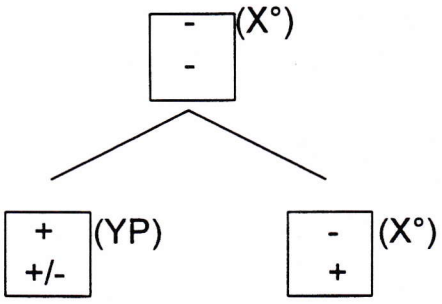
There are 16 configurations of values for the immediate dominance relation:

A	- -	B	+ -	C	- -	D	- +
	- -		- -		+ -		- -
E	- -	F	+ -	G	+ +	H	+ -
	- +		+ -		- -		- +
I	- +	J	- -	K	- +	L	+ +
	+ -		+ +		- +		+ -
M	+ -	N	+ +	O	- +	P	+ +
	+ +		- +		+ +		+ +

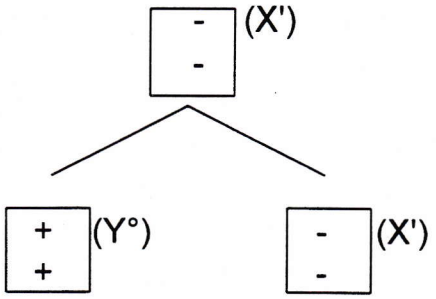
(the left pair of values immediately dominates the right pair; values standing for the terms that carry them)

The following are "illicit configurations" crucial to the "recognition problem"

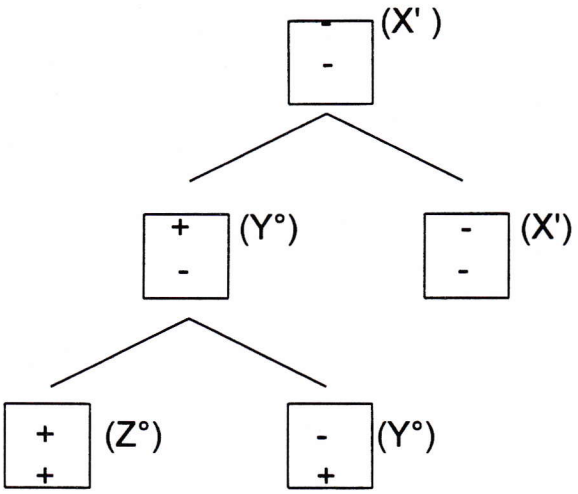
1) XP-adjunction to X°



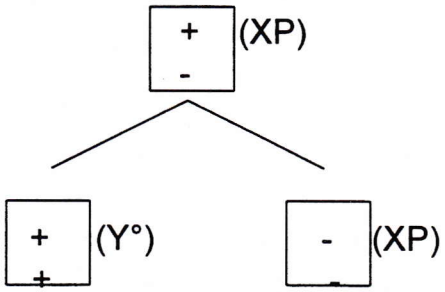
2) X°-adjunction to X'



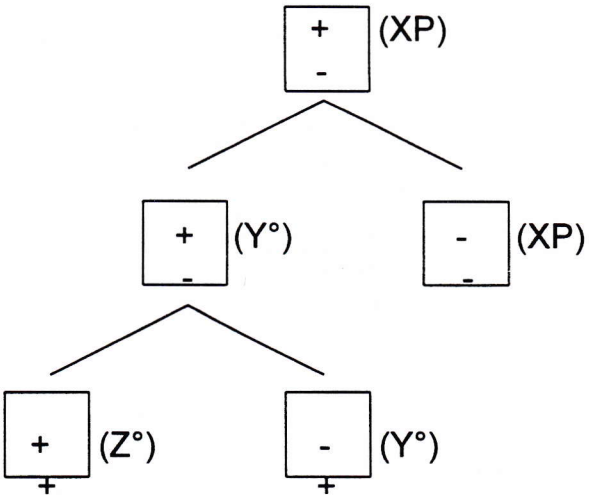
3) Complex X°- adjunction to X'



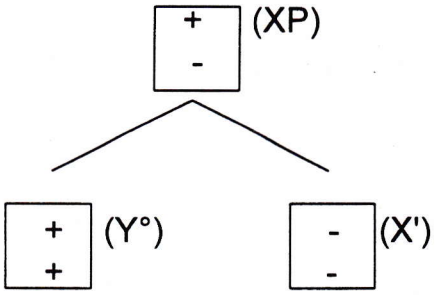
4) X°-adjunction to XP



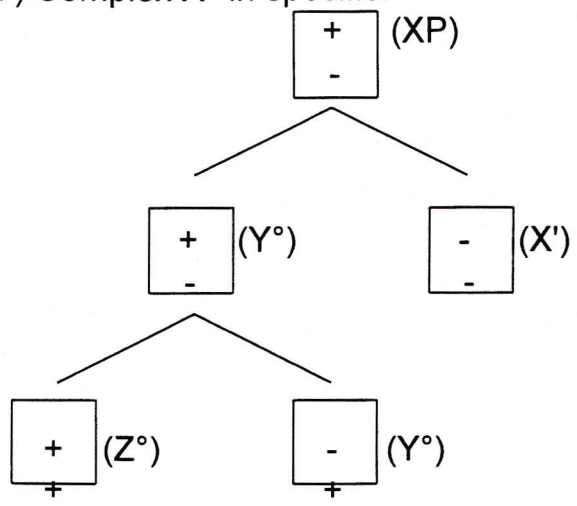
5) Complex X°-adjunction to XP



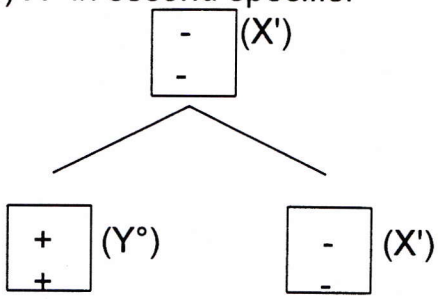
6) X° in specifier



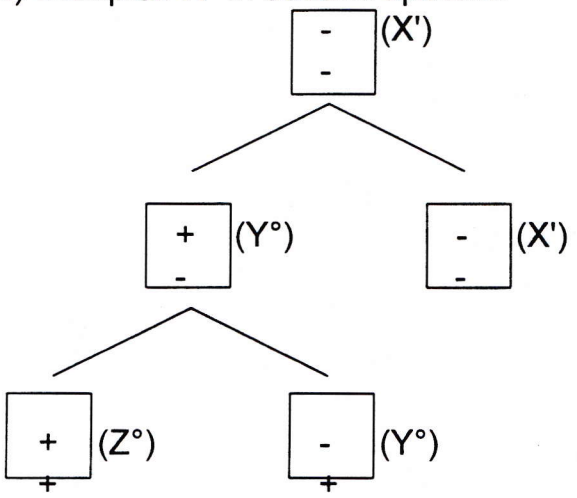
7) Complex X° in specifier



8) X° in second specifier



9) Complex X° in second specifier



The immediate dominance (ID) relations for (1)-(9) according to types A-P are the following:

ID1 = { D/K, E }
ID2 = { A, K }
ID3 = { A, D, H, N }
ID4 = { B, N }
ID5 = { B, G, H, N }
ID6 = { B, N }
ID7 = { B, G, H, N }
ID8 = { A, K }
ID9 = { A, D, H, N }

The following configurations are supposed to be licensed (among others):

10) $[_{XP}$ YP $[_{X'}$ . . .]] (XP in specifier)

11) $[_{XP}$ YP $[_{X'}$ ZP $[_{X'}$ . . .]]] (XPs in multiple specifiers)
12) $[_{XP}$ YP $[_{XP}$ . . .]] (XP adjoined to XP)
13) $[_{X'}$ YP $[_{X'}$ . . .]] (XP adjoined to X')
14) $[_{X'}$ X° YP] (X° and complement)
15) $[_{XP}$ X° YP] (X° and complement)
16) $[_{X'}$ $[_{X°}$ Y°X°] YP] (complex X° and complement)
17) $[_{X°}$ Y°X°] (X° adjunction)
18) $[_{X°}$ $[_{Y°}$ Z° Y°] X°] (complex X° adjunction)
19) $[_{X°}$ Z° $[_{X°}$ Y° X°]] (multiple X° adjunction)

The resulting ID-type sets are:

ID10 = { B, G/N }
ID11 = { A, B, G/N, D/K }
ID12 = { B, G/N }
ID13 = { A, D/K }
ID14 = { D/K, E }
ID15 = { H, G/N }
ID16 = { A, E, D/K, K }
ID17 = { E, K }
ID18 = { D, E, H, N }
ID19 = { A, E, K }

Result:

- ID1 cannot be ruled out because it is matched by ID14.
- ID2/ID8 cannot be ruled out because they are overlapped by ID11/ID13/ID16/ID19.
-ID4/ID6 cannot be ruled out because they are matched by ID10/ID12 and overlapped by ID11

- ID3/ID5/ID7/ID9 cannot be ruled out because they are overlapped by an extended ID11': $[_{XP}$ YP $[_{X'}$ $[_{ZP}$ Z° WP] $[_{X'}$ . . .]]]
= { A, B, D, G/N, (G/N), H }

- (ID18 can be extended to yield ID18': $[_{X'}$ $[_{X°}$ $[_{Y°}$ Z° Y°] X°] WP]
= { A, D, D/K, E, H, N }
to overlap and rule out ID3/ID9)
- etc.

Recoding these results in terms of sisterhood or c-command-relations does not alter the picture.

The illicit structures cannot be distinguished from licensed configurations.

II) It might alternatively be attempted to distinguish the unwanted from the licensed structures by identifying their configuration of labels. Again, the structures in (1)-(9) find their match among (10)-(19), since the only distinction are straight (X) labels and labels resulting from adjunction (<X,X>).

1/2/4) $[_{<X,X>}\ Y\ X]$
3/5) $[_{<X,X>}\ [_{<Y,Y>}\ Z\ Y]\ X]$
6/8) $[_{X}\ Y\ X]$
(7/9) $[_{X}\ [_{<Y,Y>}\ Z\ Y]\ X]$

12/13/17) $[_{<X,X>}\ Y\ X]$
18) $[_{<X,X>}\ [_{<Y,Y>}\ Z\ Y]\ X]$
10) $[_{X}\ Y\ X]$
extended 10'
$= [_{XP}\ [_{YP}\ ZP\ [_{YP}\ \ldots]]\ [_{X'}\ \ldots]]$
$= [_{X}\ [_{<Y,Y>}\ Z\ Y]\ X]$

The illicit structures cannot be distinguished from licensed configurations.

References:

CHOMSKY, Noam (1986): Barriers. Cambridge MA.
CHOMSKY, Noam (1995): The Minimalist Program. Cambridge MA.
EPSTEIN, Samuel D. (1995): "Un-Principled Syntax and the Derivation of Syntactic Relations", ms. Harvard.
FRANK, Robert & K. VIJAY-SHANKER (1995): "Foundations of Phrase Markers", talk presented in Berlin.
GAERTNER, Hans-Martin & Markus STEINBACH (1994): "Economy, Verb Second, and the SVO-SOV Distinction"; in: Working Papers in Scandinavian Syntax Vol.53, 1-59.
KAYNE, Richard S. (1994): The Antisymmetry of Syntax. Cambridge MA.
MUYSKEN, Pieter (1982): "Parametrizing the notion "Head""; in: Journal of Linguistic Research 2, 57-75.
PARTEE, Barbara et.al. (1993): Mathematical Methods in Linguistics. Dordrecht.
ZWART, Jan-Wouter (1994): "The Minimalist Program and Germanic Syntax. A Reply to Gaertner and Steinbach"; in: Working Papers in Scandinavian Syntax Vol.54, 1-41.

Hans-Martin Gaertner
Max-Planck Gesellschaft
Arbeitsgruppe Strukturelle Grammatik
Jaegerstrasse 10/11
10117 Berlin
e-mail: gaertner@asg.ag-berlin.mpg.de

---

*CONSTRAINTS ON FREE RELATIVE CLAUSES IN TURKISH**

Jaklin Kornfilt
FAS-Berlin and Syracuse University

O. Introduction:

This paper investigates aspects of Turkish relative constructions that have attracted little, or no, attention previously: 1. the order of morphemes--in particular, of agreement morphemes of the participle and of the plural morpheme of the (understood) head--in the verbal complex of Free Relative Clauses (FRCs) with regular factive participial morphology; 2. the lack of so-called "matching effects" in Turkish HRCs; 3. lack of genuine infinitival relative clauses (both headed and non-headed) and the existence of (headed) "future tense" or "irrealis" relative constructions instead; 4. lack of FRCs in the irrealis.

I shall suggest that the the explanation for the observations in 1, 2 and 4 has to do with a constraint on empty operators originally proposed in Levin (1983): Empty operators cannot be governed. If on the right track, this study provides an additional piece of evidence in favor of head government--a notion abandoned in recent work (e.g. Chomsky 1993) but very recently defended again (cf. Rizzi 1995). This paper also attempts to provide evidence for the existence of syntactic rules that are motivated by the necessity to overtly mark the scope for certain rules taking place at LF.

I now turn to individual discussions of the four observations listed above.

1. "Reordering" between participial agreement and inherent plural:

Note the order of the morphemes between the inherent plurality marker **-lAr** and the agreement marker in (1)a.:

(1)a. [Geçen yaz ada-da gör-dük-**ler**-**im**] bu yaz gel-me-di (-ler)
last summer island-Loc. see-Partic.-pl.-1.sg. this summer come-Neg.-past(-3.pl.)
'(Those) who(m) I saw on the island last summer didn't come this summer'

This order is unexpected. The expected order would be for agreement to precede plurality, given the corresponding relative construction with an overt head:

---

* I would like to thank Artemis Alexiadou and Paul Law for enjoyable discussions of some of the material studied in this paper. I should further like to thank Ewald Lang and Renate Steinitz for the pleasant and fruitful time I spent at FAS-Berlin in 1995, during which time this paper was written. Special thanks are due to the members of the syntax project, i.e. Artemis Alexiadou, André Meinunger, Paul Law, and Chris Wilder, for welcoming me in their group as a fellow syntactician.

(1)b.[[Geçen yaz ada-da gör-düğ-üm] kişi-ler]
last summer island-Loc. see-Partic.-1.sg. person-pl.
'The people who(m) I saw on the island last summer'

The agreement is on the embedded verb and expresses person and number features of the embedded subject. The plurality morpheme, on the other hand, expresses the inherent plurality feature of the head. Thus, in a FRC without an overt head noun, if the plural feature of the head is to be expressed at all, it should show up after the agreement morpheme and not before.

Why, then, is the original order between agreement and plural morphemes not preserved, i.e. why is the shape of the participle in (1)a. not as in (1)c.?

(1)c. *gör-düğ-üm-ler?

I will return to attempting an answer in the last section of this paper, and I now turn to an illustration of the second observation made in the introduction.

2. No matching effects in Turkish FRCs:
In Turkish relative clauses, there is neither an overt complementizer, nor an overt relative pronoun. Instead, the verb of the modifying clause bears participial morphology (in the examples above, the general factive morpheme -DIK), and there is a gap in the position of the constituent in the modifying clause that corresponds to the head. Now note the following example:

(2)[sevgili-m-in e sev-me-dik -ler-in]-den hediye al -ma-m
lover -1.sg.-Gen. love-Neg.-Partic.-pl.-3.sg.-Abl. present take-Neg.-Aor.+1.sg.
'I don't take presents from (those) who(m) my lover doesn't like'

A clash might be expected between the Accusative empty category in the modifier in (2) and the Ablative in the matrix--especially, if we assume (as would be quite reasonable) that there is an abstract operator instead of an overt wh-element which moves to Spec/CP. In this particular example, the operator would be marked Accusative; the whole NP (i.e. the FRC) would get Ablative case assigned in the matrix clause. No clash seems to arise, however, and this and similar examples are fine.

3. "Infinitival" Relative Clauses:

Turkish does not have infinitival relative constructions; this is illustrated by the ungrammatical (3):

(3) *Ahmet Ayşe$_j$ - ye [$_{NP}$ [$_{CP}$ PRO$_j$ e$_i$ oku - mak] bir kitap$_i$] al-dı
-Dat. read-Inf. a book buy-past
Intended reading: 'Ahmet bought Ayşe a book to read'

This is a surprising fact, given that infinitival complements with Control semantics are very productive in Turkish; e.g.:

(4) Ahmet$_i$ [PRO$_i$ bir kitap oku - mak] isti - yor
a book read - Inf. want - Pres. Progr.
'Ahmet wants to read a book'

Example (4) illustrates subject control; object control by accusative and dative objects is illustrated by the next two examples:

(5) Ahmet Ayşe-yi$_i$ [PRO$_i$ sinema-ya git-meğ]-e zorla - dı
-Acc. cinema-Dat. go-Inf. -Dat. force-past
'Ahmet forced Ayşe to go to the movies'

(6) Ahmet Ayşe-ye$_i$ [PRO$_i$ sinema-ya git-meğ]-i tavsiye et-ti
-Dat. cinema-Dat. go-Inf.-Acc. recommend -past
'Ahmet recommended to Ayşe to go to the movies'

Given these facts, the question becomes particularly intriguing why infinitival clauses in Turkish, which do seem to parallel their English counterparts so closely, allow neither type of WH-construction to penetrate them.

On the way towards an account of this mystery, we have to note that Turkish does have a relative construction which is equivalent to English infinitival relatives (cf. translation of (3)) semantically and which is formed by replacing the infinitive morphology by the morpheme for Future, which, in this case, has the semantics of irrealis/potentiality:

(7) Ahmet Ayşe$_j$ - ye [$_{NP}$ [$_{CP}$ PRO$_j$ e$_i$ oku - yacak] bir kitap$_i$] al-dı
-Dat. read-Fut. a book buy-past
'Ahmet bought Ayşe a book to read'

These "potentiality" relative clauses have the following properties: their modifying clause never has an overt subject, and the verb of that clause is never inflected for agreement. I claim in this paper that examples like (3) are ungrammatical due to Rizzi's WH-Criterion: the infinitive affix **-mAK** is marked as [-WH], as is the non-factive affix **-mA** of which the infinitive is a sub-class. The future affix, however, just like the factive **-DIK**, is [+WH].

These are not just idiosyncratic, arbitrary assignments of feature values: The factive markers (which show up in the morphological slot reserved for the tense morpheme(s) in fully finite verbs) do have tense connotations, albeit to an

impoverished degree. Complement clauses marked with one of these factive morphemes are, from the point of view of tense, independent from the matrix:

(8)a. [viski iç - tiğ - iniz]-i bil - iyor - uz
- DIK - 2.pl. -Acc. know -Pres.Progr.-1.pl.
'We know that you drink/drank whisky'

b. [viski iç - eceğ - iniz]-i bil - iyor - uz
- AcAK -
'We know that you will drink whisky'

Note that (8) a. and b. are identical, up to the nominalization marker on the subordinate ("nominalized") clause. In particular, the tense/aspect marking of the matrix clause is the same. Hence, it is clear that the different tense/aspect interpretation of the subordinate clauses is due to the different "nominalizers", i.e. **-DIK** and **-AcAK**. This type of complement is often referred to in the literature as "Factive Nominal" (cf. Lees 1963, Underhill 1976), and "Personal Participle" (Lewis 1967). While the rich array of tense and aspect of Turkish root clauses is not found in these complements, there is nevertheless a remainder of tense, as seen by comparing (8) a. and b.: we can describe this subsystem as having the (impoverished) tense array of future (-AcAK) and non-future (-DIK). It should be further pointed out that, despite the difference in shape, such examples do, indeed, form one type, since they are selected by the same matrix verbs (essentially, factive verbs), and they have, indeed, factive semantics themselves, as can be seen by their translations.

The independence of the tense of the complement from that of the matrix, which is typical for the "factive", is not true for the second complement type, illustrated in (9), where clearly the non-factive complement is dependent of the matrix where tense is concerned:

(9) a. [viski iç - me - niz] - e karşı - y - ız
- mA - 2.pl.-Dat. against-Cop.+Aorist-1.pl.
'We are against your drinking whisky'

b. [viski iç - me - niz] - e karşı - ydı - k
- mA - 2.pl.-Dat. against-Cop.+past-1.pl.
'We were against your drinking whisky'

c. [viski iç - me - niz] - e karşı ol-acağ - ız
- mA - 2.pl.-Dat. against be-Fut. -1.pl.
'We will be against your drinking whisky'

The same is true of infinitival complements:

(10) a. pro_i [PRO_i viski iç - mek] isti - yor - uz
drink- Infin. want-Progr.-1.pl.

'We want to drink whisky'

(10) b. pro_i [PRO_i viski iç - mek] iste - di - k
drink- Infin. want-Past-1.pl.
'We wanted to drink whisky'

c. pro_i [PRO_i viski iç - mek] isti - yeceğ - iz
drink- Infin. want -Fut. -1.pl.
'We will want to drink whisky'

The fact that, just like the non-factive complements, infinitival complements are also dependent on the matrix clause from the point of view of tense is not surprising, given that the verbs selecting for infinitival complements are a subset of those selecting non-factive complements. As a matter of fact, the non-factive morpheme **-mA** and the infinitival morpheme **-mAK** have an obvious formal resemblance. Some researchers (e.g. Kural 1993) have called the non-factive morpheme an inflected infinitive. There is some justification to that, given that the non-factive morpheme is, indeed, inflected for agreement with its subject, while the regular infinitive cannot be, since its subject is PRO. This would follow under the PRO-Theorem, if we assume that the subject is governed by AGR. On the other hand, it is not the full infinitival morpheme which is inflected, but the form without the **-K**. Hence, it makes better sense to view the infinitive as an agreement-less non-factive gerundive, and one which lacks, as all such gerundives in Turkish do, tense and aspect.

In what follows, I argue that the inflectional element raises to Comp at LF and offer arguments in favor. To this end, I discuss first WH-questions, and then relative clauses.

3. 1. WH-Questions in Turkish:

Wh-elements in Turkish are, essentially, in situ. In other words, there is no evidence of a syntactic movement to a clause-peripheral position--say, to Spec/CP. While the preferred position for these elements is immediately pre-verbal (like in Hungarian, cf. Horvath 1986), this is not obligatory (cf. Bechhofer 1975). Therefore, differences in the scope of wh-words, which are expressed in terms of surface order in a language like English with overt syntactic wh-movement, must be expressed differently in Turkish. This is done by intonational differences. In the following discussion, I shall be mainly concerned with narrow scope wh-questions (i.e. embedded wh-questions), but I shall also address their wide scope counterparts, i.e. matrix questions whereby a wh-element is "extracted" (but not overtly, only with respect to scopal semantics) out of a complement clause.

It appears that wh-constituents in both Factive and Action Nominals can have wide scope:

Factive Nominal:

(11) [Parti-ye kim -in gel - diğ - in] - i duy - du - n?
party-Dat. who-Gen. come-DIK-3.sg.-Acc. hear -Past-2.sg.
'Who did you hear came to the party?'

Action Nominal:

(12) [Parti-ye kim -in gel - me - sin] -e kız - dı - n?
-mA -3.sg.-Dat. angry-Past-2.sg.
'Who were you angry that came to the party (i.e. about whose coming to the party were you angry)?'

However, while some -DIK complements allow for narrow-scope questions, -mA complements never do for some speakers. While some other speakers are more permissive in this regard, only very few matrix verbs that select for -mA complements are allowed to take embedded wh-questions even by these more permissive speakers. In other words, embedded questions can always be of the -DIK-type, but they are heavily restricted at best when they are of the -mA-type:

(13) [Parti-ye kim -in gel-diğ-in] -i bil-iyor-um (sor-; duy-; etc.
party-Dat. who -Gen. come-DIK-3.sg.-Acc. know-Pr.-1.sg. (ask; hear...)
'I know (asked; heard ...) who came to the party'

(14) *[Parti-ye kim-in davet ed-il-me-sin]-i tembih et -ti-m
party-Dat. who -Gen. invite -Pass.-mA-3.sg.-Acc. insistently tell-past-1.sg.
'I insistently/urgently said who was to be/should be/for whom to be invited to the party'

A corresponding example with a similar, but more widely used verb is better:

(15) ?(?) [Parti-ye kim-in davet ed-il-me-sin]-i söyle - di - m
party-Dat. who -Gen. invite -Pass.-mA-3.sg.-Acc. say -past-1.sg.
'I said who was to be/should be/for whom to be invited to the party'

The contrast between Factive and Action Nominals with respect to embedded wh-questions can be seen in a particularly clear fashion with some of those matrix verbs that take either -DIK- or -mA-complements:

(16)a. [Parti-ye Ahmed-in gel - diğ - in] -i söyle - di - m
party-Dat. -Gen. come-DIK-3.sg.-Acc. say-Past-1.sg.
'I said/told that Ahmet came to the party'

b. [Parti-ye Ahmed-in gel - me - sin] -i söyle - di - m
- mA -

'I said that Ahmet should come to the party (for A. to come to the party)'

Now, while the -DIK-complement in (16)a. can always host a narrow-scope WH-element with the same ease for all speakers, the -mA-complement in (15) and in (16)b. is not accepted by all speakers:

(17)a. [Parti-ye kim - in gel - diğ - in] - i söyle-di-m
who - Gen. - DIK -
'I said/told who came to the party'

b. ?(?)[Parti-ye kim - in gel - me - sin] - i söyle-di-m
who - Gen. - mA -
'I said who should come/for whom to come to the party'

As mentioned before, infinitivals occur with a subset of those matrix verbs that select -mA complements. Even those speakers who are otherwise rather permissive with respect to narrow-scope wh-questions in -mA complements under widely used matrix verbs like söyle 'say' don't allow for infinitival narrow-scope wh-questions. I give some further examples for the sake of convenience, using one of the matrix verbs which were introduced earlier:

(18) a. pro_i [PRO_i doktor -a git - meğ] - e karar ver - di - m
physician-Dat. go -Inf. - Dat. decision give-past-1.sg.
'I decided to go to the doctor'

b. *pro_i [PRO_i kim -e git - meğ] - e karar ver - di - m
who -Dat. go -Inf. - Dat. decision give-past-1.sg.
'I decided to whom to go'

3. 2. Relative Clauses:

Turkish relative clauses are head-final, as are all phrases. The modifier clause is headed by a "nominalized" predicate--indeed, our familiar **-DIK** form, i.e. what I have called "Factive Nominal" earlier in this paper:

Non-future:
(19) a. [Hasan - ın iç - tiğ - i] viski
-Gen. drink-DIK-3.sg. whisky
'The whisky that Hasan drinks/drank'

Future:
(19) b. [Hasan - ın iç - eceğ - i] viski
-Gen. drink-AcAG-3.sg. whisky
'The whisky that Hasan will drink'

Note that not only is the verbal morphology the same as that of our Factive Nominals, we also have the same division into future/non-future forms, and the Genitive marking on the subject.

While the morphology is different for those instances where a subject or part of a subject is "relativized", those intricacies are not relevant for our purposes here. What is relevant and interesting, however, is the fact that neither the "Action Nominal" with **-mA**, nor (as we saw earlier in the paper) the infinitive with **-mAK** are ever part of the verbal morphology that heads the modifier clause of a relative clause:

(20) *[PRO çal - mak] bir sonat
play-inf. a sonata
Intended reading: 'A sonata to play'

(21) *[Cem-in çal - ma - sı] bir sonat
-Gen. play-mA -3.sg. a sonata
Intended reading: 'A sonata for Cem to play/which Cem should play'

The corresponding constructions with the **-DIK** morphology (and the factive semantics that go along with it) are perfect:

(22) a. [pro çal - dığ - ım] bir sonat
play -DIK-1.sg. a sonata
'A sonata which I play/played'

(23) a. [Cem-in çal - dığ - ı] bir sonat
-Gen. play -DIK-3.sg. a sonata
'A sonata which Cem plays/played'

Both of these examples are fine with the Future version of the Factive Nominal, as well:

(22) b. [pro çal - acağ - ım] bir sonat
play -Fut.-1.sg. a sonata
'A sonata which I will play'

(23) b. [Cem-in çal - acağ - ı] bir sonat
-Gen. play -Fut-3.sg. a sonata
'A sonata which Cem will play'

Note that both the Action Nominal and the infinitival, although they cannot head the modifier clause in a relative clause construction, can both be found on intermediate predicates:

(24) [pro$_i$[PRO$_i$[PRO$_i$ çal - mağ]- a başla - mak] iste - diğ - im] bir sonat
play -Inf. -Dat. begin -Inf. want-DIK -1.sg. a sonata
'A sonata which I want/wanted to begin to play'

(25) [[[Cem - in çal - mağ]- a başla - ma - sın]-ı iste - diğ - im] bir sonat
-Gen. play-Inf. -Dat. begin-mA -3.sg.-Acc. want-DIK -1.sg. a sonata
'A sonata which I want/wanted that Cem should begin to play'

3. 3. Towards an Explanation:

In attempting to account for the lack of infinitival (as well as "subjunctive", i.e. non-factive) embedded wh-questions and relative clauses in Turkish, I would like to explore a proposal by Rizzi (1991), where the following principle is proposed:

(26) The Wh-Criterion

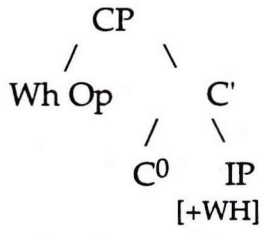
A. A Wh-Operator must be in a Spec-head configuration with an X^0 .
[+WH]

B. An X^0 must be in a Spec-head configuration with a Wh-operator.
[+WH]

I shall first discuss how the Wh-Criterion might provide an explanation for the facts we discussed concerning wh-questions before turning to relative clauses. Also, I shall first disregard the permissive dialect which freely accepts narrow-scope wh-questions with non-factive complements but shall return to that dialect later on.

In an attempt to make the Wh-Criterion more intuitive, Rizzi offers the following explanation:

"As the feature +WH on a clausal head (most typically a C^0) designates the fact that the projection of that head (CP) is a question, the Wh-Criterion simply expresses the fact that at the appropriate level of representation interrogative operators must be in the spec of CPs which are interpreted as questions and, reciprocally, CPs interpreted as questions must have interrogative operators as specifiers. The Wh-Criterion thus requires configurations of the following shape:

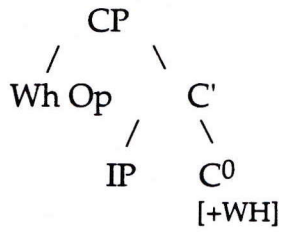


As a general well formedness principle on the scope of wh-operators, [the Wh-Criterion] can be taken as a criterial condition applying universally at LF. So, in languages lacking syntactic wh-movement, such as Chinese and Japanese, question

operators must be moved in the syntax of LF to satisfy the Wh-Criterion at this level, ..." (Rizzi 1991, p. 24)

I would like to claim that Turkish, as a language where, at least for wh-questions, LF-movement has to be posited, applies the Wh-Criterion at that level--i.e. at LF; the lack of infinitival (as well as non-factive) narrow-scope wh-questions follows as a consequence. The LF-configuration in which the Wh-Criterion would apply in Turkish would be as follows:

(27)



I do not take the fact that Spec/CP and the head of CP are at opposite peripheries of the CP to be a problem. There is nothing about the Wh-Criterion (or other principles of grammar, for that matter) which would render it (or them) inapplicable in such a configuration.

There might not be very much overt evidence in favor of this particular configuration, as opposed to one where C^0 and Spec/CP would be on the same side of CP, given that the language has no overt complementizers (although this is open to debate, as we shall see shortly), and also given that the wh-movement I am assuming here is syntactically abstract. As a matter of fact, such an alternative configuration can be assumed, as well, as far as the purposes of this paper are concerned. My main reason for positing the configuration in (27) is my assumption that the directionality between the Spec and the head of a phrase should be the same as the directionality between the head and its complement[1], and the latter configuration is doubtlessly head-final.

My specific proposal is quite simple at this point: The C^0 of non-factive complements and, as a special subset, of infinitival complements is [-WH]. Thus, wh-movement at LF to Spec/CP in non-factives (and infinitivals) gives rise to a violation of the Wh-Criterion. On the other hand, the C^0 of factive complements is [+WH], and thus the same type of movement will lead to a felicitous result--and, indeed, will be obligatory.

This particular interpretation does need some further motivation and justification, however, since the C^0 I am assuming is empty for all the "nominalized" complement types under discussion in this paper, and we shall therefore need some means to differentiate between [+WH] and [-WH] C^0s in a motivated way.

---

[1]For a similar view, see Georgopoulos 1991.

Before turning to such motivation, let me first mention--and then dismiss--another logical possibility that comes to mind. Suppose we said that while Factive complements are, indeed, CPs, non-factive and infinitival complements are not. The latter claim would be in line with traditional views which treat **-mA** and **-mAK** complements as verbal nouns and thus not fully clausal.

First of all, it was mentioned earlier that the verbal predicates of these complements do not lose their transitivity (if the verbs in question are, indeed, transitive); thus, they are not really verbal nouns, strictly speaking, and the complements in question can have complex clausal characteristics: a full array of verbal arguments, passive, causative, negation etc. In other words, the internal structure of these complements is, indeed, clausal and not different syntactically from that of factive complements.

Secondly, assume that non-factive and infinitival complements were not CPs. Depending on what we take the lower maximal projection under CP to be, such complements would be IPs of some kind--AgrPs or T(ense)Ps. Given that there is no tense in these complements, and that infinitivals don't have any agreement, such (a) lower projection(s) would obviously be of a defective type. We should therefore expect that such complements would be easy to penetrate from the outside--in particular, they should be transparent to government by the matrix verb whose complement they are.

This would mean a governed PRO-subject of infinitivals, leading to a violation of the PRO-Theorem, and subjects of **-mA** complements that bear the Case assigned by the matrix verb. However, we have seen earlier that there is no reason to assume that the PRO-Theorem can be successfully violated in Turkish, and we have also seen that the subjects of **-mA** complements are marked Genitive within their clause and never bear the Case assigned by the matrix verb (the latter being assigned to the complement as a whole).

Thus, we conclude that non-factive and infinitival complements are CPs, just as their factive counterparts.[2] If so, we do have to posit a C^0-head for them.

In order to differentiate between the C^0-head of factive complements on the one hand and the head of "action" and infinitival complements, on the other, I shall adopt a suggestion made in Rizzi (1991), namely that the basic locus of the [+WH] feature can be, in some languages, Tense. Rizzi hypothesizes that, while [+WH] features are "scattered" in the clausal structure, they can "gravitate", metaphorically speaking, to Tense in those languages (or structures) where that element is, indeed, a rich "gravitational center". In a language (or structure) where there is no Tense, or where that element is weak, that graviation will not take place.

---

[2]For a similar conclusion about German infinitivals, based on careful argumentation, see Sabel (1993).

Once we make this assumption, we have to somehow transmit the [+WH] features to the C^0-head of the clause. There are a variety of ways to implement this. We could move the verb to Tense, Agr, and then to C^0, if we are working within a system where we build morphologically complex words in the syntax, or we could percolate the relevant features up.

Note, incidentally, that whatever mechanism we choose, we will need to use it for more purposes than just for applying the Wh-Criterion. Given that matrix verbs select for certain complement types and not others, we must make sure that those verbs have access to Tense (and perhaps Aspect, Modality etc.) features in the complement clause.

This problem (if it is one) might be circumvented, if we said that the locale of these features, i.e. the "nominalization" markers, are placed in C^0. In and of itself, this might be problematic, since we want to place these markers into the Tense position, to capture the fact that at least for factive complements, this is indeed where tense differentiations are made, and also to capture the fact that these markers show up in the same slot within the verbal complex where full-fledged tense markers occur in fully finite clauses.

An intriguing idea is advanced in Kural (1993), where it is proposed that the C^0 position in Turkish complements is filled by **-k**. In other words, while the remainder of the markers are, indeed, in Tense, their final **-k** is really part of a different category, namely of C^0. For **-mA** complements (which he claims are simply inflected infinitives rather than a distinct complement type), which have no **-k**, Kural assumes that they are not CPs. He still assumes, as I have done, that the Genitive marking on the subjects of such complements is assigned by Infl (or Tense), but he follows Raposo (1987) in claiming that an infinitival Infl (i.e. Tense or Agr) cannot assign Case to its specifier unless it is Case-marked by the higher verb.

Obviously, this is against Stowell's (1981) Case Resistance Principle. There are further problems with this proposal (one of which was mentioned before, namely that the putative **-dI**, which would be the simple past in an embedded context, denotes both past and present in complement clauses, but is limited to the past in root contexts, thus casting doubt on the claim that it is one and the same morpheme), the most serious one being the order of morphemes: Agr would have to be outside of IP, since it follows the putative C^0. While Kural does recognize this problem, and devises ways to deal with it, the proposal remains problematic, especially with respect to the various relationships between Agr and the subject within the complement--e.g. Case assignment, the possibility of PRO in infinitives but not elsewhere, and, most seriously, the role of Agr in the licensing and identification of pro-subjects (since, being outside of CP, Agr would be too far removed from the subject of the embedded IP to qualify as a local identifier). If Agr is analyzed as the head of IP, all these problems disappear, and syntax as well as the morphology become straightforward.

I shall therefore retain my analysis with empty C^0s and with either projection of Tense (and the corresponding wh-) features to that C^0, or else with V-to Tense-to Agr^0-to C^0-movement; I shall not take a stand between these alternatives in this paper. Once the C^0 has the appropriate features, the explanation for the facts we have encountered follow: If C^0 has received [+WH] features, wh-movement to Spec/CP will be possible (and necessary), since it is both allowed and enforced by the Wh-Criterion; if C^0 has received [-WH] features, such movement will be ruled out by the same criterion.

The same account will also explain the fact that matrix wh-questions are always possible--both when limited to simple questions and when applying, in a complex construction, to move an embedded wh-constituent at LF, giving it wide scope. The account does need some modification, however.

This modification will rely on Rizzi's notion of "dynamic agreement". The problem is as follows: In order to land in Spec/CP of a matrix clause, a wh element will need a C^0 head with [+WH] features; however, since a matrix CP is never selected by another verb, the head of such a CP will not have "inherent' [+WH] features. How can the head of a matrix CP acquire the relevant features?

Rizzi advances a notion of "dynamic agreement", in order to account for some facts in French, whose nature do not concern us for our purposes. He proposes that some languages have available an extra option of an agreement process between a wh-operator and a licensing head:

(28) Wh-Op X^0--> Wh-Op X^0
[+WH]

This "dynamic" agreement between the element in Spec and the head in terms of features is not the same as the "static" agreement between the same elements that we had seen earlier. Static agreement for the purposes of the Wh-Criterion obtains always; there, each element has features inherently, and independently from each other; those features must agree in a given configuration. In a situation of dynamic agreement, on the other hand, the specifier is able to transmit its own features to the head. This type of agreement is limited to certain languages and certain syntactic contexts.

We have to make sure, however, that we do not run into problems. In particular, we have to block dynamic agreement from applying in embedded contexts where the head of CP lacks [+WH] features, because otherwise we would undo the beneficial, explanatory effects of the WH-Criterion.

What we shall say is that complement clauses lacking such features are specified for [-WH] features; the morphological markers **-mA** and **-mAK** express just that specification. In other words, it is not the case that such complements simply lack

[+WH] specification; rather, they are actually specified, namely for [-WH]. Such inherent specification cannot be overridden by dynamic agreement. Matrix CPs, on the other hand, are not selected for either [+WH] or [-WH], and they simply lack such inherent specification. However, having rich Tense, Aspect, and Modality, they have the potential for receiving the positive value of the [WH] feature. Therefore, in a configuration where dynamic agreement can apply, the C^0 head of such CPs can receive the [+WH] specification from a [+WH] element in Spec/CP. In this fashion, we account for both simple and complex matrix questions, without endangering the explanation we had achieved with respect to embedded complements.

There is one apparent problem that emerges with respect to matrix wh-questions involving a wh-element in a [-WH] complement. How does such an element manage to escape its own clause? Is LF-movement not restricted by Subjacency?

There are various ways to address this issue: We might say, along with Huang (1982), that LF-movement does indeed not obey Subjacency. However, in recent years more evidence has emerged to show that this view is probably not correct. Alternatively, we might say that intermediate traces are not operators (cf. Kornfilt 1984) and are therefore not affected by the Wh-Criterion. Lastly, we could say that the LF-moved embedded wh element actually does not leave the Spec of its own clause, but rather induces Pied Piping of the whole complement clause to Spec/CP of the matrix clause [3]. I shall leave the decision between these last two alternatives to future research. At any rate, we see that there is no real problem in this regard.[4]

I now turn to a brief discussion of the permissive dialect, i.e. the dialect which allows for narrow-scope wh-questions in non-factive complements.

---

[3]Such an approach based on Pied Piping is adopted by Nishigauchi (1990) for Japanese, Ortiz de Urbina (1992) for Basque, and Özsoy (1991) for Turkish.

[4]There seems to be a typological difference between Turkish and some other languages--e.g. Romance and Basque--in this respect. In Picallo (1984), it is reported that in Catalan, QPs embedded within subjunctive complements (which seem to correspond to the Turkish non-factive complements in general) cannot take wide scope over a matrix quantifier, while QPs embedded within indicative clauses (roughly corresponding to the Turkish factive complements) can. The difference is related, according to Picallo, to ECP-effects: the INFL-node of a subjunctive clause, lacking Tense, cannot act as a proper governor, while the corresponding node of an indicative clause, possessing Tense, can do so. Ortiz de Urbina (1992) reports that in Basque, certain complement clauses overtly marked for [-WH], and thus not allowing for narrow-scope wh-questions, also do not allow for wide-scope questions. This interesting typological difference between Romance and Basque, on the one hand, and Turkish, on the other, will have to be addressed in future work.

If we look at the examples for such questions and their attempted (and, depending on speaker, successful) readings in their English translations, we see that there is an aspectual or modal meaning attached to them. As a matter of fact, there is a root modal suffix, which is used to express the same meaning, and which is overtly related to the **-ma/-mAK** suffixes of the non-factive and infinitival complements: **-mAlI**. Thus, we get examples like the following:

(29) Bugün doktor - a git - meli - yim
today physician-Dat. go-Necess.-1.sg.
'I have to/I must/I am to go to the doctor today'

(30) Bugün kim - e git - meli - yim?
today who -Dat. go-Necess.-1.sg.
'Who do I have to/mustI/am I to go to today?'

This modal cannot show up in nominalized complements. Factive complements, however, by virtue of having Tense, retain some aspectual properties nevertheless, and thus can "summon" the wh-features in the clause and transmit them to the head of CP, as we saw earlier. However, non-factive and infinitival complements, devoid of both Tense and Aspect/Modality, lack a "gravitational center" to attract and transmit wh-features.

Suppose, then, that we have, in addition to--and higher than--Tense, also a Mod(ality) or Asp(ect) node, which is empty in **-mA** complements. Since Tense is empty, as well, no [+WH] features are attracted, and consequently there is nothing to transmit to the head of CP. However, for some speakers, **-mA** does have similar modality features as the corresponding root **-mAlI**. For such speakers, the Mod-node has features, although there is no distinct marker for them. Such a feature-filled Mod-node acts as Rizzi's gravitational center with respect to [+WH]-features, which end up in the CP-head, thus enabling LF-movement of the WH-element to Spec/CP.

Remember that even for such permissive speakers, however, it is impossible to have narrow-scope wh-questions in infinitivals, i.e. complements marked with **-mAK**. I suggest that the **-k** occupies the Mod (or the higher Agr) node, thus making it impossible for the modality features to get realized and transmitted, and thus also blocking the "gravitation" of the [+WH]-features to the location of modality. The usual Wh-Criterion effects will follow from this, i.e. no narrow-scope wh-question will be permitted due to lack of licensing of any wh-element in Spec/CP. Note, incidentally, that if this explanation is on the right track, non-factive complements wouldn't just be inflected infinitives, since they would have aspectual and modal features which infinitives, presumably, lack.

To summarize what we have done so far: We have posited a principled distinction between factive complements on the one hand, and non-factive and infinitival complements on the other, based on a difference in Tense and Aspect/Modality. We have adopted Rizzi's suggestion that an INFL node (or cluster of nodes) which have

rich tense (or other relevant) features attract the wh-features in the clause and make them somehow accessible to the head of the CP. Further, we have also adopted Rizzi's Wh-Criterion and have used it to explain the fact that Turkish does not allow for embedded infinitival wh-questions (and, for some speakers, for embedded non-factive wh-questions, either).

Let us now turn to a discussion of relative clauses.

It is not immediately obvious how we can apply the Wh-Criterion to explain the lack of infinitival (and non-factive) relative clauses in Turkish. This is because the CPs, i.e. the modifier clauses of the head of the relative clause, are not selected complements, and we are not dealing with question semantics. Hence, it is not plausible to posit inherent [+WH]-marking to the head of the CP.[5]

I would like to suggest nevertheless that the explanation is provided by the Wh-Criterion. First of all, subsequent work has shown the relevance of something like Rizzi's Wh-Criterion to other phenomena than questions--most notably, to negation (cf. Haegeman & Zanuttini 1990, Haegeman (1994)). Thus, perhaps, the most appropriate and general label for Rizzi's principle might be the Operator Criterion. If we treat the wh-element in relative clauses as operators, we would expect these constructions to exhibit appropriate effects.

Turning to the feature specification of the head of the CP, it will have to be in agreement with the features of the "relative operator". Since the CPs in these constructions are not selected, their head would not have inherent features, and we could exploit, once again, Rizzi's notion of "dynamic agreement", which we had used for matrix wh-questions. If we pursue this direction for an explanation, we would, once again, expect for such a head to have properties which, even though not inherently [+"Relative"], would have the potential to be thus marked under dynamic agreement. This potential could not come from non-factive or infinitival complements, which would be inherently marked with the negative value of the feature, and thus would have to come from complements with some tense and aspect/modality marking; this is what we find.

Alternatively, we might exploit the idea that there is some kind of predication relationship between the head of a relative clause and the modifying CP. We would have to specify that this kind of predication cannot take place between a head of a relative clause and a CP devoid of Aspect and Modality features. This relationship would impart the relevant positive feature to the head of CP, which would then have

---

[5]Sabel (1993) makes a suggestion which is akin in spirit to the one in the text. He suggests the existence of a [+rel] operator, on a par with [+WH] operators in questions.

to agree with respect to that feature with the "relative operator" which moves to Spec/CP.[6]

Once again, intermediate traces don't count as operators, only the wh-element in the highest Spec/CP of the relative clause does. This, then, accounts for the fact that the wh-element in a relative clause can be part of a non-factive or infinitival complement, as long as it ends up in the Spec of a CP which is headed by a "factive" marker. However, the operator cannot end up in Spec/CP of a non-factive or infinitival, for the obvious reasons spelled out above.

Now, a WH-element which moves to Spec/CP at LF has to obey Rizzi's WH-criterion, i.e. the element in C of that CP must be [+WH]. Therefore, constructions with the Future morpheme **-AcAK** are grammatical in a relative construction, since it is [+WH], while constructions with the [-WH] infinitive morpheme are ungrammatical.

4. The "potentiality" relative clauses have no counterparts as FRCs:

We have seen in the previous section that Turkish does not have genuine infinitival relative clauses, and why. We have also seen that, instead, there are irrealis relative clauses, morphologically marked with the Future tense marker, but with the syntax (PRO-subject, no agreement) and semantics (potentiality/irrealis) of infinitival relative clauses in better-studied languages like English. An example for this construction was (7), repeated here for convenience:

(7) Ahmet Ayşe$_j$ - ye [$_{NP}$ [$_{CP}$ PRO$_j$ e$_i$ oku - yacak] bir kitap$_i$] al-dı
-Dat. read-Fut. a book buy-past
'Ahmet bought Ayşe a book to read'

Interestingly, such irrealis relative clauses do not have FRCs as counterparts:

(31) *Ahmet Ayşe - ye [$_{NP}$ [$_{CP}$ PRO e$_i$ oku - yacak] Op$_i$] al-dı
-Dat. read-Fut. buy-past
Attempted reading: 'Ahmet bought Ayşe what (i.e. something) to read'

Given that Turkish does have FRCs otherwise, as we saw at the beginning of this paper, this restriction looks mysterious and calls for an explanation.

---

[6]For arguments to the effect that, although Turkish lacks relative pronouns and overt complements, its relative clauses involve some kind of operator movement which obeys Subjacency, see Kornfilt (1984).

5. Towards an Explanation:

I claim that the phenomena in 1, 2, and 4 are linked via one common explanation, sketched below: Turkish FRCs have a structurally realized, but phonologically empty head position.[7]

In addition, there is LF movement of the agreement morpheme to head position of the relative construction, motivated by the requirement that empty operators (like the abstract WH-element in Turkish) should not be governed. That there is such a requirement has been also proposed (for English) by Levin (1983), who uses it to explain contrasts like the one between the following two examples:

(32) a. * I know [$_{CP}$ op$_i$ [$_{IP}$ PRO to talk to t$_i$]]
b. I know [$_{DP}$ [$_{DP}$ the man] [$_{CP}$ op$_i$ [$_{IP}$ PRO to talk to t$_i$]]

In (32)a. the operator is governed by the matrix verb, yielding a violation of the non-government requirement on operators posited by Levin. In (32)b., on the other hand, the operator is not governed by any head. (Note that, if such an approach is on the right track, the notion of head government is a central one and cannot be subsumed under the notion of antecedent-government.)

Why should the phonologically empty operator in Turkish FRCs be governed by a governor (usually a verb) outside the DP? Even if we assume that Spec/CP is accessible to government from the outside in principle, it is not clear that it should be thus accessible in FRCs, if these constructions are analyzed as DPs with a DP head. Groos & van Riemsdijk (1979) suggest an analysis according to which FRCs have an empty head position, and that COMP (i.e. Spec/CP) in FRCs is universally accessible to government from outside in principle. They claim that languages differ with respect to whether government (and "rules" that would refer to government) refers to lexical material or structural positions, whether lexically filled or not. The Matching Effect would thus be explained.

Suppose, now, that Turkish is a kind of language where government would, in fact, refer to lexical material. Where the head of a relative clause is filled overtly, the Spec/CP position would be inaccessible to government from the outside. However, where the head would be phonologically empty, the Spec/CP position would be accessible to such government, and the prohibition against governed empty operators would be violated.

There are a number of conceptual and empirical problems with this account, pointed out in Bonneau (1990). I will not discuss those here and refer the reader to that paper. I do want to mention, however, Bonneau's alternative proposal, which has some appeal. He posits a rule of "wh-hopping" at LF, which shifts a wh-element from

---

[7]For a more detailed discussion of this particular claim, see Kornfilt (1984).

Spec/CP in a relative clause to the head position of the FRC. Matching Effects are thus explained, and the account has the virtue of treating those effects as a local phenomenon.

For the purposes of Turkish and the facts studied in the present paper, we do not have to take a stand between these two analyses. We saw in section 3 that there are good reasons to assume LF movement of a wh-operator to Spec/CP in embedded and matrix questions. Clearly, the same must be true for LF movement of an operator to Spec/CP in relative clauses.[8]

Now, whether we assume that this empty operator is accessible to outside government while located in Spec/CP, or whether we want to claim that if undergoes further movement to head position of the FRC, we have to ensure that it gets protected from such outside government, so as to avoid violating the constraint against governing empty operators.

I am suggesting here that such protection comes from the AGR morpheme, which moves to the head of the head position of the FRC (or else, under an analysis like Bonneau's, where the operator would occupy the head of the FRC, would adjoin to the head). In order to mark the scope of that LF movement, the corresponding morpheme must move to that position in the syntax. Thus, the agreement morpheme "protects" the empty WH-operator from government and thus also from Case assignment from outside. No Case clash can thus arise between the Case on the WH-element and the Case assigned from the matrix governor, and there are no Matching Effects as a result. In addition, the movement of Agr into the head position explains why we see the unexpected morpheme order **plural-agreement**, rather than the expected **agreement-plural**. Finally, the irrealis/potentiality construction cannot occur as a FRC, because it has no agreement morphology; therefore, there is no possibility for the rescue operation of moving the agreement morpheme into the head. As a result, the empty operator in Spec/CP is accessible to (indirect) government, leading to ungrammaticality as seen in (31).

One further property of Turkish FRCs should be mentioned here, albeit very briefly: They have a definite/specific interpretation, as opposed to an "operator-interpretation",[9] by which I mean the interpretation linked to elements like "whatever" in examples as the following one:

(33) I will wear whatever you will wear

---

[8]For arguments in favor of such movement of an empty operator in relative clauses, based in part on island-sensitivity, see Kornfilt (1984).

[9]Turkish does have a FRC construction with an "operator-interpretation". That construction is unrelated to the "regular" construction discussed in the body of this paper; it is not nominalized, but rather fully finite and bears the morphology of the conditional.

I would like to claim that this interpretation is due to the fact that Turkish FRCs are headed by AGR, which acts like a pronominal clitic. Had he construction been headed by an operator, the "operator-interpretation" would have obtained.[10]

Note that some irrealis participials have been relexicalized as full lexical nominals; under such readings, examples very similar to (31) are fine:

(33) Ahmet Ayşe - ye [$_{NP}$ [$_{CP}$ PRO e_i giy - ecek] Op_i] al-dı
-Dat. wear-Fut. buy-past

'Ahmet bought Ayşe some clothes' (rather than: 'Ahmet bought Ayşe what (i.e. something) to wear')

As such, these irrealis FRCs are sources of new lexical formations (e.g. yiyecek 'food=eat+Fut.', içecek 'drink=drink+Fut.', açacak 'can or bottle opener=open+Fut.', tutacak 'potholder=hold+Fut.'). Once lexicalized, we are not dealing with a relative construction any longer, and there is no (empty) operator present in the construction that needs to be protected from government.

Bibliography:

Alexiadou, A. & S. Varlokosta (1995) "On Free Relatives in Modern Greek"; ms., FAS-Berlin and the University of Pennsylvania.

Bechhofer, R. (1975) "WHO Said WHAT to WHOM?? ... in Turkish", in S. Kuno (ed.) Harvard Studies in Syntax and Semantics, vol. 1, Department of Linguistics, Harvard University, Cambridge, Massachusetts; 349-403.

Bonneau, J. (1990) "Logical Form and an Analysis of the Matching Effect in Free Relatives", in McGill Working Papers in Linguistics 6(2); 137-166.

Chomsky, N. (1981) Lectures on Government and Binding; Foris Publications, Dordrecht, The Netherlands.

Chomsky, N. (1993) "A Minimalist Program for Linguistic Theory", in K. Hale & S. J. Keyser (eds.), The View from Building 20, MIT Press: Cambridge, Mass.

Georgopoulos, C. (1991) "Canonical Government and the Specifier Parameter: An ECP Account of Weak Crossover"; Natural Language and Linguistic Theory 9, 1-46.

Groos, A. & H. van Riemsdijk (1981) "Matching Efect in Free Relatives: A Parameter of Core Grammar"; in Theory of Markedness in Generative Grammar; A. Belletti, L. Brandi, L. Rizzi (eds.); Scuola Normale Superiore di Pisa: Pisa

[10]A language which appears to be typologically the opposite of Turkish in this respect is Modern Greek, where the canonical FRC construction can only have "operator-interpretation", rather than the definite/specific interpretation. For a detailed study of Greek FRCs, see Alexiadou & Varlokosta (1995).

Haegeman, L. (1994) "Negative Heads and Negative Operators: The Neg Criterion", in Syntactic Theory and First Language Acquisition: Cross Linguistic Perspectives; vol. 1: Heads, Projections and Learnability; B. Lust, M. Suñer, J. Whitman (ed.s); Lawrence Erlbaum Associates: Hillsdale, New Jersey.

Haegeman, L. & R. Zanuttini (1990) "Negative Concord in West Flemish"; ms., Université de Genève.

Horvath, J. (1986) Focus in the Theory of Grammar and the Syntax of Hungarian; Foris Publications, Dordrecht, The Netherlands.

Huang, C.-T. J. (1982) Logical Relations in Chinese and the Theory of Grammar; unpublished doctoral dissertation, MIT, Cambridge, Massachusetts.

Kenesei, I. (1992) "Infinitival Complements and Related Non-finite Clauses in Finnish and Hungarian"; ms., University of Szeged, and University of Delaware.

Kornfilt, J. (1984) Case Marking, Agreement, and Empty Categories in Turkish; unpublished doctoral dissertation, Harvard University, Cambridge, Mass.

Kornfilt, J. (1988) " 'NP'-Deletion' and Case Marking in Turkish", in S. Koç (ed.); Studies on Turkish Linguistics (Proceedings of the Fourth Conference on Turkish Linguistics), Middle East Technical University, Ankara; 187-215.

Kornfilt, J. (1991) "Some Current Issues in Turkish Syntax", in H. Boeschoten & L. Verhoeven (eds.) Turkish Linguistics Today, Brill Publishers, Leiden, The Netherlands; 60-92.

Kornfilt, J. (1993) "Infinitival WH-Constructions and Complementation in Turkish"; in Eurotyp Working Papers; Group 3: Subordination and Complementation, vol. 4; K. Börjars and N. Vincent (ed.s); European Science Foundation; published at the University of Manchester; 66-83.

Kural, M. (1993) "V-To(-I-To)-C in Turkish", in UCLA Occasional Papers in Linguistics, vol. 11; Department of Linguistics, UCLA, Los Angeles, Cal.

Lees, Robert B. (1963) The Grammar of English Nominalizations; Indiana University, Bloomington; Mouton, The Hague, The Netherlands.

Levin, J. (1983) "Government Relations and the Structure of INFL." MIT Working Papers in Linguistics 5.

Lewis, G. L. (1967) Turkish Grammar; Clarendon Press, Oxford.

Nishigauchi, T. (1990) Quantification in the Theory of Grammar; Kluwer Academic Publishers, Dordrecht, The Netherlands.

Ortiz de Urbina (1992) "Tenseless Verbal Forms in Basque", ms., University of Deusto.

Özsoy, A. Sumru (1991) "A'-Dependencies in Turkish", ms., Boğaziçi University, Istanbul.

Picallo, C. (1984) "The INFL Node and the Null Subject Parameter," Linguistic Inquiry, 15; 75-102.

Raposo, E. (1987) "Case Theory and Infl-to-Comp: The Inflected Infinitives in European Portuguese", Linguistic Inquiry, 18, 85-109.

Rizzi, L. (1991) "Residual Verb Second and the Wh-Criterion", Technical Reports in Formal and Computational Linguistics; Faculté des Lettres, Université de Genève.

Rizzi, L. (1995) "The Fine Structure of the Left Periphery", ms., Université de Genève

Sabel, J. (1993) "W-Infinitive, Nominativsubjekte und das Wh-Criterion", Arbeitspapier Nr. 11, Universität Frankfurt, Institut für Deutsche Sprache und Literatur II: Frankfurt a.M.

Stowell, T. (1981) Origins of Phrase Structure, unpublished doctoral dissertation, MIT, Cambridge, Mass.

Stowell, T. (1982) "The Tense of Infinitives," Linguistic Inquiry, 13, 561-570.

Underhill, R. (1976) Turkish Grammar; MIT Press, Cambridge, Mass.

On learning and grammatical theory*

Paul Law
FAS, Berlin

0. Introduction

The central issue in linguistic theory is the question of what the speaker knows when he or she knows a language; in oher words, what constitutes grammar? Given that a child acquiring a language must be exposed to some linguistic environment, the problem of learning naturally arises. Therefore, an issue that is often raised in linguistic theory is whether a grammatical prime like the Subjacency condition on movement (Chomsky 1973) or the Empty Category Principle (ECP, Chomsky 1981) is learnable. In other words, the question is how the learner comes to acquire these (often abstract) grammatical principles.

The issue is all the more pressing and serious in the light of the fact that a competent speaker of a language has intuitions about what is possible and what is not. Suppose we idealize the linguistic ambiance that the learner is exposed to as containing *all and only* grammatical sentences, then one might claim that the learner would judge some example as possible if he or she has heard it before, and as impossible otherwise. But as Chomsky argued almost four decades ago, speakers have the potential of understanding and producing examples that they have never heard or spoken before. The linguistic competence must therefore contain a system of (abstract) principles which gives the speaker the intuitions about what is impossible.

However, no one would disagree that the linguistic ambiance that the learner is exposed to is rather impoverished, including grammatical and ungrammatical sentences. Worse still, the learner is generally not explicitly told which examples are grammatically possible, and which examples are not. Chomsky therefore concluded *on logical grounds* that this system of principles that constitute linguistic competence must be innate.

The concern about learning and grammatical theory is ultimately related to the interest in the make-up of the grammatical system. One thus hopes to be able to have some *independent measures* of justifying (abstract) grammatical primes of the linguistic competence. It is in this context that the issue of learning seems to be of relevance; after all, the learner must be exposed to linguistic data. We can think of a grammatical prime being learnable if and only if there is a learning algorithm that gives rise to that prime.

The issue of learnability of a grammatical prime and the question of whether it should be in the grammar are a priori independent. Therefore, we have four cases to consider, as tabled in (1):[1]

(1)	learnable	in the grammar	
a.	+	+	variations
b.	+	–	eg the null subject parameter
c.	–	+	universals, eg no movement of non-constituents
d.	–	–	eg surface filters

(1a) represents cases of language variations. (1b) represents cases where a prime can be shown to be learnable, but makes incorrect empirical predictions; hence, it is not in the grammar. (1c) represents cases of universals invariant across languages; these need not and perhaps cannot be learned. They are thus innate on logical grounds. (1d) represents uninteresting cases from the linguistic point of view. If we can show on empirical grounds that a grammatical prime

should not be in the grammar, then the fact that it is learnable or unlearnable is of no linguistic interest.

As it can be easily seen in the table, the viability of a grammatical prime is logically independent of its learnability. Being learnable does not imply being in the grammar (cf. (1a-b)), and being unlearnable does not imply not being in the grammar (cf. (1c-d)). Conversely, being in the grammar does not imply being learnable (cf. (1a-c)), and not being in grammar does not imply being unlearnable (cf. (1b-d)).

In what follows, I would like to consider each case in (1) with some concrete examples and argue that learning does not really tell us much about grammatical theory. I discuss a case of surface filter, the Doubly Filled COMP Filter (DFCF) of Chomsky and Lasnik (1977), and show that it is language variations that render it unviable, rather than learnability. I also give an analysis of how language variations with respect to the DFCF can be accommodated in Universal Grammar (UG, Chomsky 1957) without positing language-specific assumptions (section 1). I show how a grammatical prime like the null subject parameter can be shown to be learnable, and yet should not be taken to be in the grammar and suggest an alternative without any parameter to account for the same set of facts in terms of acquisition of phrase structure (section 2). I conclude the paper with a case of an unlearnable grammatical prime that should nevertheless be taken to be in the grammar, showing that learnability does not reveal the constitution of grammar (section 3).

1. Doubly Filled COMP

Chomsky and Lasnik (1977) pointed out that languages like English do not allow the co-occurrence of a *wh*-phrase and a complementizer in a local environment, as illustrated in (2):

(2) a. John wondered what (*that) Mary bought.
b. The man who (*that) John saw.

They thus suggested a surface filter like that in (3) to rule out examples of the type in (2):[2]

(3) Doubly Filled COMP Filter (Chomsky and Lasnik 1977:446)
*[$_{\text{COMP}}$ *wh*-phrase φ], φ≠e

A system of surface filters faces a number of conceptual problems, however. As Chomsky and Lasnik themselves pointed out, the filter in (3) is language-specific. Earlier stages of English allowed violations of the DFCF. In addition, the very specific syntactic description of the DFCF renders it construction-specific. It is quite obvious that language-specific and construction-specific grammatical primes seriously undermine explanatory adequacy. Language-specificity is in direct conflict with the conception of UG as a system of principles that are invariant across languages, and construction-specificity reduces the generality of the grammar. The DFCF in (3) is in fact a special case of the filter in (4), which has more specific conditions; others filters discussed in Chomsky and Lasnik (1977) like those in (5) and (6) also have construction-specific conditions:[3]

(4) *[$_{S'}$ ±WH [$_{NP}$ e] …], unless S′ or its trace is in the context: [$_{NP}$ NP ___ …]

(5) *[$_{NP}$ NP tense VP] except in the context
[$_{\alpha}$ [$_{\beta}$ X] ___], where X is a verb or *for*, β is its immediately dominating category, and α immediately dominates β.

(6) *[$_{S'}$ COMP NP …], where S′ is a root sentence.

It is not at all obvious why the very conditions as stated in the filters should hold. In the absence of a theory that take this question into consideration, and explains how the various filters are related, the filters are but a collection of arbitrary constraints. Another serious conceptual problem for the system of filters is that they do not seem to be related to anything else in the grammar, but are simply restatement of facts.

Chomsky and Lasnik (1977) remarked that it is hardly imaginable that the learner of English is explicitly told that examples of the type in (2) are impossible. Even if we assume that these examples might be in the input stream of data by performance errors, given that the learner is not told which examples are possible, and which are not, there seems to be no way the learner could deduce the existence of the DFCF from the linguistic input. If there is indeed no algorithm with which the learner comes to realize a constraint like the DFCF, then the filter cannot be learned.

By itself, the unlearnability of the filter is not necessarily problematic for linguistic theory, since they might very well be innate, hard-wired to the language faculty from the start. What is most problematic for the system of surface filters is the fact that it does not square well with the conception of UG as a system of cross-linguistic invariant principles on the one hand and language variations on the other. Consider Bavarian, for instance. Bavarian allows cooccurrence of a *wh*-phrase and a complementizer, as illustrated in (7) (Bayer 1984):

(7) a. I woaß daß Xaver des toa hod.
I know that Xaver this done has
'I know that Xaver has done this.'
b. I woaß ned wer daß des toa hod.
I know not who that this done has
'I don't know who has done this.'

(8) a. I wui wissn, wià schnäi daßd fahsd.
I want know how fast that-you drive
'I want to know how fast you drive.'
b. Schàmmà muàß mà si, wià gschlambbàd daßds ees dàheàkemds.
shame must man self how sloppy that you come around
'One must be ashamed of how sloppy you go around.'

Now, since the filter cannot be learned, it can only come from UG. The problem is that particular languages would have different UGs according as whether they allow violations of the DFCF. We thus arrive at an unacceptable conclusion. UG is, by hypothesis, invariant across languages.

An easy out of this problem is to assume that filters are in fact in UG as some sort of parameter with a negative value from the start, and that the value can be changed to positive when the learner has evidence for it from the linguistic input. Along these lines, speakers of English has the DFCF with the negative value in UG; they need not be explicitly instructed that examples of the type in (2) are impossible. For speakers of Bavarian, however, given that

they are exposed to examples like those in (7), they change the negative value of the DFCF to the positive one. Apart from the issue of whether particular grammars having different values for a specific filter have the same UG, the problems of construction-specificity and of explanatory adequacy still remain. It is hardly a satisfactory solution to replace filters with parameters that themselves do not have independent motivation.

From the point of view of Occam's Razor, a theory without filters is better than one with filters. I would like to pursue a better alternative in what follows. I would like to suggest that facts about DFCF be accounted for in terms of the different structures constructed. In particular, I claim that the non-interrogative complementizer *that* in English or *daß* in Bavarian, it has the *universal* property of disallowing a *wh*-phrase in its Spec, perhaps due to the independent fact that it may never head a *wh*-clausal complement (ie, embedded questions):[4]

(9) a. *John wondered that Mary bought the book.
b. John wondered what Mary bought.
c. *John wondered what that Mary bought.

(10) a. *Der Hans fragt si, daß de Maria das Buch kaffd.
the Hans ask himself that the Maria das book bought
'Hans wonders that Maria the book.'
b. Der Hans fragt si, was de Maria kaffd.
the Hans asks himself what the Maria bought.
'Hans wonders what Maria bought.'

Thus, knowing that a *wh*-phrase must land in SpecCP and that the complementizer *daß* does not permit a *wh*-phrase in its Spec, speakers of Bavarian on exposure to examples like those in (7) and (8) would have no choice but to build another CP on top of the one headed by *daß*, as in (11a), in order to accommodate the facts:

(11) a. … $[_{CP}$ was [$[_{CP}$ [daß $[_{IP}$ …
b. … $[_{CP}$ was [$[_{IP}$ …
c. *… $[_{CP}$ was [$[_{CP}$ [$[_{IP}$ …

When the learner hears examples like that in (10b) where in the embedded clause a *wh*-phrase appears without an overt complementizer, however, he or she would assign them single-CP structures like that in (11b). The reason for this single-CP structure, rather than a double-CP structure with an empty C position as in (11c), is that there is neither reason nor evidence for the latter.

Apart from embedded questions where a *wh*-phrase may appear in the Spec of an overt complementizer, Bavarian relative clauses permit a relative pronoun co-occurring with what is otherwise a *wh*-phrase:[5]

(12) a. Der Mantl den wo i kaffd hob.
the.NOM coat 3SG.MASC.ACC I bought have
'The coat which I bought.'
b. Der Mõ dem wo mir g'hoifa hom.
the.NOM man 3SG.MASC.DAT we helped have
'The man whom we have helped.'

c. A spuizeig, mit dem wo des Kind g'spuit hod.
a toy with 3SG.MASC.DAT the child played has
'A toy with which the child played.'

The relative pronoun *den* is homophonous with the singular masculine definite article in the accusative case and the singular masculine accusative case pronoun, and *dem* is homophonous with the one in dative case:

(13) a. I hob den Mõ g'säng.
I have the.ACC man seen
'I saw the man.'
b. Den kenn i ned.
3SG.MASC know I not
'I don't know it/him.'

and *wo* 'where' otherwise has the distribution of a *wh*-phrase in questions:

(14) a. Wo hod Xaver den Mantl kaffd?
where has Xaver the coat bought
'Where did Xaver buy the coat?'
b. I woaß ned wo Xaver den Mantl kaffd hod.
I know not where Xaver the coat bought has
'I don't know where Xaver bought the coat.'

If relative pronouns and *wh*-phrases universally appear in SpecCP, then the sequence of a relative pronoun and a following *wo* must involved two Spec positions. The structure of a relative COMP would be a double-CP structure like that in (15), where the relative pronoun and *wo* each occupies a Spec position:[6]

(15) ... $[_{CP}$ d-pronoun [$[_{CP}$ *wo* [$[_{IP}$ ...

A child learning Bavarian, with the knowledge of the syntax of relative pronouns and *wh*-phrases, he or she would again have no choice but to project a double-CP structure as in (15) in order to accommodate the examples in (12).

By contrast, English speakers lacking exposure to examples allowing doubly filled COMPs would have no evidence for building double-CP structures like those in Bavarian. They would have to move the *wh*-phrase to the Spec of *that*, violating the lexical property of the non-interrogative complementizer:

(16) *... $[_{CP}$ what [that $[_{IP}$ ...

My analysis of the doubly filled COMP facts has three advantages. First, languages that allow doubly filled COMPs and those that do not have the same UG in that their particular grammars do not contain a language-specific assumption distinguishing one from the other.[7] Second, despite their difference with respect to the DFCF, the non-interrogative complementizer (English *that* and Bavarian *daß*) has the same property in both types of languages, namely, it universally disallows a *wh*-phrase in its Spec. Third, it explains why

children learning languages obeying the DFCF like English never make mistakes in erroneously allowing an overt *wh*-phrase in the Spec of the complementizer *that*.

To the extent my analysis of doubly filled COMP facts is correct, it makes a case for the independence of learning and grammatical theory in that learnability by itself does not argue for or against a grammatical prime being in the grammar. Surface filters are not learnable, but are not in the grammar for reasons independently of learning. Rather, it is primarily language variations and the conception of UG as a system of principles invariant across languages that argue against the filters. It thus appears that quite generally we need not appeal to learning to decide whether a grammatical prime is in the grammar; linguistic evidence (syntax, semantics, phonology, etc) suffices.

2. Null subject

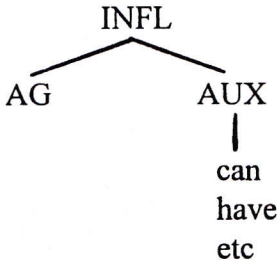
It is well-known that children systematically allow null subjects (NSs) in early stages of language development independently of whether the ambiant language allows NSs or not. Thus, the fact that children learning English and German, for instance, consistently have NSs cannot possibly attributed to the linguistic input they get, since these languages generally require an overt subject. Linguistic theory not only must bring other independent differences between NS languages (NSLs) and non-NSL to bear on NS, and explain why it is that children choose to have NS in early stages of language development, even if the ambiant language is a non-NSL, but also must provide an account of how they come to realize that the linguistic environment to which they are exposed comes from a NSL or a non-NSL.

In this section, I consider some analyses of NS that rely on a parameter, and discuss them in the context of language development (sections 2.1 and 2.2). I argue on empirical grounds independently of learning, however, that there is no need for such a parameter and propose to account for NS by pronoun incorporation, a process that crucially hinges on V-to-I verb movement (section 2.3). The issue of how a NSL or non-NSL is identified is then discussed (section 2.4), and an account for NS in the initial state of the language is suggested to be due to the process of acquisition of phrase structure (section 2.5).

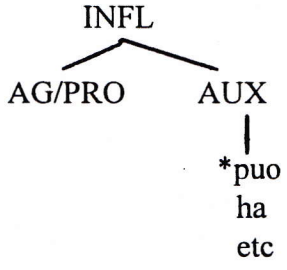
2.1. Null subject and INFL

Hyams (1989) claimed that language contains a NS parameter that she refers to as the AG/PRO parameter, which universally has the NS option as the initial value, and that the conditions sanctioning NSs are related to the syntactic properties of modals and auxiliary verbs. More specifically, she suggested that the structure of INFL is as in (17) where AG=PRO in NSLs like Italian or AG≠PRO in non-NSLs like English:

(17) a.



b.



NSLs cannot have modals in INFL since PRO would be governed, violating the constraint against government of PRO (Chomsky 1981:191), a result of the interaction of binding and government modules of the grammar:

(18) PRO is ungoverned.

The connection between NS on the one hand, and modals and auxiliary verbs on the other is thus established.

Zagona (1982) gave several arguments for a separate Aux constituent in English, including facts about tag-formation, negative placement, VP deletion and Subject-Aux inversion:

(19) a. Peter hasn't eaten, has he?
b. John will not finish his paper.
c. Mary isn't coming tonight, but Sue is.
d. Will Robert find his sunglasses?

But none of these is possible in Italian, a NSL. The fact that neither the negation *non* 'not' nor a pronominal clitic may intervene between the auxiliary *avere* 'have' apparently suggests that the Italian auxiliary verb does not occur in Aux:

(20) a. *Mario ha non mangiato. (cf. Mario non ha mangiato.)
'Mario has not eaten.'
b. *Mario ha lo mangiato. (cf. Mario lo ha mangiato.)
'Mario has eaten it.'

In addition, the impossibility of deleting a VP, stranding the auxiliary *essere* 'be' behind as shown in (21) might be taken to be evidence that the auxiliary does not appear in INFL:

(21) *Maria non è arrivata ancora, ma Gianni è.
'Maria hasn't arrived yet, but Gianni has.'

If the occurrence of the modals and auxiliary verbs to the left of the subject (cf. matrix questions in English) involves left-ward movement from INFL, then the ungrammaticality of the examples in (22) would follow from the assumption that Italian modals and auxiliary verbs do not appear in INFL:

(22) a. *Ha Gianni mangiato.
'Has Gianni eaten?'
b. *E Gianni arrivato.
'Has Gianni arrived?'
c. *Puó Gianni aiutarci.
'Can Gianni help us?'

The non-existence of tag-formation in Italian would be expected as well, if the construction indeed involves left-ward movement of INFL.

On the basis of the facts given in (20)-(22), Hyams concluded that the Italian INFL has the structure as in (17b) where modals and auxiliary verbs do not appear in INFL, but as main verbs inside VPs.

From the acquisition point of view, the emergence of modals and auxiliary verbs in the Aux position would exclude NS, as they would govern PRO in INFL (cf. the structures in (17)). The expectation seems to be borne out. Examples like those in (23) are quite common in

child language at the point where modals and auxiliary verbs have not emerged (Bloom, Lightbown and Hood 1975):

(23) Play it (Eric II)
Want more apple (Eric III)
Ride Dumbo (Gia II)
Eat piece (Gia II)
Touch milk (Kathryn II)
Want go get it (Kathryn III)

Combining facts in (23) with the absence of the English auxiliary *be* and the modals in the same period of time as well as the later (more or less) concurrent development of lexical subjects and the modals and the auxiliary *be*, as shown in (24):

(24) a. There's a birdie in there (Eric V)
There's a little ball (Gia V)
It's nice and clean (Gia V)
There's Humpty Dumpty up there (Kathryn III)
b. It doesn't fit (Eric IV)
What is the baby doing? (Gia V)
Foot goes over here (Kathryn III)

the connection between NS on the one hand and the auxiliary *be* and modals on the other seems quite natural.

If language universally has NS as the initial value of the NS parameter, then there must be some trigger in the linguistic input that signals to the learner that the ambiant language is a non-NSL. Hyams suggested that the detection of overt expletives and unstressed pronouns be the trigger for the resetting of the NS parameter to the non-NS value.

Some facts about acquisition of German appear to further corroborate Hyams' analysis of NS. Children learning German systematically allow NS (Clahsen and Muysken 1983) and OV word-order in early stages of acquisition (Park 1970, cited in Roeper 1973), despite the fact that a lexical subject is required and main clauses have the VO word-order in the adult language. Thus, examples of the type in (25) are quite common in early stages of acquisition of German (from Mills 1987):

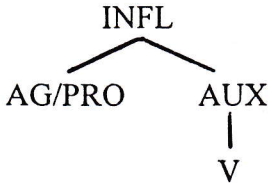
(25) Teddy holen
'fetch teddy'
Hause gehen
'go home'
Meike abmachen
'take (it) off Meike'
Teddy sofa fahren
'teddy drives the moped'
Meike fenster gucken
'Meike is looking out the window.'

If the VO order is derived from the OV order by first moving the verb to INFL, which is then preposed past the object (Thiersch 1978), as shown in (26):

(26) a. Hans $[_{VP}$ Maria e_i $]$ $[_{INFL}$ liebt$_i$ $]$
b. Hans $[_{INFL_j}$ liebt$_i$ $]$ $[_{VP}$ Maria e_i $]$ $[_{INFL}$ e_j $]$

then the fact that NS is inconsistent with the VO order follows immediately. This is because PRO in INFL representing the NS would be governed when the verb moves to INFL:

(27)



By contrast, NS is possible with the OV order. This is because the verb is not in INFL, and PRO in INFL is not governed.

In Hyams' analysis, then, particular languages differ with respect to whether AG is PRO or not. If AG=PRO, then the language allows NS but not modals in INFL; if AG≠PRO, then language permits modals in INFL but not NS. For the learner, it is not difficult to figure out whether the ambiant language is a NSL or not. If it is a NSL, then he need not do anything. The initial setting of the NS parameter would remain as is. However, if the learner detects the existence of modals as evidenced in tag-formation, negative placement, VP deletion and Subject-Aux inversion, or the presence of overt expletives and unstressed pronouns, or if the learner realizes that the verb moves to INFL, then he or she must set the NS parameter to the non-NS value. As what the learner needs in order to reset the NS parameter is positive evidence, identifying a non-NSL is unproblematic.

Given that the parameter is learnable, should it be in the grammar? How do we decide? Does learnability bear on the issue? I would like to argue that there are both conceptual and empirical reasons to suppose that the parameter is not in the grammar, despite its learnability.

Conceptually, Hyams' analysis of NS crucially hinges on the assumption that PRO cannot be governed. Thus, if it turns out that the distribution of PRO is not incompatible with PRO being governed (Jaeggli 1980, Bouchard 1982), then we can no longer attribute the impossibility of NS to government of PRO in INFL. In addition, it is not at all clear that PRO should be allowed to occur inside INFL as in (17b). PRO seems to have the distribution of a maximal projection, as illustrated in (28)-(29):

(28) a. [PRO to please John] is easy.
b. [for [Bill to please John]] is easy.
c. [for [[Bill's friends] to please John]] is easy.

(29) a. John bought the violin [PRO to play with]
b. John bought the violin [for [Mary to play with]]
c. John bought the violin [for [[the child he taught] to play with]]

one has to make some additional assumption to the effect that either a maximal projection may occur inside a X° as in (17b), or PRO in this case is in fact a X° which does not project maximally. It is difficult to see how one can go about justifying such an additional assumption.

Empirically, the facts presented in (20)-(22) do not seem to bear on NS. In particular, the impossibility of these examples certainly does not establish the link between NS on the

one hand and the syntax of modals and auxiliary verbs on the other. French pretty much patterns with Italian with respect to VP deletion and Subject-Aux inversion as shown in (30) and (31):

(30) *Marie n'est pas encore arrivé, mais Jean est.
'Marie hasn't arrived, but Jean has.'

(31) a. *A Jean mangé?
'Has Jean eaten?'
b. *Est Jean arrivé?
'Has Jean arrived?'
c. *Peut Jean nous aider?
'Can Jean help us?'

The lack of tag-formation in French follows from the fact that it generally does not allow an inflected verb to appear to the left of a non-pronominal subject, as shown in (31). Since French clearly does not allow NS, the impossibility of the French examples in (30) and (31) certainly does not bear on the conditions licensing NS. Therefore, the ungrammaticality of the comparable Italian examples in (20)-(22) tells us something about the positioning of verbs, negation and clitics in the language, rather than about NS. It does not reveal anything about NS in Italian any more than it does in French.

In fact, there are problems with the arguments based on the Italian data presented above. Although it is true that the negation *non* 'not ' and clitics may not intervene between a modal and a main verb as shown in (20), some other adverbial elements like *scrupolosamente* 'carefully', *daverro* 'indeed', *attentamente* 'attentively', *mai* 'never', *più* 'any longer' may:

(32) a. Gianni ha scrupolosamente studiato l'articolo.
'Gianni has carefully studied the article.'
b. Gianni non ha mai studiato l'articolo.
'Gianni has never studied the article.'
c. Gianni ha daverro studiato l'articolo.
'Gianni has indeed studied the article.'

The ungrammaticality of the examples in (20) is therefore due to the syntax of negation and clitics rather than that of the verb.

The relative positioning of verbs and adverbs in Italian is quite similar to that in French. An inflected main verb in French may be separated from its object by an adverbial (cf. Emonds 1978, Pollock 1989):

(33) a. Gianni studia attentamente l'articolo.
'Gianni carefully studies the article.'
b. Gianni non studia mai l'articolo.
'Gianni never studies the article.'

(34) a. Jean étudie soigneusement l'article.
'Jean carefully studies the article.'
b. Jean n'étudie jamais l'article.
'Jean never studies the article.'

It goes without saying that a clitic may not appear between an inflected auxiliary verb and an uninflected main verb in French either. The example in (35) is on a par with that in (20b):

(35) *Jean a le mangé. (cf. Jean l'a mangé.)
'Jean has eaten it.'

As the syntax of verb is more or less the same in both Italian and French, it cannot possibly be that the positioning of the verb is the reason for the difference between the two with respect to NS. Consequently, we have no reason to suppose that the grammar contains an AG/PRO parameter that relates NS to the syntax of modals and auxiliary verbs. It is quite clear that we need not appeal to learning in order to reach this conclusion.[8]

2.2. Null subject and morphological uniformity

Jaeggli and Safir (1989) claimed that NS is found in languages that have rich inflectional morphology in the verbal paradigms. That is, all inflected forms in a verbal paradigm contain some morphological marking encoding person and number features. In addition, they also claimed that NS is found in languages that have no such morphological marking, and suggested the parameter in (36) for NS:

(36) The Null Subject Parameter (NSP)
Null subjects are permitted in all and only languages with morphologically uniform inflectional paradigms.

Morphological uniformity is in turn taken to be as in (37):

(37) Morphological Uniformity
An inflected paradigm **P** in language **L** is morphologically uniform iff **P** has either only underived forms or only derived inflectional forms.

where a word **W** of category **K** is underived if it is morphologically non-distinct from the stem (or root) of **W**, and is derived if it is formed of a stem (or root) **W** plus an affix attached to **W**.

In this view, identification of a language as a NSL or not is rather trivial. All that is required of the child is to find out whether the ambiant language has a uniform inflectional paradigm or not in the sense of (37). Thus, if the ambiant language is a language like Spanish where the verbal inflectional paradigm has distinct forms for different persons and numbers, or if the ambiant language is a language like Japanese where the same verbal form obtains for all persons and numbers, then the child would realize that the ambiant language is in fact a NSL:

(38)

Spanish		Japanese
habl-o	'I speak'	ii-masu
habl-as	'you (sg) speak'	'I/you/we/they speak'
habl-a	'he/she speaks'	or 'he/she/it speaks'
habl-amos	'we speak'	
habl-áis	'you (pl) speak'	
habl-an	'they speak'	

The NSP in (36) is therefore learnable. Should it be in the grammar, though? How do we decide?

Despite the quite impressive range of languages, from Italian and Chinese on the one end to Icelandic and German on the other, that are correctly predicted by the NSP to be a NSL or not, it fails in the case of Haitian Creole (HC) and many other African languages. The inflectional paradigms in HC is as uniform as that in Chinese and Japanese in that it has the same form for all persons and numbers; yet, NS is impossible:

(39) a. mwen/ou/li/nou/yo pati.
1SG/2SG, 2PL/3SG/1PL/3PL left
'I/you(sg)/he/she/it/we/you(pl)/they left.'
b. Jan/nèg la/ *∅ pati.
Jan man DET leave
'Jan/the man left.'

According to Koopman (1984:173ff), the verb in the West African language Vata does not carry inflectional markings expressing subject verb agreement, but NS is not possible:[9]

(40) a. ń/n̄/Ó, Ú, É.../à/á/wá gbā.
'I/you/he, (3rd pers)../we/you/they speak.
'I/you/we/you/they speak.' or 'he/she/it speaks.'
b. *gbā.
'I/you/we/you/they speak.' or 'he/she/it speaks.'

HC and Vata are thus straightforward counterevidence to the NSP.

Once again, we can see that the fact that a grammatical prime is learnable does not necessarily imply that it should be in the grammar. If one is interested in the question of what constitutes grammar, one can bring cross-linguistic variations to bear on the issue. There is no need to appeal to learning.

2.3. Null subject as pronoun incorporation and null arguments

The viability of the NSP depends not only on the arguments in favor of it, but also on the alternative accounts for NS. I would like to provide an alternative account of NS with no parameter, again, without appealing to learning. To the extent that my alternative is correct, it argues for the independence of syntactic theory and learning in that one can present cross-linguistic empirical facts rather than learning-theoretic grounds to argue for a grammatical prime being in the grammar.

Law (1993) argues that NSLs are languages that have pronominal incorporation from SpecVP, rather than are the results of setting of a NS parameter. The condition for incorporation from SpecVP is crucially contigent on V-to-I verb movement, as the trace of the incorporated subject in SpecVP would only be properly governed from I:

(41) ... $V°_i$+I° [$_{VP}$ adverb [$_{VP}$ t_i ...

We have ample evidence that the verb moves to I in Italian. Examples in (32)-(33) are instantiations of the schema in (41). The lack of the *that*-trace effect (Rizzi 1982) when the subject is extracted from SpecVP further corroborates the claim that the verb moves to I in Italian:

(42) a. Chi credi che verrà?
'Who do you think will come?'
b. ... verrà$_i$+I° [$_{VP}$ t_i ...

The grammaticality of the example in (42a) is due to the fact that the subject trace in SpecVP is properly head-governed by the verb that has moved to I. The trace of a pronominal subject incorporating into I would also leave a properly head-governed trace, giving rise to NS.

The reason why English does not have NS is precisely because it does not have V-to-I verb movement. Incorporation of a pronominal subject from SpecVP would leave an ungoverned trace, violating the ECP. Why do French and Germanic languages lack NSs then, given that they do have V-to-I verb movement? The reason for this is that they do not have pronominal incorporation. Subject pronouns in these languages exhibit non-X° syntactic properties. The clearest piece of evidence for this is the fact that a pronominal subject might appear to the right or to the left of an inflected verb:

(43) a. Que suis-je? vs Je suis étudiant.
'Who am I?' vs 'I'm a student.'
b. Wen kennen sie? vs Sie kennen ihn.
'Who do they know?' vs 'They know him.'

The distribution of unstressed subject pronouns (cf. footnote 10) and verbal morphology is good diagnostics of pronominal incorporation. In languages like French that have some inflectional morphology, the verb forms may have formally different endings in accord with the number and person features and the verb stem to which the ending attaches, but the subject pronouns are always the same regardless of the verb stem:

(44)

être 'to be'		préférer 'to prefer'	
je 'I'	suis 'am'	je 'I'	préfère 'prefer'
tu 'you'	es 'are'	tu 'you'	préfères 'prefer'
il/elle 'he/she/it'	est 'is'	il/elle 'he/she/it'	préfère 'prefers'
nous 'we'	sommes 'are'	nous 'we'	préférons 'prefer'
vous 'you'	êtes 'are'	vous 'you'	préférez 'prefer'
ils/elles 'they'	sont 'are'	ils/elles 'they'	préfèrent 'prefer'

That is, the ending of the first person, plural, non-past tense of the verb *être* 'to be', for instance, is different from that of the verb *préférer* 'to prefer', but the subject pronoun is always *nous* 'we'. French unstressed subject pronouns clearly do not incorporate into the verb.

By contrast, in languages like Italian there is no necessary morphology marking the presence of a subject pronoun. There is no identifiable morpheme that always appears with the verb form and that can be taken to represent the subject:[10]

(45)

guardare 'to watch'		preferire 'to prefer'	
guardo	'I watch'	preferisco	'I prefer'
guardi	'you watch'	preferisci	'you prefer'
guarda	'he/she/it watches'	preferisce	'he/she/it prefers'
guardiamo	'we watch'	preferiamo	'we prefer'
guardate	'you watch'	preferite	'you prefer'
guardano	'they watch'	preferiscono	'they prefer'

One might take the endings *-o*, *-i*, *-a*, *-iamo*, *-ate/-ite* and *-ano/-ono* as enclitic subject pronouns, but one would then have to say that these pronouns have different morphologies according to the verb stem to which they attach. That is, we would have to say that the third person, plural subject pronoun, for instance, is the ending *-ano* when it attaches to the stem *guardare* 'to watch', but is the ending *-ono* when it is suffixed to the stem *preferire* 'to prefer'. One would then face two problems.

First, we would have to explain why subject pronouns have different forms according as the stem to which it attaches, but non-subject pronouns do not. As shown in (46), the object pronoun in Italian remains the same whether the stem verb is *guardare* 'to watch' or *preferire* 'to prefer', on a par with French:

(46) a. lo/la/le/li guard*ano*.
3SG.MASC/3SG.FEM/3PL.MASC/3PL.FEM watch.3PL
'They watch him/her/them.'

b. lo/la/le/li preferisc*ono*.
3SG.MASC/3SG.FEM/3PL.MASC/3PL.FEM prefer.3PL
'They prefer him/her/them.'

(47) a. Ils le/la/les regardent.
3PL 3SG.MASC/3SG.FEM/3PL watch.3PL
'They watch him/her/them.'

b. Ils le/la/les préfèrent.
3PL 3SG.MASC/3SG.FEM/3PL prefer.3PL
'They prefer him/her/them.'

Second, we have to account for why the form of the Italian subject pronoun may vary, but that in French does not. The best way to avoid these two problems is, it seems, to consider the different verb forms as shown in (45) as morphological manifestations of the amalgamation of the verb and the incorporated subject pronoun. The first problem is solved since morphological spell-out of the amalgam of a subject pronoun and a verb varies as a function of the conjugation class (*-are*, *-ere* or *-ire* and a few irregular verbal paradigms) to which the verb belongs. The second problem is also solved since one need not commit oneself to subject pronouns having any particular forms, as they are incorporated into the verb. The issue of different forms of the subject pronoun depending on the verb stem hence does not arise.

If this view of NS is correct, then there is simply no need for a NS parameter. Theory of grammar should then not contain such a parameter, on parsimonious grounds. What of null argument languages (NALs) like Chinese and Japanese, where both subjects and objects may be phonetically null? As evidence for V-to-I verb movement in these languages and verbal inflection for person and number are lacking in these languages:

(48) a. Zhangsan zǐ-xì-de kàn nèi-běn shū.
Zhangsan detail read that-CL book
'Zhangsan is carefully reading that book.'

b. *Zhangsan kàn zǐ-xì-de nèi-běn shū.

(49) Wǒ/nǐ/tā/wǒ-mén/nǐ-mén/tā-mén lái.
1SG/2SG/3SG/1PL/2PL/3PL come
'I/you/we/you/they come or he/she/it comes.'

Law (1993) suggested that phonetically null arguments in these cases are possibly due to the independent conditions on *use* of pronouns. In these languages, it is almost impossible to use pronouns to refer to inanimate objects; use of demonstratives for these cases is the norm:[11]

(50) Wǒ zúo-tien mǎi-le yì-běn shū.
I yesterday buy-ASP one-CL book
'I bought a book yesterday.'
a. nèi-běn shū hén yǒu qù.
that-CL book very have taste
'That book is very interesting.'
b. ?hén yǒu qù.
very have taste
'It is very interesting.'
c. *tā hén yǒu qù.
3.SG very have taste
'It is very interesting.'

It is perhaps the unproblematic recovery of inanimate referents from the discourse context that induces the use of null pronouns for animate referents as well. If this is correct, then theory of grammar should not contain even a null argument parameter, for the same reason why it does not have a NS parameter.

2.4. Identifying languages

There remains, though, the question of how a child would come to realize that the ambiant language is a NSL, a non-NSL, or a NAL. In fact, the last type of languages is most easily identified. If a language permits null object, then it necessarily allows other arguments to be phonetically null. That is, on the basis of examples like those in (51b) and (52b) the ambiant language is inevitably a NAL:[12]

(51) a. A: Zhangsan kàn-wán-le nèi-běn shū méi-yǒu?
Zhangsan see-finish-ASP that-CL book not-have
'Has Zhangsan finished reading that book?'
b. B: Kàn-wán-le.
see-finish-ASP
'He has.'

(52) a. A: Lisi ài-bú-ài Máo zhǔ-xí?
Lisi love-not-love Mao chairman
'Does Lisi love chairman Mao?'
b. B: Ài.
love
'He does.'

In the absence of examples of the sort in (51b) and (52b), the learner would have no reason to suppose that the ambiant language allows null arguments. Identifying a NSL like Italian is also unproblematic, as evidence for the presence of an object pronoun is rather

strong. In contrast with NS, the presence of the pronominal object is prominent not only in simple cases like that in (53):

(53) a. non lo vedo bene.
not 3SG.MASC see well
'I don't see him well.'
b. la vogliamo.
3SG.FEM want
'We want her.'

but also in cases of clitic-doubling and participial argeement:

(54) a. Non, questo giornale non lo conosco.
no this newspaper not 3SG.MASC know
'No, I don't know this newspaper.'
b. Sì, le capsule le prendo sempre.
yes the capsuls 3PL.FEM take always
'Yes, I always take the capsules.'

(55) a. Non, non l' ho ancora comparto.
non not 3SG have yet buy.SG.MASC
'No, I haven't bought it yet.'
b. Sì, l' ho già invitata.
yes 3SG have already invite.SG.FEM
'Yes, I have already invited her.'

(56) a. Non, non li ho ancora comparti.
non not 3PL.MASC have yet buy.PL.MASC
'No, I haven't bought them yet.'
b. Sì, le abbiamo viste ieri.
yes 3PL.FEM have see.PL.FEM yesterday
'Yes, we have seen them yesterday.'

The absence of a pronominal object induces different participial morphology:

(57) a. Sì, ho già invitato Luisa.
yes have already invite
'Yes, I have already invited Luisa.'
b. Non, non ho ancora comparto i fiori.
non not have yet buy the pens
'No, I haven't bought the pens yet.'
b. Sì, abbiamo visto Luisa e Chiara ieri.
yes have see and yesterday
'Yes, we have seen Luisa and Chiara yesterday.'

A learner of Italian would therefore have ample evidence for the NS versus overt object asymmetry, and can identify it as a NSL like without any problem.

By contrast, children learning French would have no evidence for the subject/object pronoun asymmetry of the type seen in Italian, even though the distribution of object pronoun and the concomitant agreement paradigms are very similar to those of Italian. As shown in (58), an object pronoun may double an overt full noun phrase:

(58) a. Ce journal, je ne le connais pas.
this newspaper I 3SG know not
'I don't know this journal.'
b. Les capsules, je ne les prends toujours.
the capsule I 3PL take always
'I always take the capsules.'

and the presence of an object pronoun may trigger participial agreement as well:[13]

(59) a. Non, je ne l' ai pas encore prise.
no I 3SG.FEM have not yet take
'No, I have not taken it yet.'
b. Oui,je l' ai déjà peinte.
yes I 3SG.FEM have already paint
'Yes, I have already painted it.'

(60) a. Non, je n' ai pas encore pris la photo.
no I have not yet take the picture
'No, I have not taken the picture yet.'
b. Oui,j' ai déjà peint la porte.
yes I have already paint the door
'Yes, I have already painted the door.'

For the learner, the contrast between the examples in (53)-(57) and those in (58)-(60) is not the presence of the object pronoun and the agreement patterns, but rather the subject/object pronoun asymmetry. While the subject pronoun is obviously phonetically null in Italian, that in French is prominently present. The learner of French would thus notice that the subject may not be phonetically null, and that the language is not a NSL.

Lacking exposure to examples of the type in (51b) and (52b) that are observed in Chinese where both arguments may be phonetically null, and without evidence for the NS versus overt object asymmetry of the type seen in Italian, the learner would have no choice but to conclude that an ambiant language like English or HC is a non-NSL.[14]

2.5. Acquisition of phrase structure

The most interesting question for linguistic theory and acquisition theory is why children initially permit NS *regardless* the ambiant linguistic environment. In particular, as English does not allow NS, the fact that children learning it initially permit NS cannot possible be due to the linguistic input, but must be related to the initial state of the grammar. A logical possibility is that the grammar has some parameter that has NS as the initial value, as Hyams proposed. However, if my claim is correct in that there is no such parameter, which is desirable on parsimonious grounds, then how can we account for NS in the initial state of the grammar, especially when it occurs in languages like English that do not allow NS?

I would like to suggest that the reason why the grammar has NS in the initial state be related to the acquisition of phrase structure; more specifically, children at early stages of language development have mastered the structure for VP, but not that of IP. The lack of the IP-projection implies the absence of a SpecIP position for the subject, and hence the impossibility of a lexical subject.

A range of disparate facts in child language are correctly predicted in this view. The lack of INFL explains the absence of modals, auxiliary verbs and inflectional morphology child English (Brown 1973), if these elements in fact occur in INFL as standardly assumed (cf. Chomsky 1957, Emonds 1978, Zagona 1982 and Pollock 1989). If the German verb in right-periphery position as in (25) (cf. section 2.1) is indeed in its base-position (Koster 1975), then the OV order in early child German would follow directly from the lack of an IP-projection, there being no INFL position for the verb to move to.[15]

There are some reasons to think that acquisition of phrase structure proceeds from smaller units to bigger units, ie from smaller constituents to larger constituents. Although it is intuitively clear that acquisition develops from single syllable (*'chine* for *machine* Eric I) to what can be loosely called one-word (*baby* Gia I, *nother, umbrella* Kathryn I) to multi-words (*another clown* Eric III, *baby chicken* Gia III), it is not immediately obvious why children do not produce sequences of words that do not form a constituent. While we can consider the one-word stage to be the point where children perform the fundamental task of acquisition in matching phonetic matrices from the linguistic environment with syntactic categories, taken to be bundles of abstract features (±N, ±V, ±singular, etc), it is logically possible that children could produce sequences of two or more words that are formally non-constituents in adult grammar. Thus, children produce rather complex sequences like those in (61a), but do not go through a stage where expressions like those in (61b) could have been uttered given that they have acquired the individual words:[16]

(61) a. Why you want do that? (Eric V)
I ride my bike (Gia V)
Kathryn go get book (Kathryn III)
I turn the light in (Peter VI)

b. Why do?
ride my
Kathryn get
turn the

What this means is that phrase structure must be built piece-meal in that a category must project maximally with its internal structure before it can be put in relation to some other category. In other words, knowledge of complex structures is a function of knowledge of the constituent parts.

Along these lines, then, children must acquire the structure of VP before they master that of IP, since VP is an integral part of IP. We now have an explanation for NS in early child language. NS is possible at the early stage of development knowledge since the structure of VP has, but that of IP has not yet, been acquired. The lack of a SpecIP position follows directly from the absence of the IP projection; NS is thus an immediate consequence of the fact that there is no position for the subject.

3. Conclusion

In the foregoing sections, I argued that one can decide whether a grammatical prime is in the grammar by considering the empirical facts that bear on it. In particular, language variations are good testing grounds for the empirical bearing of a grammatical prime. I showed that learnability of an abstract grammatical prime does not tell us much about syntactic theory in that one cannot accept or reject it as part of the grammar on learnability grounds.

Let us now return to the table in (1), repeated here as (62):

(62)

	learnable	in the grammar	
a.	+	+	variations
b.	+	–	the null subject parameter
c.	–	+	universals, eg no movement of non-constituents
d.	–	–	eg surface filters

We have discussed three cases in (62a), (62b) and (62d). Let us now turn to the case in (62c) by considering the principle in (63):[17]

(63) X can be moved only if X is a constituent.

The principle in (63) is to capture the grammatical contrast in (64):

(64) a. John said that he would go to the store, and [gone to the store]$_i$ he has t_i.
b. *John said that he would go to the store, and [gone to]$_i$ he has t_i the store.

If *gone to the store* has the structure in (65), then the reason why one can move *gone to the store* but not *gone to* is because the former is a constituent, but the latter is not:

(65) [$_{VP}$ gone [$_{PP}$ to [$_{NP}$ the store]]]

How can a child learn the principle in (63)? The child is certainly not explicitly told that examples of the type in (64b) are impossible, and is instructed that only constituents may be moved. That is, the child has no direct access to the principle in (63) from the linguistic environment. If the principle cannot be deduced from linguistic input, then the only possible explanation for the child's knowledge of it is that it is part of his or her innate language faculty, this being the only logical alternative for the source of such knowledge. In other words, a grammatical prime may be in the grammar, even though it cannot possibly be directly deduced from the ambiant language.

If what we are interested in is whether a grammatical prime is in the grammar, then we should look from right to left in the table in (62). It is quite easy to see that learnability does not bear on the issue. A prime may be in the grammar, whether it can be learned or not. This conclusion is not at all surprising. As Chomsky has pointed out, it is pointless to study the issue of language learning if we have no idea what it is that has to be learned. Now, if what has to be learned is *prior to* how it is learned, then one cannot answer the question of what constitutes grammar by appealing to learning.

This is of course not to say that learning has no independent interest. Insofar as children have to be exposed to some linguistic environment in order to attain the final state of grammar, theory of learning will tell us how such a process develops. But it is not revealing about the primes that constitute grammar.

Notes

* I would like to thank Artemis Alexiadou, Brigitta Hafka, Jaklin Kornfilt and André Meingunger for very helpful comments and discussion of an oral presentation of this paper. Inadequacies are my responsibility.

[1] As we will see, the constructions we discussed here are learnable *in principle*. What is meant by unlearnable in (1) is that the learner has no access to a learning procedure.

[2] In terms of current phrase structure, the DFCF would be formulated as a constraint banning an overt *wh*-phrase in the Spec position of an overt complementizer.

[3] The filter in (4) is to distinguish (i) and (ii):

(i) a. *Who$_i$ did you wonder [whether t_i saw Bill]
b. *Who$_i$ do think [that t_i saw Bill]

(ii) The man [O_i [that t_i saw Bill]]

The filter in (5) is to force the presence of *for* in the examples in (iii) and (iv):

(iii) a. His plan *(for) Bill to win.
b. It bothers me *(for) Bill to win.
c. It is illegal *(for) Bill to win.
d. I want very much *(for) Bill to win.
e. There is someone at the door *(for) you to play with.
f. *(For) John to take the job would be preferred.

(iv) a. What wage$_i$ would you work [for t_i] [*(for) your kids to have a chance to go to college]
b. Who$_i$ would it [bother t_i] [*(for) your kids to have a chance to go to college]

The filter in (6) is to rule out the examples in (v):

(v) a. *that John is here.
b. *whether John is here.
c. *who John saw.

[4] Formally, one might want to make use of a [WH] feature, eg *that* has a [–WH] feature, and hence disallows a *wh*-phrase in its Spec. It is not my immediate concern here to dwell on the technical detail.

[5] Bayer (1984) points out that when *wo* is present, the "unmarked" relative pronouns unshaded in (i) may be dropped:

i)

	masc	neut	fem	plur
NOM	*der*	*des*	*die*	*die*
ACC	*den*	*des*	*die*	*die*
DAT	*dem*	*dem*	*der(a)*	*dene(n)*

For the " marked" ones falling in the shaded areas, they may be dropped just in case they are phonetically identical to the determiner of the noun phrase which the relative clause modifies.

[6] When the relative pronoun drops, the structure of a relative CP would be a single CP-projection as in (ii):

(ii) ... $[_{CP}$ wo [$[_{IP}$ ...

The structure in (ii) is parellel to that in (11b).

[7] Along these lines, then, some cross-linguistic variations are not due to grammar-internal properties, but rather are driven by external linguistic environment. As languages vary with respect to the DFCF, it cannot be that some grammar-internal property that allows or disallows COMP to be doubly filled. Moreover, there seems to be no other property that can be linked to doubly filled COMP (cf. German, a language that is very much like Bavarian, is like English with respect to the DFCF); therefore, it does not appear likely that it is some particular grammar-internal property that gives rise to doubly filled COMP. If this is correct, then variations with respect to doubly filled COMP, if they are indeed due to the external linguistic environment, can only be taken to be accidental facts. We have seen that theory of grammar can very well accommodate these facts by allowing building of phrase structure in accord with the independent properties of the elements constituting that phrase structure.

An important issue that immediately arises is how theory of grammar-external variations can explain why the observed facts exist, but other logically possible variations do not. While it seems clear that grammatical theory should be sufficiently constrained so as to limit the possible variations (eg, no language may have nouns taking verbs as complements), it is not immediately obvious why some specific variations exist but not others. For instance, why should it be that double-CP structures are possible (to the extent that my proposal can be sustained), but triple-CP or double-IP structures, are not known to exist? An adequate answer to this question has to await future research.

[8] Hyams (1992) claimed that her earlier analysis of NS has the following problems and suggested a new analysis in terms of Jaeggli and Safir's (1989) theory of NS (cf. section 2.2):

(i) a. The development of tense accompanying the transition to a non-NS grammar is unaccounted for.
b. The emergence of infinitive marker *to* alongside modals is not predicted.
c. Modals infrequent initially.
d. NS are "unidentified."

However, I do not think that these are really problems for Hyams' (1989) analysis. First, if tense is in INFL as standardly assumed, then the emergence of non-NS alongside tense morphology would be just as expected, since tense morphology would govern PRO in INFL. Second, the emergence of the infinitive marker *to* would be predicted to be at the same time as that of modals if it is assumed to be in INFL, a not unreasonable assumption given that tense or finiteness of a verb is marked either bby inflectional morphology or an infinitival marker like*to*. Third, as Hyams (1992) also pointed out, the infrequent use of modals in the initial period might be due to lexical learning. Fourth, there is no independent justification for "identification" of null pronouns. The referent of a null pronoun is recoverable from discourse context to the same degree as the referent of an overt pronoun is. It is then uclear why null pronouns must be "identified" any more than overt pronouns are.

[9] The strong forms of the pronoun may be used by themselves as answers to a *wh*-questions, in contrast with weak forms of the pronoun (Koopman 1984:74):

(i) a. àlÓ Ò wà sàká là?
who he-R want rice WH
'Who wants some rice?'

b. àmĭ/*n.
me (strong form)/I (weak form)
'me.'
c. ànyĭ/*à.
us (strong form)/we (weak form)
'us.'

Like the weak forms, the strong forms do not trigger subject verb agreement either. Based on facts about coordination and the strong/weak distinction for object pronouns, Koopman argues that the weak forms are not syntactic clitics but are phonological clitics. In particular, there is no empty category related to the clitic in the syntactic representation. That is, Vata is not a NSL.

[10] Italian certainly has a series of subject pronouns that bear stress: *io* 'I', *tu* 'you, sg', *lui* 'he', *lei* 'she', *noi* 'we', *voi* 'you, pl', *loro* 'they'. They may, though need not, co-occur with the verb. One property that they share with stressed pronouns cross-linguistically is that they may not be bound as variables. Thus, the stressed pronoun in (ia) must be construed as referring to a specific person, not as bound by the quantifier *everyone*, and the second conjunct of the example in (ib) may not be interpreted as meaning every student has the self-belief of being smart; that is, the stressed pronoun may not be bound as a variable (cf. Montalbetti 1984):

(i) a. Everyone thinks that HE is intelligent.
b. Every teacher thinks that HE is smart, and every student does too.

[11] The slightly less than perfect status of (50b) is perhaps due to discourse factors. It is not immediately clear if the phonetically NS can pick up the discourse topic as set up by the preceding sentence.

[12] It is worth pointing out that the answers to the questions in (51a) and (52a) containing an object pronoun do not sound as natural. They are either totally unacceptable or inappropriate in the given discourse context:

(i) a. B: *Kàn-wán-le tā
see-finish-ASP 3SG
'He has read it'
b. B: ??Ài tā.
love 3SG
'He loves him.'

[13] It is of course impossible to tell the grammatical gender of the pronoun when it precedes an auxiliary beginning with a vowel, due to the elision of the vowel of the pronoun. For the point here, I assume that the pronouns in (59) have a discourse antecedents that are feminine gender, and hence have the same gender as well.

[14] There is of the question of why HC does not allow phonetically null arguments like Chinese. Although I have no satisfactory answer to this important question at this point, it seems to me that labelling languages like Chinese as discourse-oriented with some topic-prominent properties like discourse-binding of anaphors, topic-comment structures with no gap (Li and Thompson 1981, Huang 1984) does not really address the issue in an adequate manner either. It is not obvious that these properties bear on the issue of null arguments. Even if null arguments are considered to be discourse-bound in the same manner as that of discourse anaphors, it is not clear why the phonetic matrix of the

pronoun should matter. The bearing of topic-comment structures with no gap on null arguments is even more obscure.

[15] Although examples with lexical objects and null subjects are overwhelmingly common, those with a lexical subject are not unattested. Thus, the examples in (i) are found alongside those in (ii) (Bloom, Lightbown and Hood 1975):

(i)	a.	play it	(Eric II)
		eat juice	(Eric III)
		find it	(Eric III)
	b.	eat piece	(Gia II)
		fix dat	(Gia III)
		ride truck	(Gia III)
	c.	touch milk	(Kathryn I)
		untie this	(Kathryn II)
		want go get it	(Kathryn III)
	d.	Pull it	(Peter III)
		turn it	(Peter III)
		Push the button	(Peter III)
(ii)	a.	This rides	(Kathryn I)
		man ride bus	(Kathryn I)
		Kathryn read this	(Kathryn II)
		Kathryn do it	(Kathryn II)
		lamb goes	(Kathryn II)
		Kathryn sit down	(Kathryn II)
		I put this in there	(Kathryn III)
		foot goes over there	(Kathryn III)
	b.	I find it	(Eric III)
		man sit blocks	(Eric III)
		I need that	(Eric III)
	c.	Gia ride bike	(Gia III)
		Mommy work	(Gia III)
		Gia want Daddy	(Gia III)

Most analyses of NS in child language thus appear to have abstracted away from these examples. What is clear, however, is that children have NSs much more often than lexical subjects at stage I. A few exceptions to this observation are possibly in the data of Kathryn, shown in (iia) above. The claim that children permit NS would appear to hold for a very brief period of time, perhaps as a matter of a few weeks.

[16] One might argue, however, that the utterances children make must correspond to semantic or conceptual units, explaining why all the examples in (61b) except *Kathryn get* are unattested.

Along these lines, one could then conceivably explain NS by saying that the combination of a verb and an object is semantically or conceptually a property, but that of a verb and a subject is not. The problem with this view is the co-existence of relative clauses where a subject is relativized and those where an object is relativized:

(i) a. The car that Kathryn got
b. The car that hit Kathryn.

The combination of a subject and a verb in the relative clause in (ia) is conceptually as much a property as that of an object and a verb in (ib). Thus, the reason why expressions like *Kathryn get* are not found in the data cannot be explained by claiming that it is not a conceptual or semantic unit (cf. (ia)) above.

[17] Note that the universality of the constraint in (63), as well as principles like Subjacency or the ECP that I mentioned in the introduction, rests entirely on empirical grounds, whatever they may be. The relevant point here is that whatever is responsible for the contrast in (64) or grammatical differentiations of Subjacency or ECP violations is not learnable.

References

Bayer, Josef. 1984. COMP in Bavarian syntax. Linguistic Review 3: 209-274.

Bloom, Lois, Patsy Lightbown and Lois Hood. 1975. Structure and variation in child language. Chicago, IL: Society for research in child development.

Bouchard, Dennis. 1982. On the content of empty categories. Dordrecht: Foris.

Brown, Roger. 1973. A first language. Cambridge, MA: Harvard University Press.

Chomsky, Noam. 1957. Syntactic structures. The Hague: Mouton.

Chomsky, Noam. 1973. Conditions on transformations. Festschrift for Morris Halle, ed. by Stephen Anderson, Peter Culicover and Paul Kiparsky. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht: Foris.

Chomsky, Noam and Howard Lasnik. 1977. Filters and control. Linguistic Inquiry 8: 425-504.

Clahsen, Harald and Pieter Muysken. 1983. The accessibility of 'Move α' and the acquisition of German word order. Unpublished manuscript.

Emonds, Joseph. 1978. The verbal complex V'-V in French. Linguistic Inquiry 9: 191-231.

Huang, James C-T. 1984. On the distribution and reference of empty pronouns. Linguistic INquiry 15: 531-574.

Hyams, Nina. 1989. The null subject parameter in language acquisition. The null subject parameter, ed. by Osvaldo Jaeggli and Kenneth Safir. Dordrecht: Kluwer.

Hyams, Nina. 1992. A reanalysis of null subjects in child language. Theoretical issues in language acquisition, ed. by Jürgen Weissenborn, Helen Goodluck and Thomas Roeper. Hillsdale, NJ: Lawrence Erlbaum.

Jaeggli, Osvaldo. 1980. Topics on Romance syntax. Dordrecht: Foris.

Jaeggli, Osvaldo and Kenneth Safir (eds). 1989. The null subject parameter. Dordrecht: Kluwer.

Koster, Jan. 1975. Dutch as a SOV language. Linguistic Analysis 1: 111-136.

Koopman, Hilda. 1984. Syntax of verb. Dordrecht: Foris.

Law, Paul. 1993. On null subjects and null arguments. Canadian Journal of Linguistics 38: 1-41.

Li, Charles and Sandra Thompson. 1981. Mandarin Chinese: A functional reference grammar. Berkeley: Univeristy of California Press.

Merke, Ludwig. 1986. Bairische grammatik (corrected version). München: Heimeran Verlag.

Mills, Anne. 1987. The acquisition of German. The cross-linguistic study of language acquisition, ed. by Dan Slobin. Hillsdale, NJ: Lawrence Erlbaum.

Montalbetti, Mario. 1984. After binding. Doctoral dissertation. MIT, Cambridge.

Park, T-Z. 1970. The acquisition of German syntax. Unpublished paper, Psychologisches Institut, University of München.

Pollock, Jean-Yves. 1989. Verb movement, UG and the structure of IP. Linguistic Inquiry 20: 365-425.

Rizzi, Luigi. 1982. Issues in Italian syntax. Dordrecht: Foris.

Roeper, Thomas. 1973. Theoretical implications of word order, topicalization and inflections in German language acquisition. Studies of child language development, ed. by Charles Ferguson and Dan Slobin. New York: Holt, Rinehart and Winston.

Thiersch, Craig. 1978. Topics in German syntax. Doctoral dissertation. MIT, Cambridge.

Zagona, Karen. 1982. Government and proper government of verbal projections. Doctoral dissertation. University of Washington, Seattle.

Some Remarks on Argument Ordering in German-

An Endorsement for a Universal Hierarchy[1]

André Meinunger (FAS)
andre@fas.ag-berlin.mpg.de

1. What is the Basic Word Order (in German)?

It is not very clear what basic word order is supposed to mean, and consequently it is even less clear how it can be defined. According to standard approaches I will assume that arguments of a lexical head are projected within the c- or m- command domain of that head. In case of multiple arguments, these are ordered obeying a hierarchy of thematic roles. Assuming binary branching, thematically higher ranked arguments asymmetrically c-command deeper ranked ones. The thematic hierarchy in turn is a different matter of debate. Almost nobody challenges that the agent argument is located very high in the thematic hierarchy and thus stays furthest away from the deepest head position within the verbal phrase. Concerning the other arguments, and partly even adjuncts, no agreement can be found. One controversial question is the ranking of dative and accusative objects[2]. As for the basic orders it has been claimed that all possible rankings are attested (Höhle (1982), for a reprise cf. Haider (1992)). All possible rankings means: (I) dative is higher than accusative, (II) accusative is higher than dative, and neither ranks over the other or both are mutually exchangeable (III). It is claimed that the instantiation depends on the nature of the verb.

(1)

(I)	abgewöhnen, beibringen, verweigern, zutrauen...	
	wean, administer, deny,	to think somebody is able to
(II)	aussetzen, unterziehen, zuführen	
	expose, submit,	to bring to
(III)	geben, zeigen, empfehlen	
	give, show, recommand	

Indeed, at first glance this division seems to be well motivated. If one gives these verbs to speakers and asks them to build a sentence with them, they will with high probability order the

---

[1] This article is a slightly modified and shortened version of one chapter of my thesis (Meinunger 1995 a). While writing this part I got advice and helpful comments from Markus Steinbach, Ralf Vogel and Ilse Zimmermann. Susan Olsen checked the English of the original version. Thanks to all of them.

[2] I am aware of the fact that the question of whether dative ranks over accusative is not identical to the question of whether goal ranks over theme. However, the questions are related.

arguments in the way the classification predicts. That means that, whereas in sentences with verbs of class I dative objects will precede accusative ones, sentences with class II verbs will show the reverse order. Sentences that contain class III verbs will come with both orders. This is of course not sufficient for the given classification.

2. Difficulties with a Misunderstanding of Focus Projection as a Diagnostic for Basic Word Order

Höhle (1982) takes the intuitions described above only as a point of departure and develops a test to justify the 'different-order-hypothesis' theoretically. He proposes a correlation between basic word order and maximal focus spreading on the one hand, and derived word order and narrow focus on the other. Thus, his claim is that focus projection is possible for base generated structures, but impossible for derived orders. (For the mechanism of focus projection see chapter 4 and chapter 6 and references quoted there.) I, too, assume that this is the right conjecture. However, I think that one has to be very careful in using focus projection as a reliable test. Later I will come back to the reason. But first, let's look at the data.

(2) a. daß Carl_{NOM} die $\text{Lösung}_{\text{ACC}}$ fand (spreading)
that Carl_{NOM} the $\text{solution}_{\text{ACC}}$ found
b. daß die $\text{Lösung}_{\text{ACC}}$ Carl_{NOM} fand (no spreading)

(3) **class I**
a. daß er seiner Frau_{DAT} sein Geld_{ACC} nicht gönnte (spreading)
b. daß er sein Geld_{ACC} seiner Frau_{DAT} nicht gönnte (no spreading)
class II
c. daß er seine $\text{Kinder}_{\text{ACC}}$ ihrem $\text{Einfluß}_{\text{DAT}}$ aussetzte (spreading)
d. daß er ihrem $\text{Einfluß}_{\text{DAT}}$ seine $\text{Kinder}_{\text{ACC}}$ aussetzte (no spreading)
class III
e. daß er seiner Frau_{DAT} sein Geld_{ACC} gegeben hat (spreading)
f. daß er sein Geld_{ACC} seiner Frau_{DAT} gegeben hat (spreading)

(2) is uncontroversial and shows that nominative must precede accusative to make focus projection possible. This fact then is carried over to the spreading possibilities in the double object examples from (3). However, the data here are less clear. Nevertheless, I claim that the mistake lies somewhere else, namely in the misunderstanding of the relation between questions and focus projection in possible answers. It is simply not the case that an answer to a wh-question only consists of the open proposition delivered by the question plus the (exhaustive) instantiation of the open proposition. It is very well possible for the answer to contain more material, for example in order to facilitate storing of new information. What I mean is that the answer to a question of the sort 'What happened?'/ 'What's the matter?' need not necessarily be an all-new sentence. A

structured proposition in form of a categorial statement can also be a possible answer. A sentence like 'Aunt Lisa died' may have different information packagings. It can be a thetic statement, i.e. an all new sentence. In English, telicity of a one-argument clause is achieved by putting the main stress on the head of the argument. In that case the intonation pattern is:

(4) Aunt LIsa died.

Another possibility is the use of the term *aunt Lisa* as an expression for someone about who it is being asserted that she died. In that case, the expression *aunt Lisa* is (more) salient, and the stress goes on the verb. This is the intonation of a categorial statement.

(5) Aunt Lisa DIED.

Nevertheless, (5) is a possible answer to a what-happened-question. There is no necessary identity between the open proposition set by the question and the presupposed material in the answer. Otherwise, what-happened questions would only be allowed in situations where the speakers have no common ground at all, which is a very rare, if not even impossible case. It is true that presupposed material from the question cannot be used as the focus of the corresponding answer.

(6)
A: What happened to aunt Lisa?
B: *Aunt LIsa died.

However, this fact does not imply that everything contained in the answer which does not belong to the question must be focus or new information. Let me give another example:

(7)
A: (Why is Mary angry with Paul?) What did he do?
B: The day before yesterday, he slept with Marianne.

This dialog does not have the slightest flavor of oddness. The question asks for some action of Paul that causes Mary's anger. The answer to that is his sex with Marianne, encoded in the VP [$_{VP}$ slept with Marianne]. For some reason, B decided to be a bit more explicit and gave the time of the action. The sentence initial position of the temporal adjunct, together with an intonation pattern that puts little weight on it, but more on *Marianne*, indicates that the temporal information encoded in 'the day before yesterday' is a (non contrastive) topic. Thus we have two constituents that are not in focus, but only one of them is delivered by the linguistic context, namely *Paul* = *he*. The other one, which contains a deictic expression, can still be easily accommodated. Thus, we see that it is not completely conclusive to consider question-answer pairs as a reliable diagnostics for focus projection. Given a question and a felicitous answer, one cannot claim that all the material which is contained in the answer which is missing in the question must be new

information and hence in the range of focus projection. So, why this long discussion? (3 f.) claimed that focus projection is possible where accusative precedes dative. However, focus projection was understood there as question-answer felicity. Thus, (3 f.) is regarded as a possible answer to a question 'Was hat er gemacht?' (What did he do?). With the wrong theory about the focus projection test outlined above, this then leads to the conclusion that every constituent (including the verb), but *er*, must be focus. This, however, is not the case. I shall claim that the accusative argument in this case must be discourse-related and focus does not spread over it. I argue that the focus projection capacities of class I verbs are not different from class III verbs. And, therefore, the contrast between (3 b.) and (3 f.) seems to me to be spurious.

3. The Strict Word Order Hypothesis

Now, I want to show that there is a clear and more reliable test for showing that dative is ranked higher than accusative (for both class I and class III verbs). According to the work of Adger (1993) and earlier work of mine (Meinunger 1993, 1995a), which is based on Diesing's Mapping Hypothesis (1992); I will argue that linguistic material which is being introduced into the discourse frame stays in its base generated position. Thus we have to examine the order in which new material organizes. Since DPs containing ordinary nouns are not conclusive, we have to look for something else. Ordinary DPs are not conclusive because even indefinite DPs can easily obtain a presuppositional reading. However, with unstressed indefinite articles they are almost perfect indicators of what we are looking for. I think the best way of showing the linear order of arguments is to use indefinite pronouns that cannot or can hardly have a presuppositional reading. Such elements are (unstressed) *jemand, niemand, etwas, nichts, einer* (somebody, nobody, something, nothing, a/one) and their phonologically reduced forms *'was, 'ner*. When one constructs sentences with these pronouns, one sees that verbs of class I behave exactly as verbs of class III in that the dative object must precede the accusative one.

(8) class III

a. weil er jemandem (et)was {gezeigt / gegeben / empfohlen / erklärt / geschickt...} hat

since he somebody$_{DAT}$ something$_{(ACC)}$ {shown, given, recommanded, explained...} has

b. *weil er (et)was jemandem {gezeigt / gegeben / empfohlen / erklärt / geschickt...} hat (reverse order, i.e. ACC > DAT)

The same is of course the case with class I verbs, which is already predicted by Höhle's theory.

(9) class I

a. weil er jemandem (et)was {abgewöhnt / verweigert / beigebracht / zugetraut / verübelt...} hat

since he somebody$_{DAT}$ something$_{(ACC)}$ {weaned, denied, tought, blamed...} has

b. *weil er (et)was jemandem {abgewöhnt / verweigert / beigebracht / zugetraut / verübelt...} hat (reverse order, i.e. ACC > DAT)

As mentioned above, (unstressed) indefinite NPs behave similarly. However, things are more complicated here. The order ACC > DAT itself is not ungrammatical, and the unmarked stress always falls on the verb adjacent argument. In this sense (10/11) a. and (10/11) b. are equally good. What distinguishes (10/11) a. from (10/11) b. is that the former may serve for focus projection whereas the latter cannot. However, as I have argued, the focus spreading test is not appropriate. So I propose that (10/11) b. get starred when the intended reading is one where the indefinite objects are introduced into the discourse frame.

(10) class III

a. weil er einer Frau eine Rose geschenkt hat
since he a woman$_{DAT}$ a rose$_{ACC}$ given has

b. *weil er eine Rose einer Frau geschenkt hat

(11) class I

a. weil er einem Freund ein Lied beigebracht hat
since he a friend$_{DAT}$ a song$_{ACC}$ tought has

b. *weil er ein Lied einem Freund beigebracht hat

I hope to have shown that class I and class III are not different with respect to argument projection and that we therefore should not speak of two different classes.
Let us now turn to class II. If we apply our test to the verbs of this class, we will find out that the base order is ACC > DAT. However, I have to admit that the ordering test with indefinite pronouns does not work very well here.

(12) class II
  a. weil ich auf der Party niemand(en) jemandem vorgestellt habe
     since I at the party $\text{nobody}_{(ACC)}$ somebody_{DAT} presented have
  b. *[??]weil ich auf der Party niemandem jemand(en) vorgestellt habe

Yet, we may have one argument as a full DP. The claim is that the relevant indefinite pronouns must be in their base position. Thus it does no harm if the linearly following argument is a structured DP and the indefinite pronoun precedes it. The data become uncontroversial again.

(13) a. weil er jemanden einer schweren Prüfung unterzog
     since he somebody_{ACC} [a difficult exam]$_{DAT}$ submitted
  b. *weil er einer schweren Prüfung jemanden unterzog
(14) a. weil sie niemanden einer großen Gefahr aussetzen würde
     since she nobody_{ACC} [a big danger]$_{DAT}$ expose would
  b. *weil sie einer großen Gefahr niemanden aussetzen würde

Thus it seems that there are not three classes, but there may at least two: DAT > ACC and ACC > DAT. Nevertheless I would like to maintain the claim that DAT > ACC holds underlyingly. The ACC > DAT order can be seen as an epiphenomenon similar to what is going on with the so-called ill-behaved experiencer verbs (for a discussion of this story see Grimshaw 1990 and Belletti and Rizzi 1988, Pesetsky 1990 and Meinunger 1995a).

4. Some Similarities with Experiencer Verb Constructions

Generally, arguments should be projected uniformly (UTAH: Baker (1988)) and according to Grimshaw's hierarchy given in (15) (Grimshaw 1990). One class of experiencer verbs - the *fear* class (or Belletti and Rizzi's *temere* class (1988)) - is well-behaved. That means that the experiencer, located higher in the hierarchy, becomes the subject of the sentence; the theme, located deeper, becomes the object.

(15) (Agent (Experiencer (Goal / Source/ Location (Theme))))
(16) Lohengrin fears Elsa's question.
(17) Artemis likes Kayne's theory.

However, there is the class of ill-behaved verbs - the *frighten* class (Belletti and Rizzi's *preoccupare* class)

(18) Alberich frightens the Nibelungs.

Here the experiencer appears as a postverbal object, and the theme occupies the subject position. Grimshaw however presents a way out of the dilemma. Her proposal is that there is not only one scale of hierarchy but more, at least two. She shows that the ill-behaved verbs have something to them which the other class lacks. There is a causative element involved such that (19) can be paraphrased by:

(19) Alberich causes the Nibelungs to experience fear.

Then she states that the causal structure of a predicate also defines a hierarchy, just as the thematic structure does, a hierarchy in which the cause argument is most prominent:

(20) (cause (....))

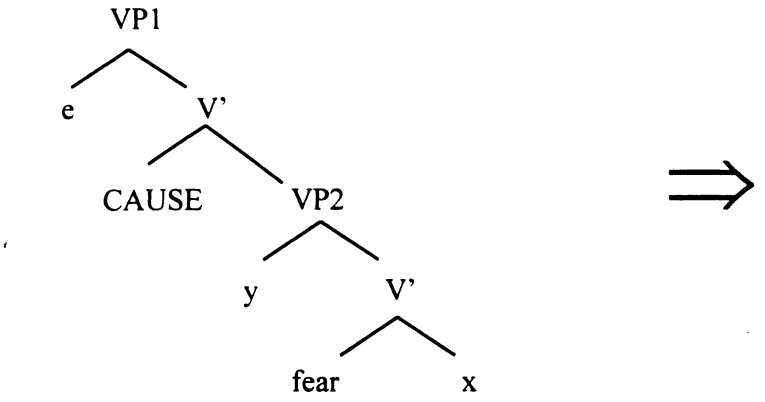
She claims that the causativity hierarchy overrides the other one(s) and imposes a structure where the causer is the most prominent argument. Another possible, and actually similar way of capturing the difference between the two classes is more along the lines of Pesetsky (1990). In his theory too, *frighten* is not equal in meaning to *fear* with the theta-roles in the reverse order. The difference lies in the additional causative component which the well-behaved class lacks, but the ill-behaved class exhibits. This can be represented in the following representation:

(21) a. like /fear: $\lambda x \lambda y$ [x E y]

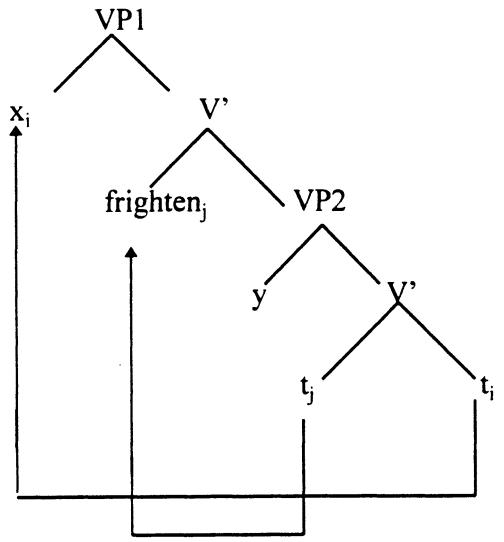
b. please/ frighten: $\lambda x \lambda y$ [y CAUSE [x E y]]

If this notation, taken from Haider (1992), is translated into a syntactic tree, we get a specifier position where the agent is licensed in the topmost argument position. Instead of making the lambda prefix unselectively bind two variables, we can handle the difference syntactically by assuming movement (or another position dependency):

(22)



(23)



Thus, similarly, to 'GIVE' = CAUSE + POSS, one might consider 'FRIGHTEN' as CAUSE + 'FEAR'.

5. The DAT > ACC > DAT / PP Asymmetry

Now, I would like to claim that this kind of argument (position) manipulation can be fruitfully carried over to the bitransitive verb asymmetry. It has been observed that (in German) there seems to exist a tendency that when the non-theme object of bitransitive verb is +animate or +human, it is realized as a dative object (24 a), (25a). On the other hand, when it is not animate or human, it is likely to be expressed in a directional PP (24 b), (25 b) (see Kaufmann (1993) among others). Another difference that Kaufmann overlooks or intentionally withholds is the fact that in the animate case the dative object appears preferably before the accusative object; in the inanimate case, the PP must appear after the accusative object.

(24) a. Sie schickte ihrer Tante ein Buch.
she sent [her aunt]$_{DAT}$ [a book]$_{ACC}$

b. Sie schickte das Buch **an** die Bibliothek.
she sent the bookACC to the library

(25) a. Sie brachte ihrem Vater einen Kuchen.
she brought [her father]$_{DAT}$ [a cake]$_{ACC}$

b. Sie brachte einen Kuchen **ins** Büro.
she brought a cake into+the office

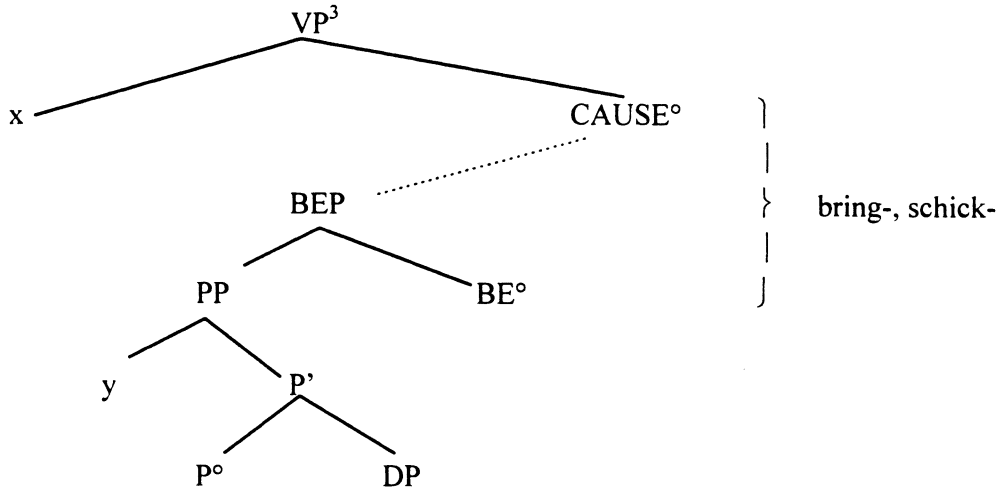
I would like to claim that it is not primarily the interaction of animacy or humanness, but that the difference is mediated through a distinction concerning the interaction of the atomic predicates. In the beginning of this chapter, I assumed POSS(ESSION) to be an atomic predicate. Now, I will

argue that it is of great advantage to analyze it as a derived one. Therefore, I have to assume a view of argument structure similar to that found in Speas (1990) and of have-be alternation much like in Kayne (1993). My claim is that many bitransitive verbs either refer to a relation between a theme and the theme's location, or express a process (or a state) in which the dative argument possesses / comes to possess the theme. I furthermore claim that the former relation (location) is underlying and the latter (possession), which contains more information, is derived. As for the constructions with a locational (secondary) predication, I assume that the lexically decomposed structure looks like:

(26) [x CAUSE [... BE [y [IN/ AT/ ON z]]]]

Thus for *bringen* (to bring) and *schicken* (to send) with a prepositional complement, we would have a tree structure like in (27).

(27)



This is the representation for sentences like (24 a) and (25 b). Now comes Kayne's idea (which goes back to earlier work by traditional grammarians). For him *have* is derived from a preposition which has incorporated into *be*. Transferred into a syntactic theory of lexical head decomposition, this means something like the deepmost locational P° incorporates into the primitive BE. This process results in the POSSESSION relation. Exactly as with the experiencer verbs, the head movement within the VP triggers the movement of an argument. In our case here, it is the former complement of the preposition which becomes the specifier of POSS. (The overt preposition disappears and a possession relationship comes across. See also Kayne.)[4]

[3] For the shake of harmony I will assume that in German also the VP internal atomic predicates project head finally. This makes the trees appear somewhat less familiar. Nevertheless I think that this is not an insurmountable problem for the reader.

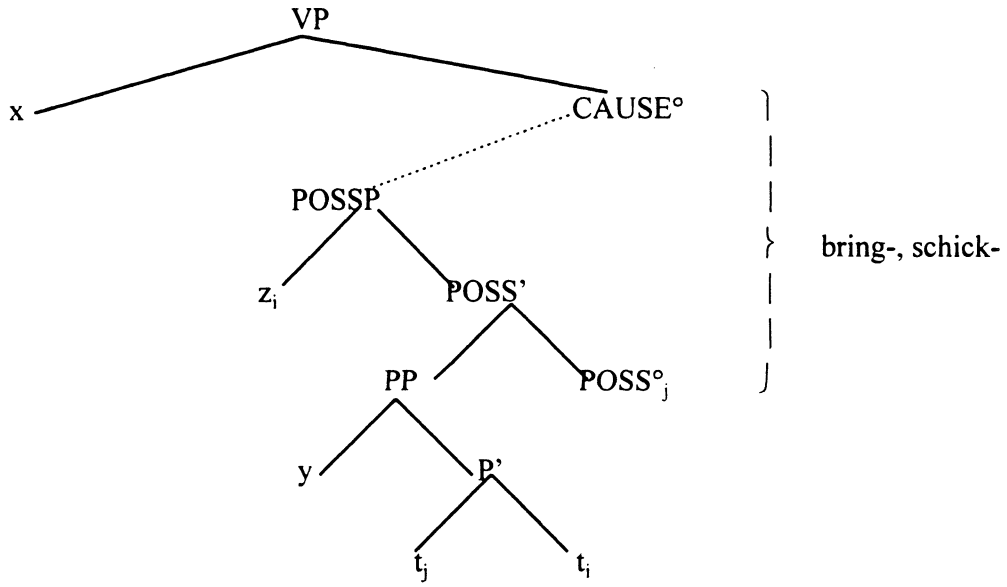
[4] Interestingly there is a fact that could be used as additional evidence for the analysis. The fact is the relation between dative Case and possession. It is well known that there is no one-to-one correspondence between morphological Cases on the one hand and thematic roles on the other. However, it is as well known that both are more than only loosely related. At any case, in many languages that have morphological dative, this case is often assigned to the possessor in a process similar to the one discussed here. For example in Hungarian (discussed in Szabolcsi (1981) and re-presented in Kayne (1993)), the possessive construction consists of a copula (BE) and a

Semantically, that means that it becomes the possessor. Thus, my claim is that the possession relation is not a semantic primitive, but that it is a result of verb phrase internal changes. Thus:

(28) [x CAUSE [e ... BE [y [IN/ AT/ ON z]]] ⇒

[x CAUSE ...[z [POSS y]]]

(29)



---

single DP containing the possessor and the possessee. When the whole DP is definite, the possessor can remain is situ carrying nominative Case, but in other cases it must or can move to the left to some specifier position where it gets assigned dative Case. Something similar also happens in my non-standard German. A DP expressing some possessive relation may come in two variants:

(i) der Garten von der Ingrid — having the structure $[_{DP}$ D° $[_{NP}$ N° $[_{PP}$ P° POSSESSOR]]]
the garden of the Ingrid

(i) somehow corresponds to the base variant in (32) involving a PP. The other, more natural, variant is (ii) where the possessor has been moved to some specifier position where it appears in dative Case. The D° element shows agreement with the phrase in the specifier position with respect to gender. Here, the dative's function is to mark the possession relation:

(ii) meiner Mutter ihr Garten — having the structure $[_{DP}$ $POSSESSOR_{DAT}$ [D° $[_{NP}$ N° t]]]
my_{DAT} mother her garden

Also sentences that refer to possession relations make use of dative Case as possessor marker. In my variety of German, it is very common to express possession by a copula (BE) with two satellite DPs (I don't want to call them arguments). If the possessee is definite, it is likely to appear in nominative case. The possessor then carries dative Case:

(iii) Dieser Garten ist meiner Mutter.
this $garden_{NOM}$ is my $mother_{DAT}$

Thus, the link of POSS and a dative DP in its specifier seems to be motivated by an akin, but different construction across languages.

This analysis is corroborated by the following facts. The alluded tendency to dativize a +animate /+human DP is only an epiphenomenon. There is nothing strange about having an +animate/ +human DP within a PP construction.

(30) weil ich ein Buch zu meinem Vater gebracht habe
since I a book to my father brought have
(31) weil ich das Farhrrad zu meiner Tante geschickt hatte
since I the bicycle to my aunt sent had

However, the meaning is different from the corresponding DAT > ACC construction. (30) and (31) do not tell us anything about possession. (30), for example, expresses that I brought some book to my father's residence. My father needn't even know of the book. In (31), there is not the slightest hint that the aunt becomes the possessor. On the other hand, the corresponding DAT > ACC constructions make a POSS reading much more likely.

(32) weil ich meinem Vater ein Buch gebracht habe
since I my father a book brought have
(33) weil ich meiner Tante das Fahrrad geschickt habe
since I my aunt the bicycle sent have

(32) strongly suggests that now my father owns the book. However, my claim is not that POSS necessarily expresses ownership. It merely means that someone is in the (perhaps temporary) possession of something. For example, (33) does not necessarily mean that the ownership of the bicycle changes from mine or someone else's to my aunt's. However, the sentence says that my aunt is somehow in conscious possession of the bike. This is not the case with the PP construction in (31). That sentence might describe a situation where I have sent a / my bike to my aunt's address in Paris. However, for the time being my aunt doesn't live there and I know that. The only reason for my sending action was that I want to go to Paris and did not want to take the bike with me in the train. Since I don't trust left-luggage offices, I wanted to pick up my bike at my aunt's place rather than at the station. In such a case, my aunt need not know anything about that. (33) cannot be used to describe such a situation.
This theory is also partly corroborated by the fact that the DAT > ACC vs. ACC > PP alternation is not freely allowed. It is not the case that to every DAT > ACC order there is a corresponding ACC > PP order. This possibility seems to me to be limited to the case with verbs where the non-accusative object can receive a locative reading. For verbs, where this is not possible, the ACC > PP construction sounds awkward.

(33a) [ok]weil ich es meiner Mutter {gezeigt / empfohlen / erklärt / zugetraut / verübelt...} habe[5]

'since I showed, recomanded, explained...it to my mother'

(33b) *weil ich es **an** meine Mutter /**zu** meiner Mutter {gezeigt / empfohlen / erklärt / zugetraut / verübelt...} habe

Now the reader might wonder why I have spent so much effort on the DAT > ACC vs. ACC > PP alternation. The answer lies in the DAT > ACC vs. ACC > DAT problem which was alluded to above, but for which a solution has still not yet been given. The following discussion revives this problem.

Above, I have shown that there is no DAT > ACC vs. DAT > ACC & ACC > DAT distinction, i.e. class I and class III collapse. The long discussion about the DAT > ACC vs. ACC > PP distinction was intended to prepare for the next verb class collapse; namely, I shall claim that the 'ill-behaved' class II verbs are hidden ACC > PP verbs. To put it in other words, the dative argument of ACC > DAT verbs (class II) is actually (the remnant of) a PP. The argumentation will not be very semantic. The only thing I want to mention is that also Müller (1993, p. 204, fn.3) admits that the dative arguments of verb II class verbs do not act as goals. I want to go further and say that the datives denote something local. Let us consider the verbs of class II. Haider (1992) gives the following examples:

(34)

aussetzen	to expose so to sth
ausliefern	to extradite
entziehen (!)	to take away from
unterziehen	to submit
unterwerfen	to subject
zuführen	to bring to

[5] Now, my argumentation could be used against me. What I did was dealing with the opposition possession vs. location. Now, I am using the lack of a locational reading with the given verbs as an argument for the lack of the ACC > PP construction. So far, so good. However, if the matters were that simple, my narrow minded opposition predicts that with the given verbs, we only get a reading where POSS plays a role. This, however, is not the case. Here we do not get any (sub)relation which could be identified as POSSESSION. So what I have to say is that my theory of location to possession change does not explain every DAT > ACC ordering. This, however, has never been my claim. What I claim is only that it covers a considerable part.

We can add:

vorstellen	to introduce
vorziehen	to prefer

All these verbs, with one exception, can be morphologically decomposed into a verbal stem and a local preposition (underlined). The only exception *entziehen* can easily be shown to be misplaced here. Even people who accept the Höhle-Haider test of focus projection admit that the order is dative > accusative[6]. Thus my claim is that ACC > DAT verbs are ACC > PP verbs where the (local) preposition has been incorporated into the verb. A clear case where this incorporation can be shown by a related construction is the acceptability of both (31) and (32) with the verb *(zu) führen.*

(35) weil sie ein neues Opfer **zu** ihrem Medizinmann _geführt haben
since they a new victim$_{ACC}$ to their wizard _lead have

(36) weil sie ∅ ihrem Medizinmann ein neues Opfer ***(zu)**geführt haben
since they [∅ their wizard]$_{DAT}$ a new victim$_{ACC}$ **to**lead have

---

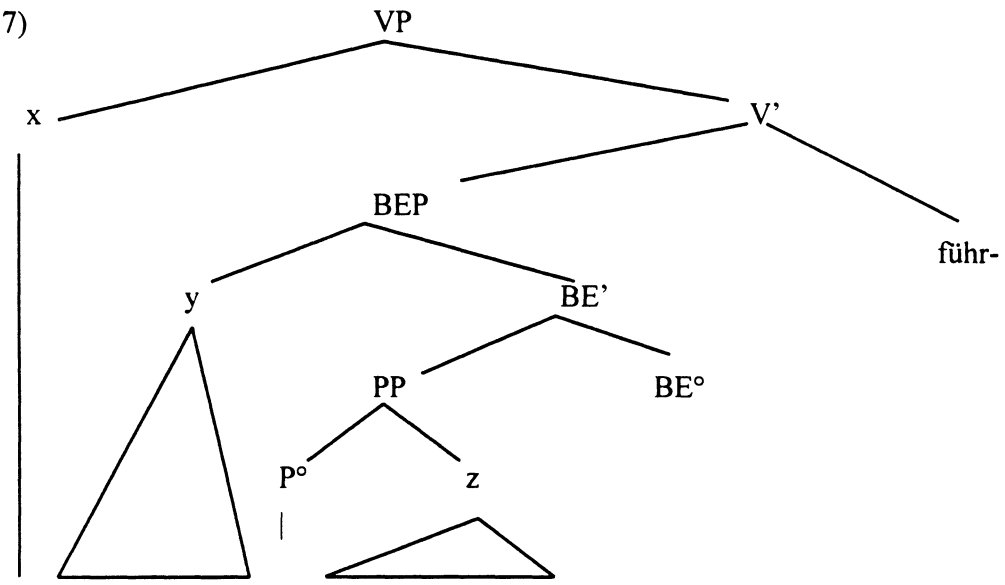
6 A: Und was hast du dann gemacht? A: And what did you do then?

B: Dann habe ich dem Wasser die Giftstoffe B: Then I depoisened the water.
then have I [the water]$_{DAT}$ [the poisenous substances]$_{ACC}$
entzogen
away-taken

Also my test of the ordering of indefinite pronouns / or DP shows that *entziehen* is an ordinary DAT > ACC verb:

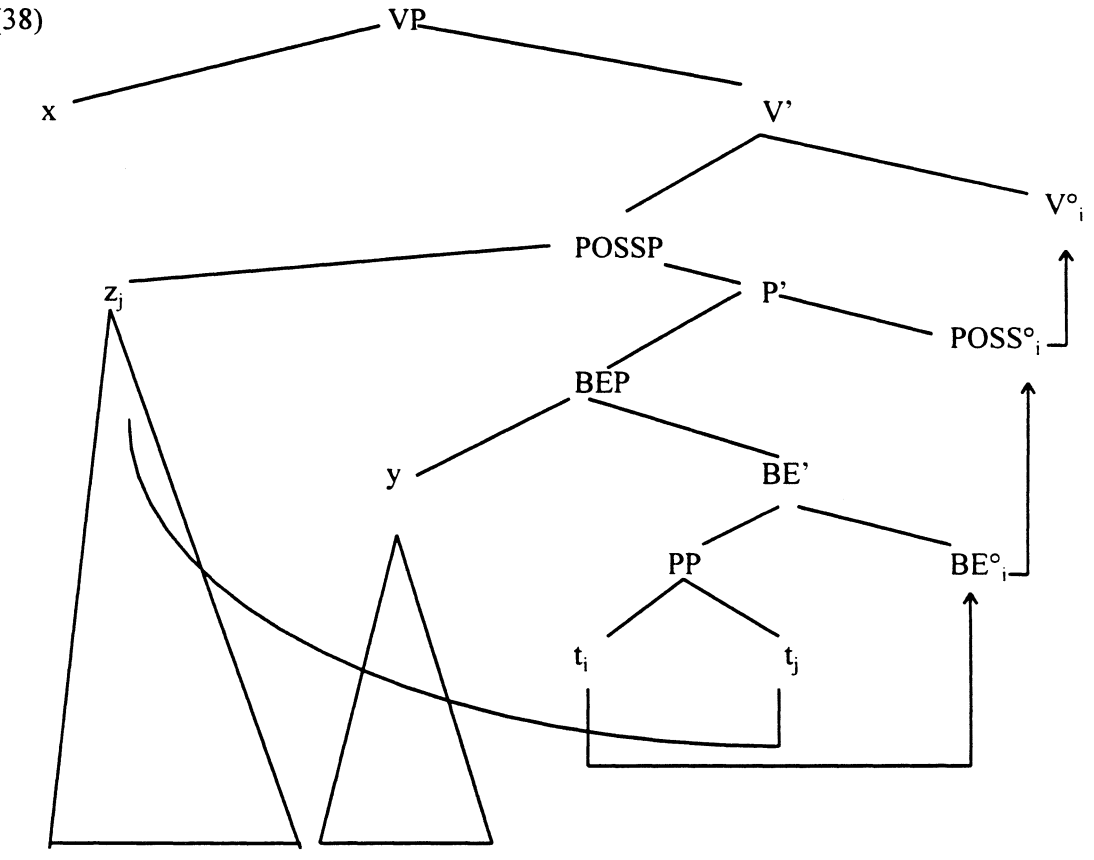
(i) ^{ok}weil ich jemandem etwas entzogen habe
since I someone$_{DAT}$ something$_{ACC}$ away-taken have
(ii) *weil ich etwas jemandem entzogen habe (reverse order)

(37)



sie ein neues Opfer zu ihrem Medizinmann geführt haben

(38)



sie ihrem Medizinmann ein neues Opfer zugeführt haben

6. Summary

Within the (German) VP, the arguments are projected according to a universal hierarchy of thematic roles. I have shown that the claim that German displays several base orders (DAT > ACC, ACC > DAT, ACC < / > DAT) cannot be maintained. The conclusion that there are different base-orders is the result of a misunderstanding of focus projection on the on hand, and the overlooking of some semantic facts with the DAT > ACC, ACC > PP alternation on the other. A closer look at the facts reveals that true dative objects generally precede and therefore c-command accusative arguments. There are no verbs which allow for both orders simultaneously. If dative objects appear to be closer to the verb than accusatives, the datives at issue are no true datives, but hidden PPs. The semantic prove comes from a lexical decomposition of the meaning. Higher ranked datives denote goal arguments, deeper ranked ones, which are actually PPs, denote locations or directions. The syntactic evidence comes from the morphological shape of the relevant class of verbs. All verbs that project an ACC > DAT VP, are particle verbs that consist of a verbal root and a prefixed (locational) preposition. I argue that this word-internal structure is the result of the incorporation of the preposition leaving the former prepositional complement surface as a(n apparent) dative argument. The internal structure of verbs projecting a goal argument is the result of an abstract incorporation of a locative/directional preposition into the semantic primitive BE. This process - similar to Kayne's *have-be* alternation (Kayne 1993) - creates a complex part of meaning denoting a possession relation: POSS, which hosts the derived goal argument in its specifier.

The conclusion of all observations is that also the German VP projects according to a familiar hierarchy proposed by many linguists for many languages:

(39) $[_{VP}$ SU [IO [DO [PP verb(]v]v]v])

References

Adger, D. (1993) *Functional Heads and Interpretation.* PhD thesis. University of Edinburgh.

Baker, M. (1988) *Incorporation. A Theory of Grammatical Function Changing.* Chicago, London: The University of Chicago Press.

Chomsky, N. (1992) A Minimalist Program for Linguistic Theory, *MIT Occational Papers 1.*

Diesing, M. (1992), *Indefinits.* LI Monographs.

Grimshaw, J. (1990) *Argument Structure.* MIT Press, Cambridge, Massachusetts.

Haider, H. (1992) *Branching and Discharge*. Universtity of Stuttgart, Progress Report 23, Arbeitspapiere des Sonderforschungsbereiches 340.

Haider, H. (1993) *Deutsche Syntax - Generativ*. Vorstudien zur Theorie einer projektiven Grammatik. Gunter Narr Verlag Tübingen.

Höhle, T. (1982) Explikationen für „normale Betonung" und „normale Wortstellung". In: W. Abraham (ed.) *Satzglieder im Deutschen*. Tübingen: Niemeyer. 75-153.

Jackendoff, R. (1972) *Semantic interpretation in generative grammar*. Cambridge: MIT Press.

Kaufmann, I. (1993) *Die Kombinatorik lokaler Verben und prädkativer Komplemente: Konzeptuelle Grundlagen semantischer Dekompositionsstrukturen*. Dissertation. Universität Düsseldorf.

Kayne, R. (1984) *Connectedness and Binary Branching*. Dortrecht: Foris.

Kayne, R. (1993) *Towards a Modular Theory of Auxiliary Selection*. Ms. CUNY, to appear in Studia Linguistica.

Larson, R. (1988) On the Double Object Construction. In: *LI 19*, 235-289.

Lenerz, J. (1977) *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Narr.

Meinunger, A. (1993) *Case configuration and referentiality*. To appear in the proceedings of CONsole II.

Meinunger, A. (1995a) *Discourse Dependent DP (De-) Placement*. Doctoral Dissertation, Universität Potsdam.

Meinunger, A. (1995b) *L'alternation DAT>ACC, ACC>PP en allemend et l'hierarchie argumentale/casuelle universelle*. Ms. FAS Berlin.

Müller, G. (1993) *On Deriving Movement Type Asymmetries*. Doctoral dissertation. SfS-Report-05-93. Tübingen.

Pesetsky, D. (1990) *Experiencer Predicates and Universal Alignment Principles*. Ms, MIT.

Schmidt, C.M. (1994b) *Zur syntaktischen Position indirekter Objekte im Deutschen: Pollockische Sätze oder Larsonische VPs?* In: Haftka, B. (ed.) Was determiniert die Wortstellunfgs-variation. Westdeutscher Verlag. 197-218.

Speas, M. (1990) *Phrase Structure in Natural Language*. Studies in Natural Language and Linguistic Theory.

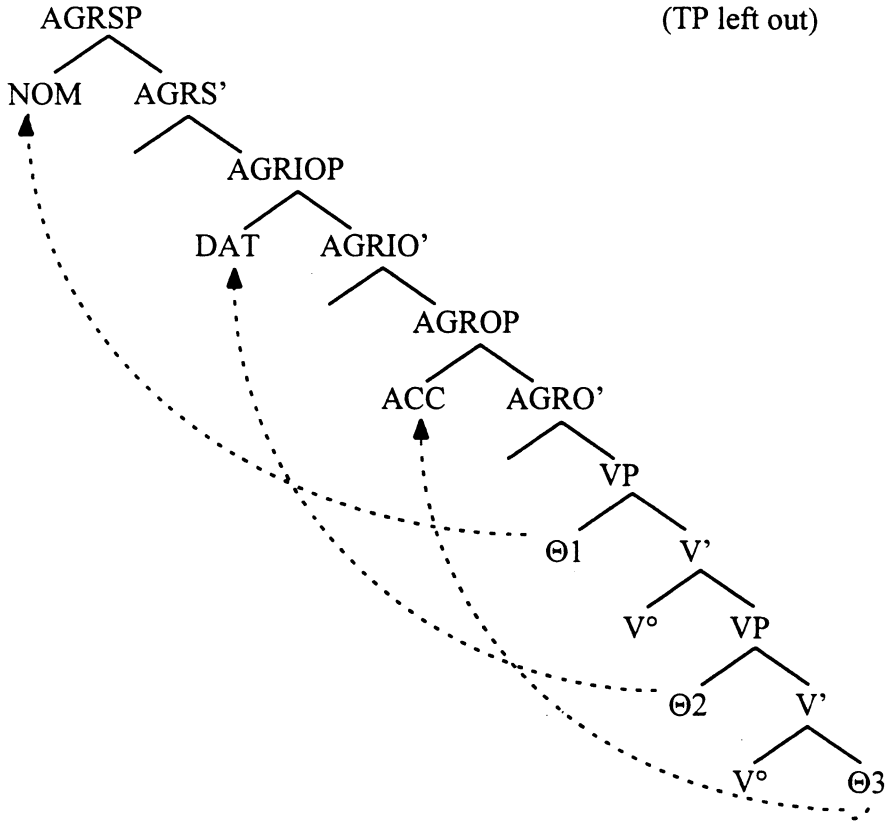
Szabolcsi, A. (1981) The semantics of Topic-Focus Articulation. In T.M.V. Jannsen & M.B.J. Stockhoj, (eds.) *Formal methods in the study of language*. Amsterdam: Mathematisch Centrum. 513-540.

On the (Absence of a) Base Position for Dative Objects in German*

Ralf Vogel, Berlin-Brandenburgische Akademie der Wissenschaften, Berlin
Markus Steinbach, Max-Planck-Gesellschaft, Berlin

The topic of this paper is the dative object case in German and the constraints that govern word order in the so-called middlefield of German clauses. In recent generative literature, the dative has very often been treated as structural case like accusative and nominative (cf. Sabel 1995, Fanselow 1995a/b, Gallmann 1992, Baker 1988, Larson 1988/1990, Müller 1993, Meinunger 1995 a.o.). The structure in (1.) is a possible realization of this idea within the minimalist program (cf. Chomsky 1995). The functional projection for the dative case is located between AGRSP and AGROP. Overt movement of all arguments (it is mostly assumed for German that the N-features have to be checked overtly) leads to the linear surface word order NOM > DAT > ACC (where x > y means that x precedes y).

1. (TP left out)



Such a solution makes clear predictions about the syntactic behavior of dative objects:

a) uniform unmarked order of the arguments (NOM > DAT > ACC)
b) equal syntactic behavior of dative and accusative objects
c) equal mode of thematic interpretation for the 3 arguments

---

* A lot of friends and colleagues have influenced the actual version of this paper, first of all the audiences at the GGS workshop in Jena in May 1995, at the Olomouc Central European Summer School in Generative Linguistics 1995 and Bierwisch, Chris Wilder, Hans-Martin Gärtner, Marie-Christine Erb, Gereon Müller, Paul Law, Ilse Zimmermann, Tolja Strigin, Kai Alter, Gerhard Jäger, Uwe Junghanns, Marcel den Dikken, Joachim Sabel, Anna Cardinaletti, Andre Meinunger, Werner Abraham, Suzan Hahnemann and everyone we tortured with masses of cryptic data. The authors thank themselves and blame each other for the mistakes.

As will be outlined in this paper, each of these three claims about dative objects has to face strong counter-evidence:

ad a) the unmarked order of German dative and accusative objects varies.
ad b) dative objects, in contrast to accusatives, cannot serve as A-binders and are extraction islands.
ad c) 'free' dative objects in German have a thematic interpretation that is independent from the verb – this never occurs with nominative and accusative.

The conclusion we will draw is therefore this the treatment of German dative objects illustrated in (1.) is not adequate – while it appears to be correct for nominative and accusative.

Our own proposal will be that German dative objects, showing clear A'-properties, have to treated as *syntactic adjuncts*. They surface, where they are inserted and do not need to undergo movement. They are only *semantic* arguments. Their thematic interpretation does not follow from theta-role assignment in the traditional sense, but from an interpretational rule that is connected with the dative object case itself.**

A. The 'base position' of German dative objects

We first want to take a closer look on the possible base positions of German dative objects. German clauses seem to show free constituent order in the middlefield (cf. 2.). On the other hand, only sentence (2.a) is neutral with respect to focus. It is the best variant in a neutral context; it can have global focus (cf. Féry 1993) or maximal focus spreading (cf. Höhle 1982; Uhmann 1987, Stechow/Uhmann 1986). Therefore, it is an optimal answer to a question like "what happened?" and also to questions on any of the constituents. (2.a) is claimed to be the *unmarked word order*. It is also the sentence that patterns most naturally with normal intonation. The only DP-movement that might have taken place in this case, is movement to case positions.

2. a. *Es hat ein Junge einem Mädchen ein Buch gegeben*
   It has a boy-NOM a girl-DAT a book-ACC given
   b. *Es hat ein Junge ein Buch einem Mädchen gegeben*
   c. *Es hat ein Buch ein Junge einem Mädchen gegeben*
   d. *Es hat ein Buch einem Mädchen ein Junge gegeben*
   e. *Es hat einem Mädchen ein Junge ein Buch gegeben*
   f. *Es hat einem Mädchen ein Buch ein Junge gegeben*

One might be sceptical about identifying unmarked word orders with basic or 'normal linear' orders. Base positions of arguments are standardly assumed to be theta and/or case positions, fixed by some theory of linking and case assignment. Nonetheless, this theory should predict empirical effects, such that it can be verified. Höhle (1982:126) observes, that any non-normal word order blocks focus projection. Unmarked argument orders with maximal focus spreading seem therefore to be a very natural and plausible candidate for

** Our approach has thus two parts: a syntactic and a semantic one. This paper focuses on the syntactic story (nevertheless the semantics is an essential part of our explanation; see section D and Vogel/Steinbach (in prep.)).

basic orders. This assumption goes along with theories that correlate DP scrambling with narrow focus effects: the marked orders are derived from the unmarked ones.

There are several intervening factors that have to be eliminated, if one tries to find the unmarked order. Lenerz (1977) has given a list of five constraints for the optimal order of the constituents:

> "a. Theme/Rheme Condition: the theme tends to precede the rheme; b. Definiteness Condition: definite tends to precede indefinite; c. Law of Growing Constituents (*Gesetz der wachsenden Glieder*, following Behaghel 1932): heavier constituents tend to follow lighter ones: d. Sentence Bracket Condition: the tendency, not to end a sentence on a light constituent if the sentence bracket is open, i.e. if the clause does not end with a verb; e. Subject/Agent Condition: subject/agent tends to precede other constituents." (Lenerz 1977:63 and 97ff; translation taken from Cooper 1994:19)

Zubin/Köpke (1985) elaborated a pragmatic account of the interaction of these constraints to explain when subject > object surface order occures, and when it is inverted. The mechanism in question is suggested to be "an instance of general cognitive problem solving rather than a real mechanism specific to language...[It lies] outside the domain of real structures of a competence grammar". (ibid: 94) What Zubin/Köpke propose is a "polycausal linearisation mechanism which operates with the 'weights' of individual factors: it calculates the cumulative weight of factors favoring S–O in an individual instance and compares this with the cumulative weight of factors favoring O–S in a competition model...The heavier side wins, and that order is produced." (ibid: 93f) Cooper (1994) concludes, with this in mind and the fact that markedness is different from grammaticality, that the word order in the middle-field of clauses in Zurich German - an Alemannic dialect spoken in Zurich/Switzerland - is governed pragmatically in toto, and that syntax has nothing more to say about it than providing the structural possibility for it. The easiest way to do this is postulating a 'flat strucure' and this is what Cooper does (with the exception of the subject position). Fanselow (1995b) comes to the same solution for Standard German with respect to the rules governing word order, but he implements non-configurationality into the bare phrase structure theory of Chomsky (1995). He proposes that the order of the arguments can be left open before Spell-Out, because case has to be checked at LF, and this will succeed, whatever the surface order of the arguments is. This ability is supposed to be the property that differentiates configurational and non-configurational languages.

What Fanselow lacks to show, is, though, that syntax really has *no* impact on surface orders. To do this, one would have to neutralize the pragmatic factors mentioned above as far as possible, and see whether restrictions on word order can be detected under these circumstances. This aim is the reason for the very special format we use for our example sentences throughout this paper: thetic sentences with indefinite DPs.

Still, contextual implications cannot completely be factored out. But they can be controlled and the unmarked order in the sense of Höhle (1982) or Stechow/Uhmann (1986) comes out quite clearly. We agree with Meinunger (1995) that the given-new distinction (equals Lenerz theme/rheme condition) has to be taken as seriously as the optimal answer criterion in detecting basic orders. To give an example (the one Meinunger uses), Haider (1992)

claims that the verb *geben,* 'give', allows for two unmarked orders of the objects and he illustrates this with the following paradigm:

3. a. *er hat seiner Frau sein Geld gegeben*
   he has his wife-DAT his money-ACC given
   b. *er hat sein Geld seiner Frau gegeben*
   he has his money-ACC his wife-DAT given

Meinunger correctly points out that (3.a) can have global focus, while in (3.b) the accusative object is necessarily discourse related and focus does not spread over it. Meinunger's proposal is therefore that *material that is introduced into the discourse frame stays in its base position.* We take this as an additional criterion for the detection of unmarked orders.

The focus of our interest here, as should have become clear, are judgments of markedness, rather than grammaticality.[1] Markedness surely is a pragmatic phenomenon, but we cannot be certain, up to now, that syntax has no impact on it. On the contrary: researchers agree, for instance, that German is a language that allows for scrambling. Scrambling within a clause might be reducible to non-configurationality along the lines of Cooper (1994) or Fanselow (1995b), but scrambling across clause boundaries certainly not – as well as extraction out of DPs, which is also possible and structurally equivalent, unless one wants to give up explanatory syntax as such. But these constructions clearly are cases of marked sentences, where the markedness is induced by a transformational process on the syntactic structure: either by movement, or, if one wants to analyse them non-derivationally, by reconstruction at LF. The marked form is syntactically derived from (or reconstructed into) the unmarked one. The contrast between the marked and the unmarked form in this case is a reflection of their different derivational effort: the marked form is syntactically more expensive than the unmarked form.

The fact that a certain phenomenon is a phenomenon of markedness, rather than grammaticality, does not necessarily imply that it has no syntactic background. So one part of what we try to do in this paper is to find out, what the impact of syntax is on the marked/unmarked contrast in the order of arguments and to give a case-theoretic explanation for our findings. For more general remarks on the issue of markedness, see section F of this article.

A1. *Variation in unmarked word orders*

If the unmarked order of subject and objects was uniformly the one given in (2.a above) for all clauses of German, there would be no problem for the structural account in (1.). But this is not the case. The unmarked word order of ditransitive constructions varies (as described in detail in Haider 1992). While (4.) illustrates the pattern we already got to know, in (5.) the unmarked order of the objects is ACC > DAT.

4. NOM > DAT > ACC

[1] It has been stated by some native speakers we asked for judgments that (2c., d. and f.) are degraded even in grammaticality. Interestingly, these are the examples where ACC precedes NOM (see section A1 for the difference between DAT–NOM order and ACC–NOM order).

a. *Es hat ein Junge einem Mädchen ein Buch geschenkt* = unmarked order
It has a boy-NOM a girl-DAT a book-ACC presented
b. *Es hat ein Junge ein Buch einem Mädchen geschenkt*
It has a boy-NOM a book-ACC a girl-DAT presented
other verbs that often occur with this unmarked order: *abgewöhnen* (to wean), *beibringen* (to administer), *gönnen* (not to begrudge), *verübeln* (to blame s.th. on s.o.), *verweigern* (deny), *zutrauen* (to think s.o. capable of s.th.)

5. NOM > ACC > DAT
a. *Es hat ein Polizist einen Zeugen einer Gefahr ausgesetzt* = unmarked order
It has a policeman-NOM a witness-ACC a danger-DAT exposed
b. *Es hat ein Polizist einer Gefahr einen Zeugen ausgesetzt*
It has a policeman-NOM a danger-DAT a witness-ACC exposed
other verbs that often occur with this unmarked order: *unterziehen* (to subject to), *ausliefern* (to extradite), *entziehen* (take away from), *unterwerfen* (to subject to), *zuführen* (to bring to)

There are paradigms with two unmarked orders of the objects, too:

6. a. *Es hat ein Freund einem Mädchen einen Jungen vorgestellt* = unmarked order
It has a friend-NOM a girl-DAT a boy-ACC introduced
b. *Es hat ein Freund einen Jungen einem Mädchen vorgestellt* = unmarked order
It has a friend-NOM a boy-ACC a girl-DAT introduced

7. a. *Es hat ein Agent einem Polizisten einen Spion übergeben* = unmarked order
It has an agent-NOM a policeman-DAT a spy-ACC handed over
b. *Es hat ein Agent einen Spion einem Polizisten übergeben* = unmarked order
It has an agent-NOM a spy-ACC a policeman-DAT handed over

Haider (1992) claims that the unmarked orders are a subcategorization property of the verb. Counterevidence against this claim are the following data that show different patterns with the verbs in (6.) and (7.). Both of them can appear with only unmarked DAT > ACC, too:

8. a. *Es hat ein Autor einem Journalisten ein Buch vorgestellt* = unmarked order
It has an author a journalist-DAT a book-ACC introduced
b. *Es hat ein Autor ein Buch einem Journalisten vorgestellt*
It has an author a book-ACC a journalist-DAT introduced

9. a. *Es hat ein Agent einem Polizisten eine Geheimakte übergeben* = unmarked order
It has an agent a policeman-DAT a classified document-ACC handed over
b. *Es hat ein Agent eine Geheimakte einem Polizisten übergeben*
It has an agent a classified document-ACC a policeman-DAT handed over

To give another example, the verb *entziehen* (to withdraw), shows different unmarked orders with different objects:

10. a. *Es hat ein Mann ein Kind einem schlechten Einfluß entzogen* = unmarked order
It has a man a child-ACC a bad influence-DAT withdrawn
b. *Es hat ein Mann einem schlechtem Einfluß ein Kind entzogen*

It has a man a bad influence-DAT a child-ACC withdrawn

11. a. *Es hat eine Frau einen Mordfall einem Detektiven entzogen*
It has a woman a murder case-ACC a detective-DAT withdrawn
b. *Es hat eine Frau einem Detektiven einen Mordfall entzogen* = unmarked order
It has a woman a detective-DAT a murder case-ACC withdrawn

The examples (8.-11.) suggest that the unmarked order of the objects is determined by the animacy of the nouns in question. This observation has been made by Fanselow (1995a): animate objects precede unanimate objects in the unmarked case. Thus, ditransitive constructions with two animate objects have two unmarked orders (cf. 6. and 7.). As expected, ditransitive constructions with two unanimate objects also have two unmarked orders:

12.a. *Es hat ein Junge einem Ball eine Kugel entgegengerollt* = unmarked order
It has a boy a ball-DAT eine shot-ACC towards-rolled
b. *Es hat ein Junge eine Kugel einem Ball entgegengerollt* = unmarked order
It has a boy eine shot-ACC a ball-DAT towards-rolled

In the case of transitive clauses with two animate objects, we have two unmarked orders, when the object has dative case, while we have only one, when the object has accusative case (cf. also Cooper 1994:29f):

13. a. *Auf dem Markt ist ein Nomade einem Römer begegnet* = unmarked order
at the market is a nomad-NOM a Roman-DAT met
b. *Auf dem Markt ist einem Römer ein Nomade begegnet* = unmarked order
at the market is a Roman-DAT a nomad-NOM met

14. a. *Auf dem Markt hat ein Nomade einen Römer getroffen* = unmarked order
at the market has a nomad-NOM a Roman-ACC met
b. *Auf dem Markt hat einen Römer ein Nomade getroffen*
at the market has a Roman-ACC a nomad-NOM met

In some contexts, e.g. when we have experiencer objects, an animate dative object precedes an unanimate nominative in the unmarked case. This is impossible for animate accusative experiencer objects. They never precede the nominative in the unmarked case:

15. a. *Es ist einem Kind ein Stein aufgefallen* = unmarked order
It is a child-DAT a stone-NOM attracted attention
b. *Es ist ein Stein einem Kind aufgefallen*
It is a stone-NOM a child-DAT attracted attention

16. a. *Es hat ein Lied einen Jungen begeistert* = unmarked order
It has a song-NOM a boy-ACC amazed
b. *Es hat einen Jungen ein Lied begeistert*
It has a boy-ACC a song-NOM amazed

A2. *Definitness effects and quantifier scope inversion*

Another piece of evidence for the principled difference between dative and accusative is their different sensitivity to the definiteness hierarchy. It appears to be the case that the definiteness hierarchy 'overrides' the animacy hierarchy in unmarked clauses, but not the 'NOM > ACC-constraint', that itself is stronger than the animacy hierarchy. So in (17.) the unmarked order of dative and nominative is reversed, compared with (15.), if the nominative is definite and the dative indefinite. Again, this reversal does not occur with indefinite nominative and definite accusative in (18.), compared with (16.):

17. a. *Gestern hat einem Kind das Konzert gefallen*
yesterday has a child-DAT the concert-NOM pleased
b. *Gestern hat das Konzert einem Kind gefallen* = unmarked order
yesterday has the concert-NOM a child-DAT pleased

18. a. *Gestern hat ein Konzert den Jungen begeistert* = unmarked order
yesterday has a concert-NOM the boy-ACC amazed
b. *Gestern hat den Jungen ein Konzert begeistert*
yesterday has the boy-ACC a concert-NOM amazed

The qualification as marked order here is due to the fact that in these cases the definite DP has to be interpreted as discourse topic (cf. Jäger 1995). The unmarked interpretation should only occur, if especially the indefinite arguments occupy their 'base position' (cf. also Meinunger 1995). As we see, the 'base position' of the dative in (17.b) then differs from that of the dative in (15.), though the only significant difference lies in the determiners of the nominative DP. The 'base position' of the accusative case on the other hand, is the same in all cases. Thus, if we presuppose that positions in unmarked orders are base positions, then the base position of an indefinite dative object changes, when another argument has a definite determiner (or, presumably, vice versa), but that of an accusative object does not change. This is expected under a theory of datives that gives up the postulation of a unique dative position in the clause, but keeps the assumption of unique case positions for nominative and accusative. This is exactly the picture that we want to draw in this article.

Another equally subtle phenomenon concerns the possibility of scope inversion with quantifiers. A well-known fact about quantifiers in German is that (19.) has two relatively easily accessible readings, while in (20.) the inversed scope reading of the quantifiers is very hard to get, if at all (we abbreviate the possible scope readings by adding $\exists\forall$ and $\forall\exists$, respectively, indicating which quantifier has scope over which one, the $^{?}$ is not a marker of ungrammaticality, but a marker of a reading that is only very hard to get):

19. *Alle Mädchen lieben einen Jungen* $\forall\exists$ $\exists\forall$
All girls-NOM love a boy-ACC
$\forall x \exists y \mid \text{girl}(x) \,\&\, \text{boy}(y) \,\&\, \text{love}(x,y)$
$\exists y \forall x \mid \text{girl}(x) \,\&\, \text{boy}(y) \,\&\, \text{love}(x,y)$

20. *Ein Junge liebt alle Mädchen* $\exists\forall$ $^{?}\forall\exists$
A boy-NOM loves all girls-ACC

d. **Ich zeigte sein$_1$ Foto jedem meiner Freunde$_1$*
I showed his fotograph$_1$ -ACC [each my friends$_1$] -DAT

27. each ... other construction
a. *Ich gab jedem Mann des anderen Uhr*
I gave each man-DAT the other's watch-ACC
b. **Ich gab dem Trainer des anderen jeden Löwen*
I gave [the trainer of the other]-DAT each lion-ACC

28. Negative Polarity
a. *Ich gab niemandem/*jemandem auch nur ein Buch*
I gave noone-DAT/*someone even only one book-ACC
b. *Ich gab kein/*ein Buch auch nur einer Person*
I gave no/*a book-ACC even only one person-DAT
c. *Ich gab auch nur ein Buch niemandem
I gave even only one book-ACC noone-DAT
d. **Ich gab auch nur einer Person kein Buch*
I gave even only one person-DAT no book-ACC

We observe that dative objects can c-command accusatives, but only from a position with A'-properties.

B2. *Datives are extraction islands*

The second contrast between datives and accusatives in German is that datives are extraction islands, but accusatives are not. This holds both for wh-extraction, as shown in (29.), and PP-extraction, as shown in (30.).[4]

29. Wh-extraction (Müller 1993)
a. *[$_{PP}$ *Über wen*]$_i$ *hat der Verleger* [*einem Buch* t_i] *keine Chance gegeben?*
about whom has the editor-NOM a book-DAT no chance given
b. [$_{PP}$ *Über wen*]$_i$ *hat der Fritz der Anna* [*ein Buch* t_i] *gegeben?*
about whom has ART Fritz-NOM ART Anna-DAT a book-ACC given

---

[4] Müller further claims that extraction out of scrambled objects is impossible.We do not agree totally with Müller's judgments. Extraction from a 'scrambled' direct object still seems for us to be much better, if not perfectly grammatical, than extraction from a dative, as in (29.a.):
? [Über wen]$_i$ hat [ein Buch t_i] der Fritz der Anna gegeben
About whom has a book-ACC the F-NOM. the A.-DAT given
Even extraction out of subject is sometimes grammatical. Consider the following examples:
i. *Von Thomas Mann hat mich noch kein Roman überzeugt*
[By Thomas Mann]$_i$ has me-ACC yet [no novel t_i]-NOM convinced
ii. *Von Thomas Mann habe ich noch keinen Roman gelesen*
[By Thomas Mann]$_i$ have I-NOM yet [no novel t_i]-ACC read
iii. **Von Thomas Mann habe ich noch keinem Roman neue Einsichten abgewonnen*
[By Thomas Mann]$_i$ have I-NOM yet [no novel t_i]-DAT new insights won from
'I didn't get new insights from any novel by Thomas Mann yet'
Fanselow (1995) gives the following counterexample against the claim that scrambled DPs are islands for PP-extraction:
iv. *Worüber hätte* [*einen solchen Schmähartikel* t] *selbst der Peter nicht aus Wut verfassen können*
About what would have such a diatribe-ACC t even the Peter not in anger write can
The only clear contrast we can see is with extraction from datives on the one hand and extraction from accusatives on the other hand, at least if we deal with extraction within a clause and not across clause boundaries.

30. PP-extraction[5]
    a. [*Über Scrambling*]$_i$ *habe ich einem Buch über Optionalität* [*einen Aufsatz* t_i] *hinzugefügt*
       'About scrambling have I the book-DAT about optionality [an article-ACC t] added'
    b. *[*Über Optionalität*]$_i$ *habe ich einen Aufsatz über Scrambling* [*einem Buch* t_i] *hinzugefügt*
       'About optionality have I an article-ACC about scrambling [a book-DAT t] added'

To illustrate the problems an account for German dative as structural case leads into, we want to take a short look at the most elaborated theory of this kind, that we could find in the literature, the one from Müller (1993).

B3. *A derivational account: Müller (1993)*

In Müller's model, dative case is assigned in the specifier of a VP-shell, called μP. The dative object moves there from its VP-internal Θ-position:

31. ... [$_{\mu P}$ DAT$_i$ [$_{VP}$ ACC [$_{V'}$ t_i V°]]]

The two positions are illustrated in (32.). The directional PP surfaces in the Θ-position, the dative object in the case position:[6]

32. a. *daß der Fritz* [$_{\mu P}$ [$_{VP}$ *einen Brief* [$_{V'}$ [$_{PP}$ *an den Vermieter*] *geschickt*]]] *hat*
       that ART Fritz a letter to the landlord sent has
    b. *daß der Fritz* [$_{\mu P}$ *dem Vermieter*$_i$ [$_{VP}$ *einen Brief* [$_{V'}$ t_i *geschickt*]]] *hat*
       that ART Fritz the landlord-DAT a letter sent has

How does Müller account for the syntactic facts? With respect to binding he claims that the dative anaphor remains in situ, because it doesn't need case, hence, it can be bound by the accusative object, which is situated in SpecVP. SpecμP, in turn, is an A'-position *per definition,* at least in German. A-binding from this position, as well as extraction out of it, is excluded in German, not e.g. in English. The A/A'-status of Spec-μP, thus, is open for language specific parametrization.[7]

Müller also integrates the phenomenon of free datives into his model. Free dative objects are mostly possessors or beneficiaries. They can be inserted rather freely in German clauses. It wouldn't make sense to consider them as subcategorized by the verbs.

---

5 For further discussion, see Gärtner (1995, this volume), footnote 5.

6 Note that 32.a. and 32.b. are not full thematic paraphrases. Only in 32.a. the landlord might not have been the addressee, while only in 32.b. the letter might have gone to a place different from the landlord's current address. See section D below. Cf. also Meinunger (1995:53) for this effect.

7 In his reply to Larson (1988), Jackendoff (1990) points to the non-productivity of the relationship between oblique and double object structures. But compared to English 'Dative Shift', 'Dative Shift' in German is less productive.

It is typical especially of free datives to have more than one possible reading, as is glossed below the examples in (33.), which are typical examples of possessor or beneficiary datives.

33. a. *Peter hat Maria ein Buch auf den Tisch gelegt*
P. has M.-DAT a book-ACC on the desk put
'Peter put a book on Maria's table' or
'Peter put a book for Maria on some (or Maria's) table' or
'Peter put a book (for Maria) on some (or Maria's) table, because Mary wants him to do so'

b. *Peter backte Maria einen Kuchen*
P. baked M.-DAT a cake-ACC
'Peter baked a cake that is supposed for Maria' or
'Peter baked a cake, because Maria wanted him to do so'

In Müller's account, free dative objects are inserted directly in SpecμP, as illustrated below. Note that the dative here has to be considered as free, because there is no Θ-role of the verb left for it. The directional PP *nach Hause*, 'home', receives the GOAL role:

34. *daß der Fritz* [$_{\mu P}$ *dem Vermieter* [$_{VP}$ *einen Brief nach Hause*]] *geschickt hat*
that ART Fritz the landlord-DAT a letter home sent has
Either 'it is the landlord's home (and not Fritz')' or
'the letter is for the landlord (and it is Fritz' home)' or
'it is the landlord's home and the letter is for him, too'

What Müller canot derive is binding of free dative anaphors, which is as possible as binding of subcategorized datives, as we see in (35.). Because free datives are inserted in SpecμP, they can never occur in a position where they can be bound from the VP-internal case position of the accusative object. If this case position was higher than μP, on the other hand, we would get the wrong unmarked word order for these sentences.

35. a. *Maria setzte die Kinder$_i$ einander$_i$ auf den Schoß*
Maria sat the children-ACC each other-DAT on the lap
b. **Maria setzte den Kindern$_i$ einander$_i$ auf den Schoß*
Maria sat the children-DAT each other-ACC on the lap

Extraction out of free datives is also impossible, as expected:

36. *[*Von wem*]$_i$ *hast du* [*dem Vermieter* t_i] *die Haare geschnitten?*
of who have you the landlord-DAT the hair cut?

We conclude that free datives syntactically behave like subcategorized ones. A derivational account cannot avoid to falsely predict syntactic differences between the two, because one has a VP-internal Θ-position and the other doesnot. Müller further restricts his theory to those verbs that are assumed to have underlying DAT > ACC order – stipulating that the other cases are lexical - which he has no empirical argument for - and result from inherent case marking.

Also, a derivational account, and any structural account, falsely rules out *multiple appearance of datives*. These cases are rarer, but not impossible. Presumably, they are restricted only semantically:

37. a. *Ich habe dir die Wurst dem Oliver auf den Teller gelegt*
I have you-DAT the sausage the Oliver-DAT onto his plate put
'Oliver gets a sausage onto his plate, but the sausage is for you' or
'I put a sausage onto Oliver's plate, as you ordered'
b. *Dem Peter habe ich gestern seinem Auto einen neuen Motor eingebaut*
The Peter-DAT have I yesterday his car-DAT a new engine-ACC built-in
'For Peter's benefit or because of his order, I inserted a new engine into his car'
c. *Helf mir mal deinem Vater in der Küche*
Help me-DAT a minute your father-DAT in the kitchen
'I want you to help your father in the kitchen'
d. *Der David hat mir der Claudia schon zuviele Geschenke gegeben*
the David has me-DAT the Claudia-DAT already too many presents given
'For me, David has already given too many presents to Claudia'

The following properties of dative objects have been demonstrated sofar:

1. dative objects cannot A-bind, but they can A'-bind.
2. dative objects are extraction islands, which also is an A'-property.
3. free datives and subcategorized datives have identical syntactic properties.
4. multiple appearance of dative objects is possible

We conclude from this that dative objects are syntactic adjuncts in German. They surface, where they are inserted. Hence, there is no 'dative movement' in German.[8] Each of the indicated positions in (38.) is a possible site for insertion of the dative object. Which of these is actually chosen, is determined by independent cognitive constraints, e.g. animacy, definitness, and agentivity hierarchy, cf. section A.

38. ... $[_{AGRS}$ (DAT) $[_{AGRS}$ NOM Agrs° $[_{TP}$ T° $[_{AGROP}$ (DAT) $[_{AGROP}$ ACC Agro° $[_{VP}$ (DAT) $[_{VP}$ ...

This seems to us to be the only way to keep the correlation between unmarked order and basic syntactic order in minimalism. Only with the assumption of direct insertion of the dative object into its surface position, it is possible that two different but equally unmarked constructions, like (6.), (7.) and (12.), are also equal in the number of derivational steps. The picture that we draw is thus not only coherent, but also fits into an economy-based theory of syntactic derivations.

An account that treats dative as a structural case postulates a fixed case position and because of this cannot analyse two different but equally unmarked orders as also economically equal. One order is always derived from the other.

An account that takes the opposite direction, and base generates not only datives, but all arguments in their surface positions, is Fanselow's (1993 and 1995a/b) and also Cooper's

[8] We have in mind a relative simple version of syntactic structures that correlates A-properties with case positions.

(1994). What both, and similarly Haider (1992), cannot explain, are the strong syntactic differences between datives and accusatives – without additional stipulations.[9]

Our move leads us to two questions that we would like to address in a more or less sketchy way in the final sections of this paper. The first question is, how we get the thematic interpretation of dative objects, if it can not be theta-role assignment in the usual way. This is the topic of section D. And the second question is, whether the syntax of German datives is totally exceptional or whether there is some systematicity behind it, and how this fits into the larger picture that we have in generative syntax. This issue is raised in section E. The next section, though, is reserved for an additional empirical area where the dative-accusative asymmetry also occurs.

C. Morphology, some effects

The fundamental difference between accusative and dative can also be observed in the area of morphology. Several phenomena illustrate this. The general tendency is: somehow the special morphological properties of datives seem to prevent that they undergo processes that are no problem for accusatives.

As case morphology in German is mostly realized by determiners, the definite article is a good example to illustrate the different patterns:

The German definite article:

	singular			plural
	masculinum	**femininum**	**neuter**	
nominative	*der*	*die*	*das*	*die*
accusative	*den*	*die*	*das*	*die*
dative	*dem*	*der*	*dem*	*den*
genitive	*des*	*der*	*des*	*der*

If we consider nominative as the unmarked form, which is usual, then we have a clear distinction between 'unmarked' nominative and accusative (with the exception of the masculinum singular) and 'marked' dative and genitive. Dative morphology never patterns together with another case form (with the exception of femininum singular, where dative and genitive are alike).[10,11]

---

[9] Fanselow admits this: " ... Insofern wird man wohl kaum darum herumkommen, das Verbot der Bindung DAT_i $einander_i$ als rätselhafte Sonderbeschränkung festzuhalten." (Fanselow 1993: 46)

[10] Another contrast shows up in the inflection of adjectives. German adjectives have a strong and a weak inflection, when they modify nouns, depending on the preceding element, e.g. definite vs. indefinite articles or weak vs. strong quantifiers:

i. viele/einige dumme Männer – alle/die dummen Männer
many/some stupid-STRONG men – all/the stupid-WEAK men (NOM/ACC)

This difference does not occur in dative DPs:

ii. vielen/einigen dummen Männern – allen/den dummen Männern
many/some stupid men – all/the stupid men (DAT)

[11] Hale/Bittner's (1995) case theory makes use of Lamontagne/Travis's (1986) concept of case phrase (KP) to distinguish morphologically simple and morphologically more complex case forms syntactically. Structural or unmarked cases are simple DPs, while non-structural or marked cases like datives are KPs:

a. $[_{DP}\ D°\ [_{NP}\ N\ ]]$ unmarked case
b. $[_{KP}\ K°\ [_{DP}\ D°\ [_{NP}\ N\ ]]]$ marked case

Many German dialects have an even more restricted pattern. Cooper (1994) gives the example of Zurich German. This dialect distinguishes, like many other German dialects, e.g. Upper Hessian and Middle Suebian, only between two morphological case forms. Nominative and accusative are indistinguishable, and the genitive does not exist, its function has been taken over partly by the dative, partly by PPs:

Definite articles in Zurich German (cf. Cooper 1994:15)

	masculinum	femininum	neuter	plural
nom./acc.	*de*	*d*	*s*	*d*
dative	*em*	*de*	*em*	*de*

C1. *Incorporation and nominalization:*

Incorporation of the indirect object is impossible in German in contrast to direct object incorporation. We can find nominalization structures where the direct object is incorporated into the verb.

39. a. *Das Bücher-Schenken machte Spaß*
The books-ACC-presenting made fun
b. *Das Geschichten-Erzählen ist lustig*
The stories-ACC -telling is funny
c. *Das Kuchen-Backen hat gut geklappt*
The cake-ACC -baking has well worked

Such compounds are impossible with dative arguments:

40. a. **Das Kindern-Schenken machte Spaß*
The children-DAT-presenting made fun
b. **Das Kindern-Erzählen ist lustig*
The children-DAT-telling is funny
c. **Das Gästen-Backen hat gut geklappt*
The guests-DAT-baking has well worked

Note that this is true only for processes of true word-formation. *Kindern-Schenken* in (40.a) has to be considered as one phonetic word, with only one word stress (*KINdern-Schenken*). This is ruled out; the sentence is well-formed as long as *Kindern* and *Schenken* remain two phonetic words.[12]

---

[12] This is one possibility to capture the observed morphological data.
An attempt to explain such contrasts has been given by Grimshaw (1990). Her generalization is that only the argument that is the lowest in the thematic hierarchy can incorporate into the verb. The 'theme' is the lowest thematic role in her theory, hence, for every verb that has a theme argument, this is the only argument that can form a nominal compound with the verb. While this analysis can explain why datives cannot incorporate with ditransitive verbs – assuming that theme is standardly linked to the direct object and goal to the dative object–, it cannot explain why dative incorporation is also ruled out when the verb does not select a theme argument, as is the case e.g. with *helfen* ('to help') and *folgen* ('to follow') – see below.

One might suggest that incorporation is possible only with the first object. This would falsely predict that the following sentences with transitive dative-object verbs are well-formed:

41. a. **Beim Kindern-Helfen*
At children-DAT-helping
b. **Beim Eltern-folgen*
At parents-DAT-following

The correct empirical generalization is that incorporation of dative objects is ruled out in principle. What is more, this is independent from thematic properties – which can be shown e.g. with spray-load-alternation-verbs, cf. the following contrast:

42. a. **Beim Kindern-schenken*
At children-DAT-presenting
b. *Beim Kinder-Beschenken*
At children-ACC-*be*-presenting

The verbal prefix *be-* very often marks the 'goal-as-direct object'-variant of the spray/load alternation in German. With respect to Θ-roles, however, *schenken* and *beschenken* are indistinguishable. The children are the goal argument in both cases here. The difference lies only in the (total) affectedness of the respective direct object (cf. Rappaport/Levin 1988 for an extensive discussion) – or maybe the 'centrality' of it (cf. section D below).

As expected under any approach, free datives cannot incorporate either (cf. (40c.)). Again, this phenomenon would find a natural explanation under the assumption that dative objects in general are syntactic adjuncts – incorporation being restricted to internal arguments of the verb.[13] There is no need to assume this, though. A purely morphological explanation might be possible in terms of morphological complexity of datives vs. morphological simplicity of accusatives. The following phenomena point strongly towards the latter strategy.

C2. *Idioms and datives*

This point is very simple. While there are idioms in general in German that have dative-DPs within them, there is one special type of idiom that seems to be impossible with datives in principle. These are idioms of the type 'take care of', 'take advantage of'. In German, these are possible with accusatives, but any imaginable construction with datives is completely odd:

43. a. *Maria hielt Abstand von Peter*
M.-NOM kept distance-ACC of P.
b. *Peter nahm Rücksicht auf Maria*
P.-NOM took consideration-ACC on M.

[13] We thank Hans-Martin Gärtner (p.c.) for the following observation: Nominalized verbs cannot assign dative case any longer, as well as nominative and accusative. But while the latter are realized by genitive case, which can be viewed as the DP-internal structural case, former dative objects can only be realized as PPs.

44. a. **Maria hat ihr Geld Wohlfahrt gespendet*
M.-NOM has her money-ACC charity-DAT donated
b. **Maria folgt Pfad zu Peter*
M.-NOM follows path-DAT to P.

Common analyses of these idiomatic constructions treat the noun parts of them as defective: Presumably, they lack a determiner, structurally we are dealing not with DPs, but just with NPs. It might be the case that a(n abstract) D° is necessary to carry the dative morphology, and that it remains unexpressed when there is no D°.

C3. *Uninflectable indefinites*

Peter Gallmann (1995) reports an interesting fact about certain indefinite expressions in German. Some of these expressions do not have an inflectional morphology, and interestingly they can realize accusative objects, but not datives. Two clear cases of these indefinites are *genug*, 'enough' and *nichts*, 'nothing':

45. a. *Sie hat genug verkauft*
She-NOM has enough-ACC sold
b. *Ich koche heute nichts*
I-NOM cook today nothing-ACC

46. a. **Feuchtigkeit schadet genug*
humidity-NOM harms enough-DAT
b. **Dieser Unmensch hat das Kind nichts ausgesetzt*
This monster-NOM has the child-ACC nothing-DAT exposed

C4. *Complement sentences*

Fanselow/Felix (1987: 85f) report another presumably morphologically induced difference between structural case like accusative and semantic (or, as they say: oblique) case. Some verbs that select a proposition as object in German assign accusative to this object, others dative or genitive. They differ in the possibility to realize the propositional object as a CP:

47. a. Hans leugnete den Diebstahl des Autos
H.-NOM denied the theft-ACC of the car
b. Hans leugnete, daß er das Auto gestohlen hat
H.-NOM denied that he the car stolen has

48. a. Die Darstellung entspricht nicht den Tatsachen
the presentation fits not the facts-DAT
b. *die Darstellung entspricht nicht, daß dieser Verlust uns so schwer traf
the presentation fits not that this loss us so heavily hit
'the presentation does not fit the fact that this loss hit us so heavily'

49. a. wir gedenken der Niederlage bei Waterloo
we commemorate the defeat-GEN at Waterloo

b. *wir gedenken, daß die Armee bei Waterloo geschlagen wurde
we commemorate that the armee at Waterloo defeated was

Fanselow/Felix' explanation makes use of the distinction between structural and oblique case in a different sense – they claim that CP cannot carry case with it and that structural case is not selected by verbs and can be realized optionally, while oblique case is selected by the verb and has to be realized obligatorily. Our explanation points towards the morphological differences: structural case has zero morphology, and that's why CPs – which necessarily cannot express case morphology – are compatible with it. Dative and Genitive on the other hand are morphologically 'more complex' and have to be realized by some element. The odd sentences become fine, when there is a pronoun that carries the case morphology:

50. a. die Darstellung entspricht dem nicht, daß dieser Verlust uns so schwer traf
the presentation fits that-DAT not that this loss us so heavily hit
b. wir gedenken dessen, daß die Armee bei Waterloo geschlagen wurde
we commemorate that-GEN that the armee at Waterloo defeated became

Furthermore, Webelhuth (1990) shows that CP datives are possible as nominative subjects in the *bekommen*-Passiv, some kind of dative passive construction (Webelhuth 1990: 45):[14]

51. a. *Wir messen große Bedeutung bei [$_{CP}$ daß Reagan wiedergewählt wird]
We measure great meaning to that Reagan reelected is
'We attribute great significance to *(the fact) that Reagan is reelected'
b. [$_{CP}$ Daß Reagan wiedergewählt wurde] bekam eine große Bedeutung beigemessen
that Reagan reelected was got a great significance attributed

Again, we see that no thematic or verb-idiosyncratic facts are responsible for the phenomenon, but only the morphological properties of the case forms themselves.

C5. *Split topicalization and quantifier floating*

Meinunger (1995:195) shows that quantifier floating seems to work likewise with nominative, accusative and dative:

52. a. *Frauen*$_i$ *haben da immer nur wenige* t$_i$ *gearbeitet*
women$_i$ have there always only few-NOM t$_i$ worked
b. *Frauen*$_i$ *hat er schon einige* t$_i$ *unglücklich gemacht*
women$_i$ has he already quite some-ACC t$_i$ unhappy made
c. *Frauen*$_i$ *hat er schon vielen* t$_i$ *das Gesicht geliftet*
women$_i$ has he already many-DAT t$_i$ the face lifted

There is an effect, nonetheless, that has to do with 'overt versus covert' realization of the dative morphology. *Frauen* has no overt dative marking suffix, but e.g. *Männern* has. In this case, quantifier floating with a dative is degraded:

[14] Thanks to Chris Wilder for making us aware of this.

53. ?*Männern*$_i$ *hat er schon vielen* t$_i$ *das Gesicht geliftet*
men$_i$ has he already many-DAT t$_i$ the face lifted

Without the overt case marker this sentence gets even worse:

54. ??*Männer*$_i$ *hat er schon vielen* t$_i$ *das Gesicht geliftet*
men$_i$ has he already many-DAT t$_i$ the face lifted

Things change, as soon as we float the quantifier together with an adjective. Now the dative example is at least worse than the others:[15]

55. a. *Kinder*$_i$ *sind nur wenige dumme* t$_i$ *nicht zur Schule gegangen*
children are only few stupid-NOM not to school gone
b. *Kinder*$_i$ *hat er nur wenige dumme* t$_i$ *unterrichtet*
children has he only few stupid-ACC taught
c. ?*Kinder(n)*$_i$ *hat er nur wenigen dummen* t$_i$ *das Lesen beigebracht*
children(-DAT) has he only few stupid-DAT the reading taught

The oddness of the more complex data with datives seems to result again from the bigger morphological complexity of datives. It might be reasonable to formulate the respective restrictions on split constituency in terms of derivational morphology, rather than syntax – though it is a syntactic phenomenon.

D. Some semantic properties of German dative objects

We make do with some hints at the semantics of dative objects here. Let us first consider example (34.), here repeated as (56.). In (56.a) we have a subcategorized dative, in (56.b) a free dative. One would expect that the free dative in (56.b) does not receive the same interpretation as the subcategorized dative. But what we observe is that the *addressee* interpretation of the subcategorized dative is also possible for the free dative in (56.b). We can get even more interpretations for the free dative. But in any case, the (*addressee*-) interpretation of the subcategorized dative in (56.a) – the only one for this dative object – is a proper subpart of the set of the possible interpretations for the free dative (56.b). So there are cases, where a free dative gets the same interpretation as a subcategorized one.

56. a. *daß der Fritz dem Vermieter einen Brief geschickt hat*
that the F. the landlord-DAT a letter-ACC sent has
b. *daß der Fritz dem Vermieter einen Brief nach Hause geschickt hat*
that the F. the landlord-DAT a letter-ACC home sent has

On the other hand – as noted in footnote 6 – we find some differences in interpretation between dative objects and directional PPs. Under a structural approach, both constituents are selected by the predicate and basegenerated in VP, so that we would not expect systematic differences in the interpretation of these two 'goal'-objects.

---

[15] The pattern in (55.) seems to follow that of true split topicalization:
?*Kindern habe ich nur amerikanischen geholfen* – children-DAT have I only american-DAT helped
Another case is the *was für*-split. There is a rather weak contrast, too.
i. ***Was** hast du **für Bücher** den Kindern gegeben* – What have you for books-ACC the children given
ii. ?***Was** hast du **für Kindern** die Bücher gegeben* – What have you for children-DAT the books given

The empirical evidence points towards a uniform treatment of free and subcategorized datives: there are no 'subcategorized' datives with totally idiosyncratic thematic properties. Our strategy is therefore: an account for the semantics of free datives is needed anyway and whatever we will say about free datives can be carried over to 'subcategorized' ones. We treat all 'subcategorized' datives as a proper subset of the set of free datives. We hypothesize that dative case has a semantically underspecified lexical entry that will be further specified in the course of interpretation (considering the linguistic and extralinguistic context) to yield the actual thematic interpretation of the dative object.

To illustrate the facts that have to be captured, we want to exemplify some interesting contrasts here.[16] First, datives are less 'affected' than accusatives:

57. a. *Der Blinde hat dem Hund mit seinem Stock auf den Kopf geschlagen*
The blind man has the dog-DAT with his stick on the head beaten
b. *Der Blinde hat den Hund mit seinem Stock auf den Kopf geschlagen*
The blind man has the dog-ACC with his stick on the head beaten

Accusative objects are usually assumed to be totally affected (cf. Fillmore 1968, Anderson 1971, Tenny 1988). Datives are less than totally affected: in (57.a), the prefered reading is that the blind man beat accidentally, while in (57.b) he beat on purpose (prefered reading again).

On the other hand, the possessor dative is ruled out, when the possessor is not affected:

58. a. *Arsene Lupin hat Cäsars Toga gestohlen (aus dem Museum)*
A. L. has Cesar-GEN toga-ACC stolen (from the museum)
b. **Arsene Lupin hat dem Cäsar die Toga gestohlen (aus dem Museum)*
A. L. has the cesar-DAT the toga-ACC stolen (from the museum)

Cesar cannot be affected, simply because he is dead. We observe: datives are less than totally affected, but more than not affected. A similar observation is that datives can be causers, but not direct causers:

59. *Mir ist dein Fahrrad umgefallen*
Me-DAT is your bike fallen down
≈ 'I accidentally did something wrong, such that your bike fell down'

The semantic lexical entry for 'dative object case' should be designed in such a way that it introduces the 'general direction' for the dative object's interpretation, but leaves open the 'details' which get specified via the linguistic (and extra-linguistic) context. How this specification procedure works is specified by general principles of cognitive inference.

---

[16] One of the subtle facts about datives that have to be addressed is that body-part datives, especially inanimate ones, are in some contexts better than in others:
a. **Peter reparierte dem Tisch die Beine*
P. repaired the table-DAT the legs
b. *Josef schnitt dem Tisch die Beine ab*
J. cut the table-DAT the legs off
The appearance of body-part datives has to be compatible with the conceptual interpretation of the verb. But it is far from obvious, what the respective properties of the verbs are in these cases.

Several semantic approaches to the dative case have been proposed – especially for the first object in the English double object construction (Goldberg 1992, Tremblay 1990) – that center on the notion of possession. A proposal in this direction has also been done by Abraham (1983) for German. These approaches can be characterized as prototype theories of case, maybe even in the spirit of Dowty's (1991) notion of 'proto-roles'. Prototypical datives are considered to be possessors or recipients. Those datives that do not fall into this class are assumed to fall into a class that is conceptually related to the concept of possession - as something like 'metaphorical extensions' of the prototypical meaning.

This way of treating datives would explain the oddity of (58.b). Cesar cannot be a possessor, because he is dead. The problem is, however, that not all German datives can be related to the concept of possession without stretching this concept up to insignificance. On the other hand, not all possessors receive dative case in German, so something additional has to be said anyway.

In the pre-generative era most researchers in the field agreed that cases have some semantic content with them. This holds especially for dative case. The reason this view was given up, was that everyone failed who tried to describe the semantics of a case form like the dative in such a way that she could predict e.g. for novel verbs under which circumstances an argument received this case. Case then was viewed as an epi-phenomenon, and considered more or less as verb-idiosyncratic. Obviously this cannot solve the problems we have with free datives. They need a semantic account and likewise several other forms of oblique cases, like Russian instrumental, or even certain prepositional objects like German *mit*-phrases ('with'-phrases) or partitive *an*-phrases ('at'-phrases; as in *an einem pullover stricken* – 'knit at a sweater').

Note that these three semantic case forms of German (as we would call them) have a semantic property that differentiates them from the structural cases nominative and accusative: they can accompany main verb *sein*, 'be' as predicates:[17]

60. a. *Das Buch ist dem Peter*
    the book-NOM is the Peter-DAT 'the book is Peter's'
    b. *Der Kaffee ist mit Sahne*
    the coffee-NOM is with cream
    c. *Maria ist an einem neuen Artikel*
    M.-NOM is at a new article

[17] The possessor reading of the dative in (60.a) surely is no accident: there must be something right about the possessor theory of the semantics of datives.
This kind of data may have influenced SC approaches to datives like den Dikken (1995). Den Dikken base generates the dative of the English double object construction as predicate of a verbal small clause complement. Presumably, he would base generate German free datives as modifying adjuncts and thus get the same problems with German datives as structural case approaches, namely postulating syntactic differences between free and subcategorized datives, and a lack of explanation, why different surface orders can be equally unmarked and thus equal in cost w.r.t. derivational economy. We do not think that it is impossible to overcome these problems in an SC approach, but the costs in terms of theoretical and conceptual complications are high. Here, we see a big advantage on our side. We do not claim that datives cannot be SC predicates. In fact, (60.a) should be analysed as such a case. Datives can be SC predicates because they have semantic content. But they have this content independent of the syntactic position they occur in.

It is fairly clear that this property is only possible for semantically contentful elements like locative and directional PPs, adjectives and the like. An accusative, and likewise a nominative, is odd in this case:

61. a. **Das Buch ist den Peter*
the book-NOM is the Peter-ACC
b. **Die Maria liess das Buch der Peter sein*
the M.-NOM let the book-ACC the Peter-NOM be

Structural case marked DPs as predicative complements of *sein* are possible, but they change their morphological properties and *agree* with their subject, so they can no longer be considered as having case by themselves:

62. a. *Der liebe Gott ist ein guter Mann*
the kind god-NOM is a good man-NOM
b. *Maria liess den lieben Gott einen guten Mann sein*
M.-NOM let the kind god-ACC a good man-ACC be

It might be possible to describe the semantic content of dative case as those semantic properties that all the thematic roles datives can have in German, have in common. Typical such roles are beneficiary, possessor, recipient, experiencer and other more special ones. Wegener (1985) seems to have something like this in mind. But she failed, as others before, because the features that actually came out were totally unspecific, like AFFECTED. Barnes (1985) tried to capture the facts about French free datives in a similar way. French free dative clitics have nearly the same pattern of distribution as German free datives have.

Wierzbicka (1980) is a case study about the Russian instrumental. Wierzbicka also has a content-related description of case forms in mind, but she is more strongly empirically oriented and more careful about her generalizations. Her strategy is to distinguish the different semantics of instrumental objects by the environment they occur in. In a semi-formal analysis she shows what the several instances of instrumental have in common. It might be that something similar is possible for German datives, too.

Nonetheless, it would not make much sense to do this without elaborating a theory of case systems as such. Wierzbicka points towards the same direction. The question is what one should expect from surface cases. It is clear that different languages have different case systems and although each language has a case form that covers some thematic roles that are covered by the dative in German, we very rarely find a one-to-one relation between two cases of two different languages. Therefore, we have to be sceptic about a semantic description of case forms purely in terms of semantic universals like primes or something similar. What a single case form is able to express, depends in part at least on the properties of the other case forms of a language.

Wierzbicka (1980) introduces the notion of centrality to differentiate nominative and accusative from instrumental and dative (in Russian). Arguments that are central to the described event (in the eye of the speaker), occur in nominative and accusative, arguments that are more peripheral in instrumental and dative. This way of seeing it could account for (57.) and (59.).

The center/periphery distinction could also account for a semantic effect that has been observed by Hudson (1992) for the English double object construction and that occurs with German dative objects, too. Depictive adjectival predicates cannot modify the first object in the double object construction (Hudson 1992: 263)

63. *John$_1$ gave Mary$_2$ the book drunk$_{1/*2}$*

While in (64.a) both Peter and the chancellor (realized as an accusative object) can be interpreted as drunk, only Peter can be in (64.b), not the chancellor, here realized as a dative:

64. a. *Bisher hat der Peter$_1$ den Bundeskanzler$_2$ nur betrunken$_{1/2}$ getroffen* (ambiguous)
Up to now has P.-NOM$_1$ the federal chancellor-ACC$_2$ only drunk$_{1/2}$ met
b. *Bisher ist der Peter$_1$ dem Bundeskanzler$_2$ nur betrunken$_{1/*2}$ begegnet*
Up to now is P.-NOM$_1$ the federal chancellor-DAT$_2$ only drunk$_{1/2}$ met

Being on the periphery, the dative might not be 'accessible' to the depictive adjective, because the latter is on the periphery itself. Further evidence for the case theoretic asymmetry between center and periphery might show up in the following data. First, recipient interpretation for the dative is possible only if the theme argument is in accusative case (65.a) but not if it is realised as a partitive PP (65.b):

65. a. *Peter schreibt Maria einen Brief*
P.-NOM writes M.-DAT a letter-ACC
b. **Peter schreibt Maria an einem Brief*
P.-NOM writes M.-DAT at a letter

Likewise a body part dative is possible if the body part is realized by an argument PP (66.a), but impossible with an adjunct PP (66.b):

66. a. *Zwei Fliegen haben Maria auf der Schulter gesessen*
Two flies-NOM have M.-DAT on the shoulder sat
b. **Zwei Fliegen haben Maria auf der Schulter gekämpft*
Two flies-NOM have M.-DAT on the shoulder fought

The same contrast can occur with the licensing of body part instrumentals. The body part PP can be related to the accusative object but not to the *mit*-PP ('with'-PP):

67. a. *Maria hat den Wagen$_i$ auf seinen$_i$ Rädern in das Lager gerollt*
M.-NOM has the cart-ACC on its wheels into the warehouse rolled
b. **Maria hat die Hemden mit dem Wagen auf seinen Rädern in das Lager gerollt*
M.-NOM has the shirts-ACC with the cart on its wheels into the warehouse rolled

Another difference between 'central' and 'peripheral' cases is that nearly every thematic role can be linked to nominative (if we also take into account passive and other constructions) and accusative (e.g. the subjects of the embedded infinitives in AcI- and ECM-constructions). German *mit*-phrases and datives each can be linked only to (distinct) subsets of the set of possible thematic roles. This difference can be seen as the minimal semantic background for the structural/semantic case distinction.

We leave the issue of the semantics of dative case open for further research at this point.

Sofar we have characterized German datives as semantic arguments and syntactic adjuncts. How exceptional is this behavior and how does it fit into the framework of generative syntax? A theory that has already introduced a distinction between syntactic and semantic arguments is the binding theory of Reinhart and Reuland (1993). We finally want to apply this theory to our German data and see whether we get a satisfactory result.

E. The syntax of dative objects and other oblique arguments in German

Let us first introduce the core ideas of Reinhart and Reuland.

E1. *The binding theory of Reinhart/Reuland (1993)*

The main thesis of Reinhart/Reuland (henceforth R&R) is that the application of the binding conditions should be reduced to cases of true reflexivization (which means coreference of two arguments of the same predicate).

Not all occurrences of anaphors are subject to the binding theory. This is exemplified with Dutch. Dutch has two anaphors, *zich* and *zichzelf*. Only *zichzelf* is a reflexivizer. *Zich* is used in logophoric contexts like long-distance anaphors a.o. The table in (68.) illustrates this pattern.

68.

	SELF	SE	Pronoun
Reflexivizing function	+	-	-
R(eferential independence)	-	-	+

Anaphors and pronouns are distinguished by two properties: the Reflexivizing function and the property of referential independence. Only pronouns are referentially independent, while only SELF anaphors have the reflexivizing function. SE anaphors have neither of these properties.

The difference between syntactic and semantic predicates and arguments is responsible for the contrast between (69.) and (70.).

69. a. **Henk*$_1$ *hoorde hem*$_1$
H.$_1$ heard him$_1$
b. **Henk*$_1$ *hoorde zich*$_1$
H. heard SE
c. *Henk*$_1$ *hoorde zichzelf*$_1$
H. heard SELF

70. a. **Henk*$_1$ *hoorde* [*hem*$_1$ *zingen*]
H.$_1$ heard [him$_1$ sing
b. *Henk*$_1$ *hoorde* [*zich*$_1$ *zingen*]

H. heard SE sing
c. *Henk*$_1$ *hoorde* [*zichzelf*$_1$ *zingen*]
H. heard SELF sing

In (69.) two semantic coarguments are coindexed and the SELF anaphor is required. In (70.) antecedent and pro-form are only syntactic coarguments: they are assigned case by the same syntactic predicate, the matrix verb. But they are not semantic coarguments, because they receive their thematic roles from two different verbs. In this case, only the pronoun is ruled out, while the SE anaphor is possible. This is summed up in (71.).

71.

	SELF	SE	Pronoun
pro-form is only syntactic coargument	+	+	-
pro-form is semantic coargument	+	-	-

The binding principles thus have to be formulated with respect to syntactic and semantic predicates. R&R do this in the following way (Reinhart/Reuland 1993:678):[18]

72. *Definitions*
a. The *syntactic predicate* formed of (a head) P is P, all its syntactic arguments, and an external argument of P.
The *syntactic arguments* of P are the projections assigned a Θ-role or Case by P.
b. The *semantic predicate* formed of P is P and all its arguments at the relevant semantic level.
c. A predicate is *reflexive* iff two its arguments are co-indexed.
d. A predicate (formed of P) is *reflexive-marked* iff either P is lexically reflexive or one of P's arguments is a SELF anaphor.

73. *Conditions*
A: A reflexive-marked syntactic predicate is reflexive.
B: A reflexive semantic predicate is reflexive-marked.

The formulation of the binding conditions in (73.) reflects the contrast between (69.) and (70.). Condition B says that, when two semantic co-arguments are coindexed, a SELF anaphor is required, according to the definition of reflexive-marking in (72.d). This accounts for all of the three sentences in (69.). But condition A can only account for (70.c), but not (70.a) and (70.b). It just requires that a SELF anaphor has to be coindexed with another syntactic argument. Furthermore, nominative anaphors are not excluded. R&R rule out these cases by another condition, that unifies the treatment of binding and A-movement chains. And this is the General Condition on A-chains given in (74.)

74. *General Condition on A-chains*
A maximal A-chain (α_1, ... , α_n) contains exactly one link – α_1 – that is both +R and Case-marked.

This condition claims that the head and only the head of an A-chain has to be both referentially independent and case-marked. Anaphors are not referentially independent, and thus are ruled out as heads of A-chains, which nominatives necessarily are. The pronoun in (70.a) is ruled out because it is both referentially independent and case-marked, but not the

[18] The definitions are the abbreviations R&R gave for ease of representation. The precise definitions should speak of i-reflexivity and i-coindixation. That is, two or more arguments share the same index i.

head of the chain. SE in (70.b) is still allowed, correctly, because the anaphor is not referentially independent.

E2. *An application of Reinhart/Reuland (1993) to German*

When we apply this theory to German, we have to consider that German datives do not count as A-binders, as we saw, contrary to English, and despite the fact that German datives are semantic arguments. To capture this, we relativize the definition of a syntactic predicate in (72.a) by (72.a'):

72. a.' ...
The *syntactic arguments* of P are the A-chains that are assigned structural Case in the extended projection of P and, optionally (i.e. parametrized language-specifically), the semantic arguments of P.

In German, only DPs with structural case count as syntactic arguments, while in English semantic arguments also count as syntactic arguments. This parametrized difference is illustrated in table (75.):

75.

	structural case	semantic case
English	+ syntactic argument	+ syntactic argument
German	+ syntactic argument	- syntactic argument

This parametrization leads to the following A-chain condition for German:

76. *General Condition on A-chains in German* (results from 72.a')
A maximal A-chain (α_1, ... , α_n) contains exactly one link – α_1 – that is both +R and Structural-Case-marked.

The head and only the head of an A-chain in German has to be referentially independent and marked with structural case. Condition A now correctly predicts that in German an accusative anaphor can only be bound by the nominative DP in matrix clauses.

We can now rule out (77.a) with a dative antecedent for the accusative anaphor, and likewise (77.b) with an antecedent marked with another semantic case form, a with-PP. We correctly predict further that in (78.) the pro-form has to be realized as a SELF anaphor, according to condition B. The pronoun is ruled out, because two semantic coarguments are coindexed. This again holds not just for datives, but also for with- and by-phrases.

77. a. **Peter hat Maria$_1$ sich$_1$ gezeigt*
P. has M.-DAT$_1$ SELF$_1$ shown
b. **Ich habe mit Maria$_1$ sich$_1$ beschenkt*
I have with M.$_1$ SELF$_1$ presented

78. a. *Maria$_1$ hat sich$_1$ /* ihr$_1$ einen Kuchen gebacken*
M.$_1$ has SELF-DAT$_1$ /*her a cake baked
b. *Maria$_1$ ist mit sich$_1$ /*ihr$_1$ zufrieden*

M.$_1$ is with SELF$_1$/*her$_1$ satisfied
c. *Maria$_1$ ist von sich$_1$ /*ihr$_1$ enttäuscht*
M.$_1$ is by SELF$_1$ /*her$_1$ disappointed
'Maria has been disappointed by herself/*her'

The binding conditions as stated in (72.a') and (76.) correctly predict a gap: The antecedent for a pro-form with accusative case can never be marked with semantic case in German:

79. **Peter hat den Gästen$_1$ einander$_1$ /sie$_1$ vorgestellt*
has the guests-DAT$_1$ each other$_1$/them$_1$ introduced

The anaphor is ruled out by condition A and the pronoun by condition B. We further predict that a dative anaphor can precede its antecedent more easily than an accusative anaphor. We found evidence that even this prediction might be correct, as given in the contrast between (80.a) and (80.b) below.

80. a. *Maria hat einander$_i$ Kinder$_i$ die Ohren waschen lassen*
M. has each other-DAT children-ACC the ears wash let
b. ??*Maria hat einander$_i$ Kinder$_i$ waschen lassen*
M. has each other-ACC children-ACC wash let

F. Concluding remarks

F1. *On Case theory*

The 'message' of this study is to treat surface case seriously. Postulating AGR-phrases is not sufficient, if this disables us from accounting for the differences between structural case and other case types. Different case types are classified through patterns of different syntactic behaviour, semantic interpretation and morphological properties. The distinction between structural and semantic case in German is threefold:

	structural case	semantic case
syntax	A-properties	A'-properties
semantics	dependent	independent
morphology	simple	complex

We suggest that all case systems make use of both of these case types in one or the other way, but we do not expect the same syntactic properties for them in all languages – one presumably parametrized difference has been illustrated in section E. Though the general tendency of cases seems to be that a complex morphology patterns together with a complex semantics, and the least marked cases also are the 'semantically emptiest', we do not want to propose that it always has to be like this. Nonetheless, in German and many other languages it seems to work exactly this way.

If our semantic treatment of datives is on the right track, there is no need for an underlying θ-position for dative objects, even the 'subcategorized' ones.[19] The issue of θ-role

[19] Researches on several languages came or have been brought to our attention. Up to now, we could not find a single language that has free and subcategorized datives and treats them syntactically in a different way. It has

assignment and argument interpretation can and in fact has to be be left open for the semantic/conceptual component. In some cases, the only way to overcome apparent violations of the θ-criterion – especially the requirement that all θ-roles of a verb have to be assigned syntactically – is to assume that it is not a syntactic requirement, but only a semantic one. To give one example, the verb *versprechen,* 'promise', has three Θ-roles, a speaker S, a hearer H and a proposition P. An example where one role is not realized syntactically but semantically, is the following one:

81. *David hat *(Claudia) einen Ferrari versprochen*
D.-NOM has C.-DAT a Ferrari-ACC promised

David is S, Claudia H, but the Ferrari is not P. So it is not just the case that a θ-role is not realized syntactically, namely P, there is also one DP that gets no θ-role, namely the Ferrari. The Ferrari is part of P, though, which can be paraphrased as 'Claudia will get a Ferrari'. But this is not expressed by a syntactic constituent, so P's 'realization' has to be left for conceptual inference. Claudia is also part of P. This is the reason, why in this case the dative is obligatory, while in general it need not be:

82. *David hat(Karl) versprochen, daß Claudia einen Ferrari bekommt*
D.-NOM has (K.-DAT) promised that C.-NOM a F.-ACC gets

The mechanisms that are involved in thematic interpretation seem to be much more complicated than θ-theory suggests.

Furthermore, there never has been clear independent evidence for the existence of θ-positions in syntactic structure, while we have strong evidence for the existence of structural case positions. So it is empirically more justified to keep the latter and abandon the former than vice versa.

But do we not need traditional θ-role assignment for the structural cases? – With respect to these, we rely on the theory of argument structure developed by Hale and Keyser (1991, 1993), where θ-role assignment to subject and direct object is done by predication. The subject receives its θ-role via predication of VP, and the object via predication of a verbal SC-copredicate that often is incorporated into the verb.[20] This theory has some failures and shortcomings (cf. Steinbach/Vogel 1994), but we assume that the general tendency is correct.

So for now we assume that structural cases get their thematic interpretation via predication, while semantic cases get it via independent semantic rules connected directly with the

been claimed for some Romance languages, e.g. Portuguese, that dative clitic doubling occurs only with subcategorized datives. However, as Albert Branchadell found out, this appeared to be a myth. Clitic doubling occurs either with all datives or with none of them in the Romance languages (see Branchadell 1992 and the references given there). The same holds for clitic doubling in Bulgarian (see Schick/Zimmermann 1995). Even in Basque, which is one of the rare languages that have true indirect object agreement, the agreement morphology occurs likewise with both 'free' and 'subcategorized' datives (see Wunderlich/Joopen 1994). The possibility of multiple datives has been testify for Italian, where dative clitic and dative NP may not be coreferent (Anna Cardinaletti, p.c.) and for Czech (Uwe Junghans, p.c.).

[20] This is not quite right: one type of direct objects, namely cognate objects, receives case presumably by adjacency. This holds for all cases of transitive verbs that have not even an incorporated SC complement in Hale & Keyser's theory.

respective case form. As long as the mechanisms for this have not been worked out, this is more a research proposal than a theory that can be proved and defended here.

E2: *Markedness of sentences*

In section A we claimed that one issue of this article is to find out whether syntax has any impact on the markedness of German clauses. To do this we compared unmarked orderings of arguments and formulated the constraints that have to be assumed to get the right results. We saw that dative and accusative objects differ in the constraints that govern their unmarked position in the clause. This can best be illustrated with their relationship to the nominative in unmarked sentences. We saw that the accusative always follows the nominative in the unmarked case, irrespective of other syntactic and semantic properties of these arguments. But the unmarked position of a dative related to nominative depends on the definiteness hierarchy and, if this does not help, on the agentivity and the animacy hierarchy (see sct. A). It is plausible to assume that these constraints are constraints of different sub-components of the language faculty: the order of structural cases is governed by syntax proper, the computational system in the sense of Chomsky (1995), while agentivity, animacy, and definiteness hierarchy are rules imposed by the conceptual/intentional sub-component, which lies beyond syntax proper.

Our theory predicts that with a nominative and an accusative there is only one optimal configuration with respect to economy of derivation. But with a nominative and a dative, there are several possibilities, because the insertion of the dative is equal in cost, no matter at what stage of the derivation it occurs. Let us assume that the computational system produces more than one single output in such a case. This output is now the input for the conceptual system. We have two different situations, depending on the case of the object:

I. { NOM > ACC }
II. { DAT > NOM ; NOM > DAT }

Let us further assume that the job of our cognitive constraints (definiteness and animacy hierarchy etc.) is 'disambiguation', they filter out the unmarked output. Hence, they apply only when the input consists of a multi-membered set of derivations, as in our case II. In case I nothing is to disambiguate and hence our conceptual filters need and do not apply. Construed in this way, the machinery yields the right results sofar. One addition has to be made: in the case of an ACC > NOM derivation there is again no competitor, but the derivation is marked. This is so, because the syntactic transformations involved here were not only those necessary to get a grammatically well-formed derivation; there was one additional derivational step, the topicalization of the accusative. We have to assume that the computational system can provide such structures, in fact anybody has to (e.g. for the cases of extraction, topicalisation, and extraposition). So our solution must be that these derivations leave the computational system as already marked. Let us assume this as an additional principle for markedness: *A derivation is marked if it contains more derivational steps than ultimately necesssary for convergence*[21] – alternatively, one could assume topic features, focus features, markedeness features etc.pp. to make markedness data compatible with current minimalist theories.

[21] This is not enough. At least it has to be explained, why certain 'unnecessary' transformations are 'grammatical', while others are not.

It seems natural to us to assume that markedness results from violations of constraints at all levels, syntactic as well as conceptual and, of course, prosodic. A sentence is marked, if it departs in at least one respect from the, say, functionally optimal or simplest structure.

Appendix: *Reflexive verbs with dative anaphora*

German has an interesting class of verbs that obligatorily select a reflexive pronoun. The reflexive pronoun does not always realize a Θ-role of its own, for example:

1. a. Ich schäme mich
      I feel ashamed SELF-ACC
   b. Peter beeilt sich
      P. hurries SELF-ACC

These verbs can be called proper reflexive verbs, because the reflexive pronoun can not be replaced by an NP or a pronoun. We cannot even speak sensefully of a second theta-role here. Improper reflexive verbs are those, where the anaphor can be replaced by a pronoun or an NP – because it has a Θ-role of its own:

2. a. Judith wäscht sich/ihn/das Auto
      J. washes [SELF/him/the car]-ACC
   b. Maria versorgt sich/ihn/die Oma mit Bier
      M. supplies [SELF/him/the granny]-ACC with beer

There are some reflexive verbs with dative anaphors. The traditional analysis of proper reflexive verbs is that verb and anaphor together are treated as one single lexical item. Our treatment of datives as verb-independent would predict that there are no proper reflexive verbs with dative anaphors. We classified the examples we found into four groups. The first group forms ditransitive verbs, where the dative can be understood as realizing a Θ-role of its own, like goal or possessor – in this sense they can be understood as improper reflexive verbs, though subject and dative necessarily are co-indexed:

3. a. Ich habe mir etwas angeeignet
      I have SELF-DAT something appropriated
   b. Ich habe mir etwas vorgenommen
      I have SELF-DAT something decided to do

The second class are transitive constructions with dative objects. Both verbs in the examples given here are derived from dative 'selecting' lexical heads, the verb *helfen* (to help) and the adjective *gleich* (equal), respectively. It again makes sense to assume an extra Θ-role for the anaphor here.

4. a. Ich habe mir beholfen mit einer Lüge
      I have SELF-DAT *be*-helped with a lie
   b. Ich bin mir gleichgeblieben
      I am SELF-DAT equal-remained

The anaphors in the following two groups can be omitted, but this has a semantic effect:

5. a. Ich habe (mir) etwas genommen
      I have SELF-DAT something taken
   b. Ich habe (mir) etwas gekauft
      I have SELF-DAT something bought

Only when the anaphor is overt, there is an implication that the agent is also the future owner of what was bought or taken by her.

6. a. Ich habe (mir)etwas angehört
      I have SELF-DAT something listened to
   b. Ich habe (mir) etwas angesehen
      I have SELF-DAT something looked at

In these examples the difference to the anaphor-less variants is that in the latter cases only the perception is described, while the overt anaphor induces a complete, concentrated and purposeful act of perceiving.

All the dative anaphors have some thematic properties of their own. On the other hand, the anaphors cannot be replaced by R-expressions. What is possible, though, is replacement by an anaphor that includes the subject, like a first person plural anaphor in case of a first person singular antecedent. This is impossible for proper reflexive verbs like *sich schämen*:

7. a. Ich habe uns etwas angeeignet
      I-NOM have us-DAT something appropriated
   b. Ich habe uns mit Papas Geld beholfen
      I-NOM have us-DAT with Daddy's money *be*-helped
   c. Ich habe uns etwas gekauft
      I-NOM have us-DAT something bought
   d. ?Ich habe uns etwas angesehen
      I-NOM have us-DAT something looked at

8. *Ich habe uns geschämt
      I-NOM have us-ACC ashamed

This might serve as evidence that there are no proper reflexive verbs with dative anaphors in German, as expected in our approach.

References

Abraham, Werner (1983), "Der Dativ im Deutschen", in: *Colloque du Centre de Recherches germaniques de l'université de Nancy II*, Nancy: 2-101.

Anderson, Stephen (1971), "The Role of Deep Structure in Semanrtic Injterpretation", in: *Foundations of Language 6*: 387-396.

Baker, Mark (1988), *Incorporation*, Chicago.

Barss, Andrew & Lasnik, H. (1986), "A Note on Anaphora and Double Objects", in: *Linguistic Inquiry 17:* 347-354.

Behaghel, Otto (1932), *Deutsche Syntax*, Vol. 4, Heidelberg.

Bittner, Maria and Ken Hale (1995), *The structural Determination of Case and Agreement*, ms. Rutgers and MIT (to appear in *Linguistic Inquiry*).

Branchadell, Albert (1992), *A Study of Lexical and Non-Lexical Datives*, Ph.Diss., University of Barcelona.

Chomsky, Noam (1995) *The Minimalist Program*, MIT Press, Cambridge/Mass.

Cooper, Kathrin E. (1994) *Topics in Zurich German Syntax*, Ph.Diss., University of Edinburgh, published as *GAGL 38*, 1995.

Den Dikken, Marcel (1995), *Particles: On the syntax of verb-particle, triadic and causative constructions*, OUP, Oxford/New York.

Dowty, David (1991), "Thematic Proto-Roles and Argument-Selection", in: *Language 67/3*: 547-619.

Fanselow, Gisbert (1993), "Die Rückkehr der Basisgenerierer", *GAGL* 36: 1-74.

Fanselow, Gisbert (1995a), *Scrambling and Anti-Scrambling*, Vortrag auf dem Workshop „The Role of Economy Principles in Linguistic Theory", Berlin, 9.-11.2.1995.

Fanselow, Gisbert (1995b), *A Minimalist Approach to free Constituent Order*, ms. University of Potsdam.

Fanselow, Gisbert and Sascha W. Felix (1997), *Sprachtheorie 2: Die Rektions- und Bindungs-Theorie*, Tübingen.

Fèry, Caroline (1993), *German Intonational Patterns*, Tübingen.

Fillmore, Charles J. (1968), "The Case for Case", in: Bach, E. and R.T. Harms (eds.), *Universals in Linguistic Theory*, New York: 1-88.

Gallmann, Peter (1992), "Dativanhebung?", in: *GAGL 35:* 72-91.

Gallmann, Peter (1995), "Zur Syntax und Morphologie von Indefinita des Typs *genug*", talk, given at FAS, Berlin, September 1995.

Goldberg, Adele E. (1992), "The inherent Semantics of Argument Structure: the Case of the English ditransitive Construction", *Cognitive Linguistics 3-1*: 37-74.

Grewendorf, Günter (1988), *Aspekte der deutschen Syntax,* Tübingen.

Grimshaw (1990), *Argument Structure*, MIT Press, Cambridge/Mass.

Hale, Ken and S. Jay Keyser (1991), *On the Syntax of Argument Structure*, Lexicon Project Working Papers, Cambridge/Mass.

Hale, Ken and S. Jay Keyser (1993), "On Argument Structure and the lexical Expression of Syntactic Relation", in: Hale, Ken and S. Jay Keyser (eds.), *The View from Building 20*, Cambridge/Mass.

Haider, Hubert (1992), *Branching and Discharge*, Manuscript, University of Stuttgart.

Haider, Hubert (1993), *Deutsche Syntax generativ*, Tübingen.

Höhle, Tilmann (1982), „Explikation für 'normale Betonung' und 'normale Wortstellung'", in: W. Abraham (ed.), *Satzglieder im Deutschen*, Tübingen.

Hudson, Richard (1992), "So-called 'Double Objects' and Grammatical Relations", in: *Language 68*: 251-276.

Jackendoff, Ray (1990), "On Larson's Treatment of the Double Object Construction", in: *Linguistic Inquiry 21*: 427-456.

Jäger, Gerhard (1995), *Topics in Dynamic Semantics*, Ph.Diss. Humboldt University Berlin.

Joppen, Sandra and Dieter Wunderlich (1994) *Argument Linking in Basque*, Theorie des Lexikons 63, Düsseldorf.

Lamontagne, Greg and Lisa deMena Travis (1986), "The Case Filter and the ECP", in: *McGill Working Papers in Linguistics 4*: 51-75.

Larson, Richard (1988), "On the Double Object Construction", in: *Linguistic Inquiry* ***19****:* 335-391.

Larson, Richard (1990), "Double Objects Revisited: Reply to Jackendoff", in: *Linguistic Inquiry* ***21***: 589-632.

Lenerz, Jürgen (1977), *Zur Abfolge nominale Satzglieder im Deutschen*, Tübingen.

Meinunger, André (1995), *Discourse Dependent DP (De-) Placement*, Ph.Diss., University of Potsdam.

Müller, Gereon (1993), *On Deriving Movement Type Asymmetries*, SFS-Report-05-93, University of Tuebingen.

Rappaport, Malka and Beth Levin (1988), "What to do with Theta-Roles", in: Wilkins, W. (ed.), *Syntax and Semantics 21 - ThematicRrelations*: 7-36.

Reinhart, Tanya and Eric Reuland (1993), "Reflexivity", in: *Linguistic Inquiry* ***24***:650-720.

Sabel, Joachim (1995), *Restrukturierung und Lokalität. Universelle Beschränkungen für Wortstellungsvarianten*, Ph.Diss., University of Frankfurt.

Schick, Ivanka P. and Ilse Zimmermann (1995) "Das possessive Klitikum des Bulgarischen", talk, given at *Erste Europäische Konferenz 'Formale Beschreibung slavischer Sprachen'*, Leipzig, November 1995.

Stechow, Arnim and Susanne Uhmann (1986), "Some Remarks on Focus Projection", in: Abraham, Werner and Sjaak de Meij (eds.), *Topic, Focus, and Configurationality*, Amsterdam/Phil.: 295-320.

Steinbach, Markus and Ralf Vogel (1994), *Zum Konzept der Tiefenstruktur in der Generativen Grammatik*, M.A. University of Frankfurt/Main.

Strigin, Anatoli (1995), *Abductive inference during update: the German preposition* mit *(with)*, Mscr., MPG Berlin.

Tenny, Carol (1989), *The Aspectual Interface Hypothesis*, Lexicon Project Working Papers 31, Cambridge/Mass.

Tremblay, Mireille (1990), "An Argument Sharing Approach to Ditransitive Constructions", in: *The Proceedings of the Ninth West Coast Conference on Formal Linguistics*, Stanford.

Uhmann, Susanne (1987), *Fokussierung und Intonation*, Ph.Diss., Universität Konstanz.

Vogel, Ralf and Markus Steinbach (in prep.) *The Semantics of Dative Objects in German*, ms. Max-Planck-Gesellschaft and Berlin-Brandenburgische Akademie der Wissenschaften, Berlin.

Wegener, Heide (1985), *Der Dativ im heutigen Deutsch*, Tübingen.

Webelhuth, Gert (1990), "Diagnostics for Structure", in: Grewendorf, Günther and W. Sternefeld (eds.), *Scrambling and Barriers*, Amsterdam/Phil.: 41-76.

Wierzbicka, Anna (1980), *The Case for Surface Case*, Ann Arbor.

Zubin, D.A. and K.-M. Köpcke (1985) "Cognitive Constraints on the Order of Subject and Object in German", *Studies in Language 9*: 77-107.

markus@skyline.asg.ag-berlin.mpg.de *ralf@skyline.asg.ag-berlin.mpg.de*

Antecedent-Containment and Ellipsis

Chris Wilder

Max-Planck-Gesellschaft, Arbeitsgruppe Strukturelle Grammatik
Jägerstr. 10-11, D-10117 Berlin, Germany
chris@asg.ag-berlin.mpg.de

0. Abstract

This paper [1] examines old and new data relating to the Antecedent-Contained-Deletion (ACD) phenomenon in English. The primary theoretical interest of ACDs lies in the evidence they provide for the assumption of LF movement processes - A'-movement (Quantifier Raising - May 1985 et. al.), or A-movement of objects in English (Hornstein 1994). There is one class of examples involving 'wide scope' VP-ellipsis, as in *John said that more trees had died than Mary did*, that has so far been overlooked. It is argued here that these examples must be included in the paradigm to be accounted for. The extended paradigm is then shown to support the following claims:

(i) ACDs do not support the assumption of covert A-movement of objects in English, contra Hornstein (1994).

(ii) ACDs do provide evidence for a covert A'-movement operation of the type identified as QR by May (1985) and others.

(iii) In addition to the ban on containment of a VP-ellipsis site within its antecedent VP at LF, the account of ACDs must refer to an independent principle that prohibits containment at PF of a VP-ellipsis site within the terminal string corresponding to its antecedent VP.

(iv) Certain examples have been mis-identified as ACDs arising from VP-ellipsis, which actually involve two adjacent deletions, neither antecedent-contained in surface order, one due to Pseudogapping (ellipsis of subparts of VP), the other due to the independent process of Backward Deletion (a.k.a. 'non-coordinate' Right Node Raising).

Furthermore, properties of ACDs are uncovered which require claims in May (1985), Larson & May (1990) about constraints on QR to be reconsidered.

1. The ACD phenomenon

1.1 VP ellipsis and antecedent containment

The antecedent contained deletion phenomenon forms a subcase of the VP ellipsis phenomenon illustrated in (1-2):

(1) a. John met Mary, and Bill did, too.
b. Although John met Mary, Bill didn't.
c. Although John did, Bill didn't meet Mary.
d. John met everyone that Mary did.
e. John said that Mary did.

(2) a. John [$_{VP2}$ met Mary], and Bill did [$_{VP1}$ e] , too.

b. Although John [$_{VP2}$ met Mary], Bill didn't [$_{VP1}$ e] .

c. Although John did [$_{VP1}$ e], Bill didn't [$_{VP2}$ meet Mary] .

d. John [$_{VP2}$ met everyone that Mary did [$_{VP1}$ e]] .

e. John [$_{VP2}$ said that Mary did [$_{VP1}$ e]] .

A VP ellipsis site (VPE), such as VP1 in the examples in (2), can be viewed as an anaphoric element that depends on the content of an antecedent VP for its interpretation. A VPE may take as its antecedent a VP in the same sentence, which may be arbitrarily distant from the VPE; alternatively, the antecedent may be a VP in the external discourse, which need not be linguistically explicit (cf. Sag (1976), Williams (1977)). Embedded in some suitable discourse, all examples in (2) are well-formed, with VP1 taking its antecedent from that discourse. In examples (2a-c), VP1 may take a VP in the same structure (i.e. VP2) as its antecedent, with no issue of antecedent containment arising. The special case of antecedent contained deletion is represented by (2d,e): VP1 is contained within VP2 and VP2 is understood as the antecedent of VP1. The puzzle that arises is that while (2e) is ill-formed (the intended reading is simply impossible), (2d) is not.

In section 2., three accounts of the contrast between (1d,e) under the 'antecedent-contained' reading are reviewed: May (1985) and Larson & May (1990); Baltin (1987); and Hornstein (1994). All these accounts depend on the assumption that (2) holds at some level (whether as axiom or theorem):

(3) A VPE may not be contained in its antecedent

It may be thought that there are semantic reasons for the ill-formedness of structures violating (3), namely that there is no finite interpretation that can be assigned. In (2e), for instance, the content of VP2 ('what John did') is dependent on the content of VP1 ('what Mary did'), but since the content of VP1 is in turn dependent on what John did, a circular reading results - similar to the effect found in (presumably syntactically well-formed) sentences like (4) under the reading indicated (cf. Haik 1987 for discussion):

(4) (*) [her$_j$ brother]$_k$ likes [his$_k$ sister]$_j$

A purely syntactic account of VP ellipsis that makes no reference to interpretation is also possible. For example, it is clear that the generation or licensing of VPE-constructions is contingent on some notion of syntactic identity between the the VPE and its antecedent. (3) could be made to follow simply from the requirement that VPE must be identical with its antecedent, under a plausible definition of syntactic identity. [2] There is, for example, no way for VP1, properly contained within VP2, to be identical to VP2 in terms of depth of embedding: VP2 is necessarily "deeper" than VP1. Hence, it is possible that no interface representation containing an antecedent-contained VPE that meets the identity requirement will be generable, so that questions of interpretation do not even arise.

Be that as it may, (2d) clearly is well-formed and does have a finite interpretation. Accounts of how that reading come to be possible take the form of assuming that the example has (at the relevant level) a syntactic representation different from (2d), which avoids violating (3). Details depend on general assumptions about how VPE's and their interpretation are analyzed. There are two main competing approaches: choice between these is essentially orthogonal to the issues surrounding ACDs. [3] In the proposal of Williams (1977) (cf. also May 1985, Hornstein 1994), a VPE is generated as a node lacking internal structure, dominating a designated phonologically null

terminal - [$_{VP}$ e] - which occurs in S-structures that feed PF (whence the gap). Following S-structure and prior to LF, the content of the antecedent VP is 'copied in' by a syntactic operation: a copy of the antecedent VP is substituted for [$_{VP}$ e]. The major argument for this approach lies in its ability to account for a range of parallelism effects governing the interpretive possibilities available for VPEs.

The main alternative consists in assuming that the VPE gap in PF arises through phonological deletion, i.e. material generated as a normal VP is deleted under identity with material belonging to the antecedent in the PF-wing of the grammar. To ensure the correspondence between deletion in PF and interpretive dependence at LF, material so deleted has to have been marked in some way in the syntax already, with this marking indicating dependence on its antecedent with regard to interpretation at LF (cf. Wilder 1995a for discussion).

In the context of the Minimalist Program, the latter approach is more natural, avoiding two extra assumptions: (i) the LF-copy operation, which is not reducible to Chomsky's (1995) "Move" or "Merge"; (ii) the addition of the designated terminal to the class of lexical items (given the restrictive hypothesis that structure is only licensed via projection of items from the lexicon). The 'PF-deletion' operation required in this alternative is independently assumed in the Minimalist model as the mechanism for trace-gap creation. Additionally, in examples like (5) (cf. also (2d)), the presence at S-structure of a wh-phrase in Spec,CP is not naturally combined with the assumption that VP1 dominates its trace only after LF-copying of VP2:

(5) I know who Mary [$_{VP2}$ met *t*], but I don't know who John did [$_{VP1}$ e]

More importantly, the fact that interpretive parallelism effects observed in coordinative VPE-constructions like (1a) are independently attested in full structures in which a VP is spelled-out but marked as dependent by deaccenting, means that some mechanism to ensure parallelism must be assumed independently of VPE. This removes the argument in favour of adopting the LF-copy mechanism (cf. Tancredi 1992, Chomsky 1993, Fox 1994; Fox also gives further arguments against the LF-copy approach).

In the following, I shall assume a PF-deletion approach to VPEs to be correct, although nothing hangs on this decision; in reviewing previous accounts, discussion is framed in terms of the LF-copying approach where appropriate.

Assuming the LF-copy theory, and assuming the structure (2e) as input to the copy-operation, there is a syntactic reason for the fact that VP1 cannot take VP2 as its antecedent. The output of copying results in a structure (6b) which contains a second instance of VP1, requiring to be replaced by a copy of VP2. The derivation will thus slip into infinite regress, a well-formed LF cannot be derived:

(6) a. John [$_{VP2}$ said that Mary did [$_{VP1}$ e]]
b. John [$_{VP2}$ said that Mary did **[$_{VP2}$ say that Mary did [$_{VP1}$ e]]]**
c. *etc.*

Under the PF-deletion approach, and given the merger theory of structure generation (Chomsky 1993, 1995), the example can also be excluded as not derivable. Assume that VP1 and VP2 must be structurally identical, i.e. syntactic copies of each other according to some measure of identity, for a VPE-dependency to arise. Then no representation corresponding to (6a) can be generated: VP2 cannot be generated independently of VP1 and be structurally identical to VP1, since VP2 must also properly contain VP1.

The conclusion with respect to (2d.) is that the structure given cannot be the correct one. Assuming that the antecedent VP is copied into [$_{VP}$ e], the structure given will lead to the same problem of infinite regress just sketched in (6) for (2e.):

(7) a. John [$_{VP2}$ met everyone that Mary did [$_{VP1}$ e]]
b. John [$_{VP2}$ met everyone that Mary did **[$_{VP2}$ meet everyone that Mary did [$_{VP1}$ e]]]**
c. *etc.*

Alternatively, under the PF-deletion approach, the structure giving rise to VP-deletion cannot have been generated, if VP1 is properly contained in VP2.

In the classic approach to (2d) of May (1985), the structure given is assumed to be the correct one for S-structure. Therefore "S-structure cannot be the level for the copying operation". It is then proposed that the S-structure representation is altered in the LF-component in such a way that the VPE is no longer contained within its antecedent VP prior to the copying operation needed to derive a well-formed LF. The operation that achieves this is Quantifier Raising, which raises the object NP containing the VPE in the relative clause out of VP, adjoining it to IP:

(8) a. *SS* John met everyone that Mary did [$_{VP}$ e]
b. *QR* [$_{IP}$ **[$_{NP}$ everyone that Mary did [$_{VP}$ e]]** [$_{IP}$ John met ***t***]]
c. *Copy* [$_{IP}$ [$_{NP}$ everyone [$_{CP}$ *Op* that M. *Past* **[$_{VP}$ meet *t*]**]] [$_{IP}$ J. *Past* **[$_{VP}$ meet *t*]**]]

At the stage at which copying applies, the antecedent VP consists of the verb and the trace left by QR, and the dependent VP is no longer contained within its antecedent. Copying the antecedent VP into the VPE thus avoids regress. Additionally, this analysis yields a representation which (i) satisfies other well-formedness conditions - for example, the relative clause operator c-commands a trace (variable) which it can bind; and (ii) forms an accurate basis for capturing the actual intepretation of the example. From (8c), a relatively simple translation mechanism will yield a formula such as (9), which represents the relevant aspects of the meaning of (8a) accurately enough:

(9) $\forall x$ (met(m)(x) $\rightarrow$ met(j)(x))

1.2 Conditions on Antecedent-containment

The assumption (3) is generally assumed to hold at LF (May (1985), Hornstein (1994), etc.). One exception is Baltin (1987), who claims that (3) holds at S-structure. In the following, I make the claim (somewhat related to Baltin's) that reference must be made to two independent conditions - (3), holding at LF, and (10), holding at PF:

(10) A VP ellipsis site E may not be *linearly* contained in its antecedent string

While a VPE may, at PF, apparently either follow or precede the string of terminals corresponding to its LF-antecedent (cf. (2b) vs. (2c)), (10) states that a VPE may not be properly contained within that string: i.e. for any two terminals (α, ß) belonging to the string spelling out the antecedent of E, if α precedes E, then E must not precede ß. This must be excluded independently of (3). (10) is motivated (i) by some new evidence showing that the distribution of VPE is governed by generalizations over surface orders that cannot be reduced to (3) holding at LF (sections 3-5); (ii) a reinterpretation of some familiar facts (section 6).

1.3 Pseudogapping vs. Antecedent-containment

Two further assumptions underlying the discussion of ACDs are:

(11) a. the target of VPE is a single constituent
b. that constituent is VP

One problem that threatens the whole discussion is the possibility that one or both of these assumptions is false.

Consider (11b). If the deletion were only to target the verb, then the dependent element would not be contained in its antecedent. The trace of the relative operator would then be a "remant" not contained within the deletion site:

(12) John [$_{V2}$ **met**] everyone [*Op* that Mary did [$_{V1}$ *meet*] *t*]]

The pseudo-gapping construction (13) provides evidence that such deletion of a verb alone, leaving its object as remnant, is possible:

(13)a. ? John met Bill and Mary did, Paul
b. John [$_{V2}$ **met**] Bill and Mary did [$_{V1}$ *meet*] Paul

Hence, (2d)/(7) are potentially irrelevant to the discussion of antecedent-containment.

Furthermore, the pseudo-gapping construction shows that discontinuous deletion sites (discontinuous in surface order at least) are possible. Consider:

(14)a. ? John sent the book to Bill and Mary did, the papers.
b. John [$_{V2}$ **sent**] the book [$_{PP2}$ **to Bill**] and Mary did [$_{V1}$ *send*] the papers [$_{PP2}$ *to Bill*]

Unless it can be shown that such discontinuous deletions form a constituent at some level, e.g. LF, the possibility must be countenanced that apparent VP-deletions are actually composed of adjacent deletions of smaller constituents. So (1a) (=15a) may involve adjacent V-deletion and NP-deletion sites, as in (15b):

(15)a. John met Mary, and Bill did, too.
b. John [$_{V2}$ **met**] [$_{NP1}$ **Mary**] Bill did [$_{V1}$ *meet*] [$_{NP1}$ *Mary*], too.

Hence, more complex examples of apparent ACDs that avoid the objection posed by (13) may turn out to consist of several deletion dependencies, none antecedent contained. Instead of involving an ACD (16b), (16a) may be analysed by analogy with (15b), as involving two adjacent, independent deletions, neither antecedent-contained:

(16)a. John met Mary everywhere that Bill did.
b. John [$_{VP2}$ **met Mary** everywhere that Bill did [$_{VP1}$ *meet Mary t*]]
c. John [$_{V2}$ **met**] [$_{NP1}$ **Mary**] everywhere that Bill did [$_{V1}$ *meet*] [$_{NP1}$ *Mary*] *t*.

I put this issue aside here, returning to it in section 6.1.

In section 2., I review some previous accounts of the ACD construction, highlighting crucial examples and other facts which these sought to account for. In section 3., these accounts are confronted with new data which motivate the approach to be defended here.

2. ACDs 1985-1995

2.1 May (1985)

May's classic analysis of ACDs, according to which hierarchical containment configurations holding at S-structure can be resolved at LF by QR, prior to LF-copying, supplies one of the major arguments for the existence of the quantifier raising operation in covert syntax.

One argument that May used to support the QR approach to ACDs was that only quantified expressions license ACDs in relative clauses modifying them. Thus (17), with a name modified by a nonrestrictive relative, is deviant under the antecedent-contained reading; this is predicted, since a name, not being a quantified expression, does not undergo QR (May 1985:12):

(17) *? John accused Mary, who Peter did [$_{VP}$ e]

The object NP remains in VP at LF, hence the VPE cannot take the VP headed by *accused* as its antecedent without inducing infinite regress.

This argument is put in doubt by the fact that (17) improves significantly if *too* is added, as observed by Wyngaerd & Zwart (1991):

(18) ? John accused Mary, who Peter did [$_{VP}$ e], too.

The pattern can be reproduced for VPE in coordinate contexts (19), suggesting that the cause of the deviance of (17) was incorrectly diagnosed by May, and that it is due to some completely independent factor (cf. also Hornstein 1994):

(19)a. *? John accused Mary, and Peter did.
b. John accused Mary, and Peter did, too.

The well-formedness of (18) does not, however, necessarily threaten the QR approach to ACDs in restrictive relatives. Lasnik (1995) notes some differences between ACDs in nonrestrictive relatives and those in restrictive relatives, including the contrast (21):

(21)a. John stood near everyone Mary did [$_{VP}$ e] .
b. * John stood near Sue, who Mary did [$_{VP}$ e] (too).

Lasnik suggests that non-restrictive cases involve a pseudogapping derivation of the type discussed in section 1.3. Then, (18) is to be analyzed as in (22a), and (21b) as in (22b):

(22)a. John [$_{V°}$ **accused**] Mary, who$_j$ Peter did [$_{V°}$ *accuse*] t_j too.
b. (*) John [$_{V°}$ **stood**] [$_{P°}$ **near**] Sue, who$_j$ Peter did [$_{V°}$ *stand*] [$_{P°}$ *near*] t_j too.

The deviance of (22b) can then be related to the impossibility of deletion of the V+P combination (*stand near*) in the coordinate structure (23a): [4]

(23)a. * John stood near Sue and Peter did, Mary
b. John stood near Sue and Peter did, near Mary

If the pseudogapping approach generalizes to all non-restrictive examples, then these are not ACDs at all, and the well-formedness of (18) is irrelevant. At the same time, the well-formedness of (21a), contrasting with (21b) and (22b), tells against the possibility of generalizing a pseudogapping analysis to all ACDs in restrictive relative clauses. The contrast can be handled by assuming a ellipsis of the whole VP including a stranded preposition (near) is possible in (21a), because QR breaks up the antecedent containment configuration prior to LF; while QR could not break the antecedent containment arising from VP-ellipsis in (21b).

2.2 Baltin (1987)

Baltin (1987) constructs an argument against the QR account of the resolution of ACD's from the observation that the account "overgenerates": readings are predicted to be possible that do not in fact exist.

One such example is (24a). The VPE in the relative clause may only take the embedded VP as its antecedent (24b), although the QR account predicts that the higher VP headed by *thought* should be a possible antecedent as well (i.e. that (24a) should be ambiguous between (24b) and (24c)):

(24) a. Who thought that Fred read how many of the books that Bill did?
b. = Who thought that Fred read how many of the books that Bill *read*?
c. ≠ Who thought that Fred read how many of the books that Bill *thought that Fred read* ?

As well as applying to quantificational expressions such as *everyone*, Quantifier Raising is assumed to apply in multiple interrogatives to wh-phrases located "in situ" at S-Structure, such as the the wh-phrase headed by *how (many)* in (24). This wh-phrase raises covertly to a position from which it takes scope as a wh-operator, namely the specifier of the root CP, where it adjoins to the wh-phrase *who* in the standard analysis. It is crucial to the QR account of the resolution of ACD's, that the whole quantifier phrase is raised by QR, including the relative clause containing the VPE (or at least, that pied-piping of the relative clause is possible). So, assuming that the relative clause in (24) is "pied-piped" by wh-raising at LF, the complex specifier of the root CP will contain the VPE at LF, as shown in (25a):

(25) a. [who_j [**how many of the books that Bill did** $[_{VP}$ **e** $]]_k$]

b. [SPEC C° $[_{IP}$ t_j $[_{VP1}$ thought [that Fred $[_{VP2}$ read t_k]]]]]

The VPE is contained neither in VP1 nor in VP2 at LF (25b). Hence, both VP1 and VP2 should be available as alternative antecedents to VPE. But factually, only VP2 can antecede the VPE - cf. (24).

In the following, a reading such as (24b), where a VPE takes only the first VP up as its antecedent, is termed a narrow scope reading; (24c), where a higher VP is intended as antecedent, is a wide scope reading. There is no independent reason why the VPE in the LF-representation (25) should not take VP1 instead of VP2 as its antecedent. In (26a), where the relevant wh-phrase has raised overtly, both narrow and wide scope readings are possible:

(26) a. How many of the books that Bill did, did Mary think that Fred read ?
b. = How many of the books that Bill *read*, did Mary think that Fred read ?
c. = How many of the books that Bill *thought that Fred read*, did Mary think that Fred read ?

Thus, there is a mismatch between the scope taken by a scope-bearing element E and the scope of a VPE contained within the phrase headed by E. While QR at LF may assign the quantifier wide scope, the VPE may only get a narrow scope reading. Baltin calls this restriction on ACD's the Boundedness restriction.

In his analysis of the ACD construction, Baltin makes a different assumption concerning the filter on antecedent-containment:

(27) There is no well-formed sentence in which a VPE is contained in its antecedent in the S-structure representation.

If (27) is true, then QR in the LF-component will not suffice as a mechanism by which antecedent-containment is resolved. Even if the application of QR should result in a VPE no longer being

contained in its antecedent at LF, this would happen, as it were, 'too late' in the derivation to save the sentence, which will aleady be marked as ungrammatical (in violation of (27)) prior to the QR operation. [5]

Baltin proposes that a different movement operation is responsible for resolving antecedent-containment, namely Relative Clause Extraposition (RCE). RCE is assumed to apply in the overt syntax, extracting a relative clause from the NP it modifies and right-adjoining it to the VP (or some higher projection). Supposing that RCE has applied - string-vacuously - to (1d), the S-structure representation looks like (28a), and (27) is satisfied. QR may then apply to yield (28b), on the basis of which VP1 may take VP2 as its antecedent without incurring regress (28c):

(28) a. John [$_{VP3}$ [$_{VP2}$ met everyone] [that Mary did [$_{VP1}$ e]]
b. [**everyone** [John [$_{VP3}$ [$_{VP2}$ met *t*] [that Mary did [$_{VP1}$ e]]
b. [everyone$_j$ [John [$_{VP3}$ [$_{VP2}$ **met** t_j] [that Mary did [$_{VP1}$ *meet* t_j]]

Given (27), QR is clearly insufficient to account for ACD-resolution, even if it is involved in the derivation of relevant examples. The hypothesis that RCE resolves containment, on the other hand, is not only compatible with (27), but also provides a means for accounting for the Boundedness Restriction on ACD-resolution.

Extraposition is subject to a limitation known as 'Right Roof Constraint' which prohibits it from moving a constituent out of the minimal sentence (CP) containing it. The Right Roof Constraint is usually illustrated by pointing to word order facts such as (29):

(29) a. Bill will say that John kissed everyone that he met tomorrow.
b. * Bill will say that John kissed everyone tomorrow that he met.

The relative clause, which modifies the object of the embedded clause, may not be placed to the right of the adverbial modifying the matrix VP, a linear position that corresponds on standard assumptions to a hierarchical position higher than the matrix adverbial, hence outside of the embedded clause. Given this interpretation of RCE and the Right Roof constraint, the impossibility of wide scope interpretation in (24) can be explained. Since RCE can only raise the relative clause prior to S-structure to a position within the clause containing the object *how many of the books*, the VPE inside it is no longer contained in the minimal VP headed by *read*, but is still contained within the higher VP headed by *thought*:

(30) Who [$_{VP4}$thought that Fred [$_{VP3}$[$_{VP2}$read how many of the books][that Bill did [[$_{VP1}$e]]]

Assuming that no post-S-structure operation affects the relative hierarchical positions of the VPs in (30), there is exactly one VP which does not contain VP1 at LF and which may function as its "antecedent", namely VP2. (QR may, in fact must, raise the phrase *how many...*, to ensure that at LF, VP1 contains a trace that can be bound as a variable by that phrase). In this way, Baltin is able to derive the Boundedness Restriction on ACDs.

2.3 Larson & May (1990)

In their reply to Baltin (1987), Larson and May (1990) give a series of counterarguments to the assumption (27), and to the proposal that all relative clauses containing VPEs in the ACD construction have undergone (possibly string-vacuous) extraposition.

Direct counterevidence to (27) comes in the form of examples involving a VPE both preceded and followed by terminals beloning to its antecedent at S-Structure. In (31a), the VPE is contained within a free relative that functions as the goal object in the double object construction; its antecedent includes the preceding verb and the following NP (the theme argument of *gave*): [6] ACDs

involving a VPE contained within the accusative subject in ECM constructions are also held to be acceptable (in disagreement with Baltin - cf. also Hornstein 1994:459):

(31)a. John **gave** [whoever he could _] **two dollars**
b. ? John **believes** [everyone I do _] **to be a genius**

Larson & May proceed to argue that QR alone makes the right distinctions among possible and impossible ACDs. In exactly those cases in which QR raises a constituent containing the VPE out of its antecedent VP before LF. The accusative subject in the infinitive complement (31b), for example, contrasts with the nominative subject of it finite counterpart, in that the latter may not host an ACD taking the VP headed by *believe* as its antecedent:

(32) * John **believes** [everyone I do _] **is a genius**

It is claimed that while the accusative subject may undergo QR to a position above the higher verb, the nominative subject is unable to leave its own clause. This claim correlates with the possibilities for scopal interactions between quantifiers: while the accusative subject in (33a) may take scope over the subject of the higher clause, the inverse scope reading is not possible in (33b):

(33)a. some student believes every teacher to be intelligent ok: $\exists\forall$ / $\forall\exists$
b. some student believes that every teacher is intelligent ok: $\exists\forall$ / * $\forall\exists$

2.4 Hornstein (1994)

In the context of the Minimalist framework of Chomsky (1993), Hornstein (1994) proposes what is essentially a variant of the QR account of ACD resolution. Following Larson & May (1990), Hornstein rejects (27), assuming that the sole filter on antecedent containment holds of LF-representations, and concluding that ACDs provide evidence for phrasal movement in the covert syntax. Disagreement arises as to the nature of that movement process. Instead of QR - movement to an A'-position - Hornstein argues that it is A-movement of objects at LF that resolves antecedent-containment in ACDs.

In Chomsky's (1993) proposal, direct objects in English, which stand in VP at S-structure ("Spell-Out"), must raise at LF to the specifier of a functional projection ("object agreement" - AGo) to check Case and agreement features. Hornstein seeks to establish that this assumption provides the basis for an account of ACD resolution without recourse to QR. The (simplified) derivation for (2d) is given in (34):

(34)a. John *Past* [$_{VP}$ met everyone that Mary did [$_{VP}$ e]]

b. [$_{IP}$ John *Past* [$_{AGoP}$ **[$_{DP}$ everyone that Mary did [$_{VP}$ e]]** AGo [$_{VP}$ *t* met *t*]]]

c. [$_{IP}$ John *Past* [$_{AGoP}$ [$_{DP}$ everyone that Mary did **[$_{VP}$ *t* met *t*]]** AGo **[$_{VP}$ *t* met *t*]**]]

Assuming an LF-copy approach to VPE, the Spell-Out representation is (34a). The direct object containing the VPE is in VP. Covert A-movement places the direct object DP in the specifier of AGo (34b). LF-copying may now substitute the main VP for the VPE in the relative clause of the raised object, no regress arises. [7] (Other technical assumptions on which this account depends are left aside here. [8]).

Hornstein cites conceptual and empirical reasons for rejecting the QR-based solution to the ACD issue. Firstly, within the Minimalist framework, given the hypothesis that all movement processes are driven by purely formal (morphological) rerquirements, it is not clear if QR even

exists. Aside from wh-movement, there is no apparent morphological requirement to which QR may be the response. Secondly, evidence independent of the ACD phenomenon that is used to support the postulation of QR - i.e. relative scope facts; in particular, scope inversion phenomena - does not require the assumption of QR (i.e. there are alternative semantic accounts for quantifier interpretation, including relative quantifier scope, that do not depend on qunatifier raising in the syntax).

Even if QR is adopted as the means for handling scope facts, this alone do not suffice to ensure that QR also provides for ACD-resolution. The latter depends on the further assumption that relative clauses are pied-piped under QR at LF. As Hornstein notes, the scope facts would be accounted for under the assumption that it is the quantifier alone that raises under QR, i.e. that further material is not pied-piped. That covert NP-movement, on the other hand, must pied-pipe the whole DP, Hornstein takes as given. [9]

Beyond this, there remain worries about the coextension of QR and ACD resolution. Appositive relatives modifying proper names in object position, for instance, do permit ACDs (cf. (18) in section 2.1 above), despite the assumption that names do not undergo QR. Baltin's Boundedness Restriction is another case where scope facts and ACD facts diverge, which remains unexplained in the QR approach.

The A-movement solution is claimed to capture all facts discussed so far, hence to be superior to previous accounts. Correct predictions are generated concering cases where QR targets and ACDs diverge: while names do not undergo QR, they do undergo LF-A-movement, like any other DP. Hence, the occurrence of ACDs in appositive relatives is expected. A-movement also differs from QR in that it is clause-bound; this fact can be used to explain the Boundedness Restriction observed by Baltin, without recourse to extraposition. Consider (24a), repeated here:

(24) a. Who thought that Fred read how many of the books that Bill did?

At LF, the phrase headed by *how many* undergoes LF A-movement to the AGo-projection immediately dominating *read* (35b). Even if there is an LF-operation of QR that raises *how many* to the specifier of the matrix CP (35a.) (thus fixing its scope), that operation does not pied-pipe the relative clause. The VPE thus remains inside IP2 at LF (35b.), and is unable to receive a wide scope reading (i.e. to take VP1 as its antecedent):

(35) a. [who$_i$ - **how many$_j$** C° [$_{IP1}$ t_i [$_{VP1}$ thought that IP$_2$]]]

b. [$_{IP2}$ Fred I° **[t_j of the books that Bill did [$_{VP}$ e]]$_k$** [$_{VP2}$ read t_k]]]]]

The only case where an ACD is able to take wide scope - i.e. where an VPE contained within an embedded clause is able to take the VP containing that clause as its antecedent - is where an A-movement process raises the DP containing the VPE out of the embedded clause. This happens in ECM constructions, in Chomsky's (1993) account. Consider (36): at S-structure, the embedded subject is contained within the clausal complement of *believe*:

(36) John **believes** [everyone I do _] **to be a genius**

At LF, however, this phrase raises to the AGo-projection immediately dominating *believe*, to check accusative Case (37). NP-movement brings the VPE out of the matrix VP, thus permitting that VP to act as antecedent:

(37) [$_{IP}$ John [$_{AGoP}$ **[$_{DP}$ everyone that I do [$_{VP}$ e]]** AGo [$_{VP}$ *t* believe [*t* to be a genius]]]]

The line is drawn, correctly, between accusative subjects (undergoing NP-movement at LF) and nominative subjects in finite complements to believe, which do not raise to an A-position outside their clause (36) vs. (32). Hence, the LF A-movement approach predicts VPEs which are contained within their antecedents at S-structure, to be permitted precisely where LF-A-movement undoes the containment configuration. Thus cases like (36), counterexamples to Baltin's extraposition account, fall out correctly.

While the A-movement proposal seems to make more accurate predictions than the QR-proposal with respect to wide/narrow scope readings and names, it may face other problems that arise from it being too restrictive in its choice of target. A VPE may give rise to an ACD configuration not only when contained in a relative modifier of a direct object, but also when contained in a PP-complement or an adverbial, categories which are not normally considered to be targets for A-movement: [10]

(38) a. John [talked [to everyone who Mary did _]]
b. John [recited his lines [in the same way that Bill did _]]

Hornstein suggests that the former case be handled in terms of A-movement of the DP out of PP into a higher agreement projection. In other words, such examples are used to motivate an extension of the LF-A-movement hypothesis to DPs in selected PPs. Examples involving adverbials at the right edge of VP (38b) are assumed not to constitute ACDs even at S-structure: it is assumed that the base-generated position is right-adjoined to (hence higher than) the antecedent VP.

Both of these proposals are problematic, but I do not dwell on them here, since there is a far more serious problem facing Hornstein's proposal, and the other approaches reviewed so far.

3. Wide scope ACDs

3.1 A new problem

As noted in the previous sections, a VPE in a relative clause modifying the subject of a finite complement may not take a wide scope reading (39a) (cf. (32) above). However, as Lasnik (1995) observes, citing R. Tiedeman, wide scope VP-ellipsis is possible in a relative modifying the subject of a finite complement clause - if the relative is extraposed (39b): [11]

(39) a. * John believes [everyone I do _] is a genius
b. ? John believes everyone is a genius [that I do _]

Similar pairs involving comparatives can also be constructed:

(40) a. * John claimed that more people than you did _ were stupid
b. John claimed that more people were stupid than you did _

The wide scope interpretation of the VPE in (39b) and (40b) is quite clear: the matrix VP is reconstructed into the VPE, so that they mean the same as the sentences (41) and (42) (with or without extraposition), respectively: [12]

(41) a. John believes everyone is a genius that I believe is a genius.
b. John believes everyone that I believe is a genius, is a genius.

(42) a. John claimed that more people were stupid than you claimed were stupid.
b. John claimed that more people than you claimed were stupid, were stupid.

With VPE, extraposition of the relative or comparative clause has a dramatic effect on acceptability, which none of the analyses discussed in the previous section is in a position to explain. The problem posed by the contrast is different in each case.

3.1.1 The QR account

For the QR account (May 1985), Larson & May (1990), the question raised is: why should extraposition of a modifier affect the ability of the head NP to undergo QR? In order for the QR account to be able to explain the wide scope reading of the VPE in (39b) and (40b), it must be assumed that the nominative can undergo QR, pied-piping its modifier, to a position above the matrix VP. [13] In order to account for the ungrammaticality of (39a) and (40a) under the wide scope reading, QR must be unable to raise the nominative to a position above the matrix VP. Either QR is prevented from applying, or an independent explanation must be found. Larson & May (1990:108) explicitly claim that the reason for the ill-formedness of (39a) is that the subject of a finite complement clause may not undergo QR into the matrix sentence ("quantified subjects are clause-bound in finite sentences"). Suppose ECP is the constraining principle. The application of ECP should be blind to whether or not a raised QNP contains a relative clause. How extraposition of a modifier could neutralize the ECP in this case, is completely mysterious. [14]

3.1.2 The extraposition account

For Baltin (1987), the problem is a different one. Given the S-structure constraint (27) of section 2.2 above, it appears at first glance that the paradigm (39)-(40) actually supports Baltin's approach. (39a) and (40a) clearly involve antecedent containment at S-structure, regardless of what happens at LF. It seems as if extraposition in (39b) and (40b) has the effect of removing the VPE from the containment configuration, exactly what Baltin claims. However, if extraposition obeys the Right Roof constraint, the relative clause in these examples can be adjoined no higher than the CP of the complement, where it will still be contained within the VPE-antecedent (the matrix VP), in violation of the S-structure contraint (27). In other words, the S-structure (43a), which obeys Right Roof, violates (27); while (43b) satisfies (27) but violates Right Roof:

(43)a. John [$_{VP}$ believes [$_{CP}$ [$_{CP}$ that everyone is a genius $_{CP}$] <u>that I do VP</u> $_{CP}$] $_{VP}$]
b. John [$_{VP}$ [$_{VP}$ believes [that everyone is a genius] $_{VP}$] <u>that I do VP</u> $_{VP}$]

The Right Roof Constraint plays a crucial role in deriving the Boundedness restriction discussed in connection with example (24) above. (39b) and (40b) are clear counterexamples to the Boundedness Restriction. Baltin assumed that the Boundedness restriction holds without exception. If it were the case that the Right Roof constraint were not operative in these examples, then (39b) could be assigned the S-structure (43b). Thus, the idea may seem worth exploring, that these sentences counterexemplify the Boundedness restriction for the very reason that extraposition is not constrained by the Right Roof Constraint. However, I do not follow this path here, for the following reasons, empirical and conceptual.

Firstly, it is doubtful whether the Right Roof effect is cancelled (i.e. that the "roof" is extended) in these constructions: the word order test clearly indicates that the Right Roof is still in place - no matrix adverbial may separate the extraposed relative from the rest of the complement clause:

(44)a. * John said that everyone is a genius yesterday that I did
b. * John claimed that more people were stupid when I asked him, than you did

Secondly, the very nature of the approach - establishing the satisfaction of an S-structure constraint (Baltin's (27)) by recourse to a rightward movement analysis of extraposition - conflicts with leading hypotheses of current theory. Within the minimalist framework, S-structure (in the

sense of earlier models) has been abandoned as a level of representation, so that it becomes desirable to reformulate constraints apparently holding of S-structure either as interface conditions holding at LF or PF, or as properties of derivational operations. Assuming Kayne's (1994) Linear Correspondence Axiom (or Chomsky's (1995) reinterpretation of it), linear order (precedence) loses its status as a primitive, becoming a property of PF derived from hierarchical relations (asymmetric c-command). One consequence is that UG does not permit rightward adjunction, nor rightward movement (adjuncts precede their hosts, and moved categories precede their traces, after linearization). Hence, extraposition cannot be analyzed as rightward movement of CP; and whatever the explanation is for Right Roof constraint effects (or their absence), these cannot be attributed to a bounding condition on rightward movement since this does not exist. For present purposes, I assume that extraposition is "stranding" of CP under leftward movement of the "head" of the extraposed clause, as in Kayne (1994) or Wilder (1995b). In that case, the extraposed clauses in (39b) and (40b) are actually located within the VP of the complement clause, as in (45):

(45) a. John $[_{VP}$ believes $[_{CP}$ that [everyone]$_j$ $[_{VP}$ is a genius [t_j that I do VP]]]]
b. John $[_{VP}$ claimed $[_{CP}$ that [more people]$_j$ $[_{VP}$ were stupid [t_j than I did VP]]]]

3.1.3 The A-movement account

For Hornstein (1994), the problem posed by wide scope ACDs is as simple as it is unresolvable. Hornstein's main goal is to establish the ACD phenomenon as a source of independent evidence for the "minimalist" conception of covert A-movement of objects in English. To do this, he must make the A-movement account of ACDs into a persuasive alternative to the QR account. His main hope of success in this task lies in establishing that the possible wide scope ACD readings are coextensive with the domain of LF A-movement. One contribution to that goal is provided by the case of Baltin's boundedness restriction, unexpected under the QR account. But in addition, the A-movement account depends crucially on the contrast between (46a) and (46b):

(46) a. John believes everyone I do _ to be intelligent
b. * John believes everyone I do _ is intelligent
c. John believes everyone is intelligent that I do _

Adding (46c) (=39b) to the paradigm ruins the picture. (46c) should be as ungrammatical as (46b). It is hardly plausible to assume LF A-movement of nominatives into a higher clause; even less so, if that operation depends on the presence of a relative clause that is extraposed.

In fact, Hornstein is aware of the problem. In a footnote, he addresses the issue of wide-scope ACDs and seeks to deny the facts:

> "May (p.c.) notes cases that are problematic for the above claims:
>
> (vi) Max thinks the yacht is longer than Oscar does
>
> (vi), he says, has the interpretation '...longer than Oscar thinks it is' [...] These cases all strike me as unacceptable. However, it is possible that more is going on here than present analysis accounts for [...] if matrix ACD readings are regularly possible from null VPs in embedded clauses, then this would tell against the minimalist approach sketched here. I will proceed as if this is not the case." (Hornstein 1994:462, Fn. 9 - my emphasis).

I do not think these facts are deniable.

3.1.4 Further properties of wide scope ACDs

What is the extent of the problem? Wide scope ACDs are not freely constructible, one has to hunt them. As Hornstein notes (fn. 9), there are unclear restrictions on matrix predicates and tenses:

(47) a. ?? John has heard that everyone is a genius [that I have _]
b. ?? John said that Bill was reporting everything [Sam did _]
c. * John doubted that the yacht was as long as she did
d. * John insists that my yacht is flashier than you do

This may help to explain why they have escaped notice so far (in addition, Baltin's examples illustrating the Boundedness restriction (24) might also have served to distract attention from them). Nevertheless, there are enough good examples to establish that this is indeed a productive phenomenon which cannot be simply dismissed.

The availability of a narrow scope reading can make a wide scope reading barely, if at all, accessible. For (48a), the reading (48b) obscures (48c), for example:

(48) a. John thought that Mary bought more books than Bill did
b. John thought that Mary bought more books than Bill did *buy*
c. John thought that Mary bought more books than Bill did *think that Mary bought*

In constructing examples, it is useful to be aware of, and control for, this "local antecedent" effect. In (49), where the embedded subject in the antecedent does not match the subject of the VPE in animacy, but the matrix subject does, the wide reading is at least as accessible as the narrow reading:

(49) a. John thought that the fire destroyed more books than Bill did
b. John thought that the fire destroyed more books than Bill did *destroy*
c. John thought that the fire destroyed more books than Bill did *think that it destroyed*

Prosodic factors can be important. A wide scope reading can be facilitated by parallel accent on the subject of the VPE and the subject of antecedent VP, with the subject of the complement clause in the antecedent deaccented (typographically: deaccenting in small pitch, accenting with ´ on vowel). In fact, wide and narrow readings for the VPE can be disambiguated prosodically in this way (cf. also discussion of deaccenting and VPE in Tancredi (1992)):

(50) a. Jóhn thought that the fire destroyed more books than Bíll did (wide only)
b. John thought that the fíre destroyed more books than Bíll did (narrow only)

Prosodic manipulation cannot save the wide scope reading in Baltin's example (51) (= (24) in section 2.2), however: it can only exclude the narrow scope reading as well:

(51) * Whó thought that Bill read how many of the books that Máry did

Two further properties of wide scope ACDs are relevant to the proposal to be made in section 3.2. Firstly, they are only found in right-peripheral position (cf. Hornstein 1994: fn.9):

(52) a. John said that there were more people than Mary did _
b. John said that more people were there than Mary did _
c. * John said that more people than Mary did _were there

The function of the head of the relative or comparative containing the ACD (subject, object, predicate, etc) does not seem to matter, as long as the VPE is in final position; i.e. if the wide scope reading is possible at all, then only if the VPE is in final position. Extraposition does not seem necessary (53a), (54a) - although it may operate string-vacuously, as in Baltin's proposal - unless the relative or comparative is not in final position (53b), (54b):

(53) a. John thinks that this will cause more harm than Mary does
b. * John thinks that this will cause more harm than Mary does to the proposal
c. John thinks that this will cause more harm to the proposal than Mary does

(54) a. John thinks the situation has got worse than Mary does
b. * John thinks the situation has got worse than Mary does through this
c. John thinks the situation has got worse through this than Mary does

Secondly, wide scope ACDs are found in relative clauses and comparative clauses, but they are not licensed in adjunct clauses (cf. also Hornstein 1994). A sentence like (55) is ambiguous with respect to which verbal projection the temporal adjunct modifies:

(55) John said that Mary would arrive when Peter did.

Adding the wide - narrow scope dimension for the VPE (and ignoring the possibility for an external antecedent), we expect the sentence to have four possible readings (56). In fact, it only has three.

(56) a. [[John said that Mary would arrive] when Peter did *arrive*]
b. [[John said that Mary would arrive] when Peter did *say that Mary would arrive*]
c. John said that [[Mary would arrive] when Peter did *arrive*]
d. * John said that [[Mary would arrive] when Peter did *say that Mary would arrive*]

If the temporal clause modifies *said*, then the VPE may take either the embedded (56a) or the matrix VP (56b) as antecedent. If the temporal clause is understood as modifying *would arrive*, then the VPE may only take the embedded VP as antecedent (56c). The intended reading for (56d) can be paraphrased as (57):

(57) John said at time t, that Mary would arrive at time t', t' being such that Peter said (at time t'') that Mary would arrive at t'.

This reading is utterly impossible for (55) - although it represents a possible interpretation of (58), the 'undeleted' counterpart to (55):

(58) John said that Mary would arrive when Peter said that she would arrive.

3.2 A new proposal

The evidence from section 3.1 speaks strongly against the proposal of Hornstein (1994): LF A-movement is too restrictive to be able to account for limitations on wide scope readings of ACDs. It is consistent with an A'-movement (i.e. QR) account, assuming that restrictions on QR are relaxed to permit LF-extraction of certain phrases from finite complement clauses. But that account is then too liberal to account for the distribution of well-formed and ill-formed ACDs - in particular, the contrasts involving extraposition. These in turn motivate a proposal in the spirit of Baltin's. There is not only an LF condition barring antecedent containment, which is voided by QR in certain cases; there is also an independent surface condition, which is voided by extraposition in certain cases.

That is, in addition to the condition (59) (= (3) of section 1.1 above) holding at LF, there is a separate principle (60), holding at PF:

(59) A VP-deletion site may not be contained in its antecedent **constituent** at LF

(60) A VP-deletion site may not be contained in its antecedent **string** at PF

In view of the considerations on the status of extraposition mentioned above, the surface condition cannot be stated in terms of hierarchical containment (as Baltin's condition is). Extraposition only serves to ensure that the VPE is not contained within the string corresponding to its antecedent; it does not remove the VPE from the hierarchical constituent of its antecedent. Instead, the condition is formulated in terms of "string-containment". What (60) excludes is any PF-string in which two terminals α, ß, belonging to the antecedent VP, are linearized with respect to the VPE such that α precedes the ellipsis site and ß follows the ellipsis site:

(61) * α < E < ß , where α, ß are elements of the antecedent of E

If syntactic objects (including S-structure or "Spell-Out" trees) do not contain precedence information, then (60) cannot hold of S-structure. The locution "at PF" is however intended to be vague, referring to some stage within the derivation from syntactic component to the "PF interface", in the sense discussed in Chomsky (1995). The earliest stage at which (60) can be assumed to hold is the level at which representations are linearized. The problem with assuming that (60) holds at the interface is of course that it refers to a phonetically empty element, by hypothesis not "visible" at the interface. [15]

(59) and (60) are each necessary independently of the other. The impossibility of a wide scope reading for a VPE in an embedded adjunct clause provides evidence for the necessity of the ban on structural containment (59) - since the adjunct appears string finally, the ban on string containment (60) would be insufficient. The adjunct - relative clause asymmetry (a VPE in a final embedded relative clause can receive wide scope interpretation) supports the assumption of the QR-escape hatch (assuming that an extraposed relative clause is "reconstructed", so that it can be pied-piped by LF-movement). The extraposition / intraposition asymmetry found with embedded relative clauses then provides the evidence for the existence of the additional surface filter (60). This filter must be formulated in terms of linear containment; hierarchical containment would be too strong (unless extraposition is rightward movement and the Right-Roof Constraint is lifted in the relevant cases).

Assuming that adjunct clauses do not undergo QR, wide scope VP-deletion in a final embedded adjunct (56d) violates (59), but not (60). A wide scope VPE in an embedded intraposed relative (62) violates the surface filter (60), but not (59) (assuming QR). A wide scope in an embedded final relative (63) violates neither.

(62)a. John believes everyone [that I do _] is a genius

b. *PF* * John **believes** everyone that I do (*believe (x) is a genius*) **is a genius**
c. *LF* [everyone that I do (*believe (x) is a genius*)] [John **believes x is a genius**]

(63)a. John believes everyone is a genius [that I do _]

b. *PF* John **believes everyone is a genius** that I do (*believe (everyone) is a genius*)
c. *LF* [everyone that I do (*believe (x) is a genius*)] [John **believes x is a genius**]

Additional evidence for the analysis of (62) in terms of LF-convergence but PF-crash comes from typical native speaker reaction to such examples - they are held to be "interpretable, but the word order is wrong". This reaction contrasts strongly with the reaction to embedded adjunct examples where a narrow scope reading is independently excluded: the wide scope reading is felt to be simply impossible ("the example is uninterpretable"). Their status is comparable to that of cases like (64), under the ACD reading:

(64) * John said that Mary did.

There is one issue not yet addressed. What to do with "genuine" string contained ACDs like (65), which - apparently - violate (60)?

(65) a. John gave whoever he could _ two dollars
b. ? John finds everyone I do _ easy to work with.
c. ? John believes everyone I do _ to be a genius

These cases are examined in section 6. below, where it is proposed that they do not involve antecedent-containment at all, hence do not pose any problem for (60).

4. Hierarchical containment and QR

In this section, two questions concerning the LF-side of the account are addressed: (i) what type of constituents undergo QR out of finite complement clauses (i.e. license wide scope readings for VPEs)? (ii) how far do such constituents move (what is the landing site of "long" QR)?

4.1 What undergoes QR?

Taking wide-scope VP-deletion as an indicator for QR, it turns out that QR can affect argument DPs of various types which have been claimed by other authors to be QR targets. Additionally, compared predicates must be assumed to be able to undergo QR. The VPE in each case is contained within a relative clause modifying the head noun, or, in the case of compared noun phrases and adjectival predicates, within the comparative clause selected by the degree word (e.g. the *than*-clause selected by *more*, the *as*-clause selected by equative *as* or by *the same*, etc.):

The possibility for argument noun phrases that are modified by relative clauses also seems to depend on the nature of the determiner. Wide scope VPEs occur most felicitously in relatives modifying DPs headed by "strong" determiners (*every*, *each*, *most*, etc - cf. (66)); definites, especially in combinations such as *the same* N, or *the very* N (67); and partitive DPs headed by "weak" determiners (68) ("strong" and "weak" in the sense of Milsark (1977), Diesing (1992)). In each case, the extraposition-intraposition contrast is evident, even where the extraposed variant itself is only marginally acceptable:

(66) a. John thought that most people were there that Mary did.
b. * John thought that most people that Mary did were there.
c. John thought that each proposal should be accepted that Mary did.
d. * John thought that each proposal that Mary did should be accepted.

(67) a. John said that the (very) same problem would arise that/as Mary did
b. * John said that the (very) same problem that/as Mary did would arise

(68) a. John said that many/none of the problems would arise that Mary did
b. * John said that many/none of the problems that Mary did would arise

All these DP types have been claimed by May (1985) and/or Diesing (1992) to undergo QR.

Examples involving relatives whose DP is headed by a weak determiner that is not a partitive, are marginal, bordering on the unacceptable:

(69) ??? John said that a / one / many / few / no problem(s) would arise that Mary did

In her discussion of indefinites, Diesing (1992:Ch.3) distinguishes a presuppositional and a non-presuppositional reading of indefinite DPs headed by weak determiners, claiming that presuppositional indefinites undergo QR, while non-presuppostional indefinites do not. Partitive indefinites are invariably presuppositional. Using ACDs as a diagnostic, the contrast between the two readings comes out clearly in simple examples (Diesing 1992:71):

(70)a. ?? I read few books that you did.
b. I read few of the books that you did

Under certain circumstances, (70a) can be rendered acceptable - namely when the context facilitates a presuppositional reading of the DP (in which there a salient set of books over which the determiner ranges, as in the sole possible reading of (70b)). As Diesing notes, (70a) is "unquestionably ungrammatical in the case of the cardinal (nonpartitive) reading". [16]

Plausibly, the difficulty with (69) can be attributed to the difficulty in accessing the presuppositional reading for the DP that would accompany the QR needed to resolve the VPE, just as in (70a). Additional support for this speculation can be derived from considering the weak determiner *some*. As Milsark (1977) has pointed out, the non-presuppositional ('cardinal') reading of DPs headed by *some* typically involve an unstressed variant *sm*; while stressing the determiner triggers a presuppositional reading. A wide scope VPE associated with *sm* is impossible, its counterpart with stressed *some* is relatively good:

(71)a. * John said that sm problems would arise that Mary did
b. ? John said that SOME problems would arise that Mary did

Already on the basis of simple examples like (72a), Diesing is led to suggest that comparatives undergo QR. Cases involving compared DPs seem to yield the best examples of wide scope ACDs (72b):

(72)a. I read more books than you did.
b. John said that more problems would arise than Mary did.

Compared embedded predicates also license wide scope ACDs:

(73) John said that his yacht was longer than Bill did

This indicates that it is the comparative morpheme, as a functional head governing the predicative AP analogous to the way a determiner governs NP, that licenses QR of its phrase, pied-piping the predicate along with the *than*-clause.

It is even possible for a compared adverbial modifying a verb in a finite complement to license a wide scope VPE in its *than*-clause:

(74) John said that it would rain more often than Mary did

(74) has a reading in which *more often...* modifies the embedded verb *rain*, while the *than*-clause contains a VPE taking the matrix VP *said that...* as its antecedent; which contrasts with the impossibility for a non-compared embedded temporal adverbial to contain a wide scope VPE (cf. (56) above):

(75) John said that it would rain when Mary did

The contrast (74) vs. (75) provides further support for the hypothesis that the comparative morpheme licenses QR of its phrase.

One problem for the QR approach noted above concerns that apparent possibility for appositive relatives modifying names to contain ACDs, as in (76a) - the problem being that names are not QR targets. However, appositive relatives in complement clauses do not support wide scope ACDs. (76b), which would have the reading (76c) if the DP headed by the name *Peter* could undergo QR, is deviant:

(76)a. John met Peter, who Mary did, too.
b. * John said it would please Peter, who Mary did, too
c. John said it would please Peter, who Mary said it would please, too

In section 2.1 above, it was suggested that examples like (76a) may involve pseudo-gapping, hence not bear on the issue of antecedent containment and QR at all. That solution is further supported by the fact that (77b) - with the intended reading (77a) matches the ill-formedness of (76b):

(77)a. John said it would please Peter, and Mary said it would please Bill.
b. * John said it would please Peter, and Mary did, Bill.

4.2 What is the landing site of "long-distance" QR?

The existence of wide scope ACDs motivates the claim that QR can raise constituents out of finite complement clauses into the next clause up. If this is so, then it is expected that the raised constituent will enter into scopal interactions with other quantified expressions originating in the higher clause. This expectation is not met, however. The questions that arise have to do with the analysis of scopal interactions in general, which I do not attempt to solve here. Instead, the problem will simply be laid out and left for further research.

It is already clear for simple cases (not involving wide scope ACDs) that a quantified expression originating in a finite complement clause does not interact scopally with any quantified argument (subject, direct or indirect object) in the matrix clause. The embedded subject in (78) cannot take scope over an indefinite in the matrix:

(78)a. Someone believes that everyone is a genius ok: $\exists\forall$ / * $\forall\exists$
b. John told someone that everyone is a spy

Addition of an extraposed relative clause modifiying the embedded subject, and containing a wide scope VPE does not alter this fact - the phrase headed by *everyone* cannot take scope over *someone* in (79):

(79)a. Someone believes that everyone is a genius that Mary does ok: $\exists\forall$ / * $\forall\exists$
b. John told someone that everyone is a spy that Mary did

In order to escape antecedent containment at LF, some phrase containing the wide-scope VPE (by hypothesis, the DP headed by *everyone* in (79)) must minimally gain scope over the constituent

(V') containing the matrix verb and its complement (the finite clause). At the same time, it must be ensured that rules of interpretation do not permit *everyone* to take scope over *someone*. In the theory of May (1985), quantified expressions may only move to positions adjoined to VP or IP (leaving aside movement to Spec,CP). Supposing that the phrase headed by *everyone* is adjoined to VP, and that the matrix subject *someone* is higher up the tree (e.g. in Spec,IP or adjoined to IP), an LF like (80a) is derived:

(80) a. someone$_y$ [everyone$_x$ that Mary *believes* [*x is a genius*]] [y **believes** [**x is a genius**]]
b. $\exists y\ \forall x$ (believe(m) (genius(x))) $\rightarrow$ (believe(y) (genius(x)))
c. "There is a *y* such that for all *x*, if Mary believes *x* is a genius, then *y* believes *x* is a genius"

From (80a), the formula (80b) would be derivable, which correctly represents the interpretation of (79a).

The problem posed by (80a) is that in current theories of scopal interactions, the matrix subject *someone* would be able to "get into the scope" of the raised DP, either by a rule of syntax such as quantifier lowering, or a rule of interpretation. In May's theory, for example, once an expression is adjoined to some VP, rules of interpretation enable it to interact scopally with any other expression adjoined either to that VP or to the IP immediately dominating. The lowest adjunction site consistent with resolving antecedent containment in (79) would be the VP headed by *believes* (or *told*). If the phrase headed by *everyone* has adjoined to that VP, then nothing would prevent it from being able to take scope over *someone*.

It does not seem possible however, to maintain that the landing site of QR is some lower position from which such scopal interactions would not be possible. [17] If QR in (79a) were is restricted to the finite compement clause, an LF would be generated in which the phrase headed by *everyone* would be in the scope of the matrix verb *believe*. The reading expected would be (81):

(81) a. $\exists y$ believe(y) ($\forall x$ (believe(m) (genius(x))) $\rightarrow$ (genius(x)))
b. There is a *y* who believes that for all *x*, if Mary believes *x* is a genius, then *x* is a genius.

Quite apart from the fact that this assumption about the landing site of QR would lead to an antecedent-containment configuration at LF, the reading (81) - which is available for the 'undeleted' examples (82), is <u>not</u> available for (79a):

(82) a. Someone believes that everyone that Mary believes is a genius, is a genius.
b. Someone believes that everyone is a genius that Mary believes is a genius.

The facts about wide scope ACDs suggest instead that the account of scopal interactions may be in need of revision.

There is one further constraint on wide scope ACDs that may be relevant to the issue of absent scopal interactions. The antecedent to the VPE in a wide scope ACD is restricted to the first VP dominating the finite clause containing the quantifier phrase. While (83a,b,c) are all possible, (84) is not:

(83) a. John thinks that more trees died than Mary does _
b. John thinks that more trees seem to have died than Mary does _
c. John thinks that it seems that more trees have died than Mary thinks it does _

(84) * John thinks that it seems that more trees died than Mary does _

In each of (83a,b,c), the constituent (*more trees...*) that undergoes QR is located in the clause below the VP-antecedent to the VPE in the comparative clause:

(85)a John [thinks [that more trees died than Mary does *think died*]]
b. John
[thinks [that more trees seem to have died than Mary does *think seem to have died*]]
c. John thinks that it
[seems [that more trees have died than Mary thinks it does *seem have died*]]

In (84), the constituent (*more trees...*) that undergoes QR is located in the second clause below the VP-antecedent to the VPE in the comparative clause:

(86) * John
[thinks [that it seems [that more trees died than Mary does *think that it seems died*]]]

This paradigm suggest that "long-distance" QR is strictly local, in the sense that it may not raise a constituent further than the VP immediately dominating the first finite clause containing it. If this is correct, then the other facts discussed in this section would fall out, if the theory of scopal interaction is designed in such a way that a phrase raised from a finite complement and adjoined to VP cannot interact scopally with the arguments of the verb heading that VP.

5. German comparatives

Cross-linguistic support for the account of wide-scope VP-ellipsis sketched here is not easy to come by, since many languages do not show VP-ellipsis in the English pattern at all, or only to a much more restricted extent. However, a similar contrast to the one discussed in section 3. is found in an elliptical comparative construction that is not confined to English.

Consider first the English examples (87)-(88).

(87) Peter thought about it more often than Mary

(88) a. Peter thought that more trees had died than Mary
b * Peter thought that more trees than Mary had died

Assuming that the complement to *than* in (87), superficially a bare DP, is underlyingly clausal, this clause undergoes ellipsis which deletes all daughters of I', leaving the subject as the sole remnant. (87) thus has an abstract representation like (89):

(89) Peter **thought about it** more **often** than Mary *thought about it x-often*

Assuming further that I'-ellipsis, like VP-ellipsis, is subject to the PF-constraint (60), it is clear that (60) is satisfied by (89).

Where the elliptical *than*-clause is associated with an embedded nominative, wide-scope interpretation is possible only when the comparative clause is extraposed (89a). (89b) only has the nonsensical interpretation "more trees than just Mary died", i.e. that Mary is a dead tree. The contrast can be accounted for if the examples (88) have representations as in (90).

(90)a. Peter
thought [that [more **trees** _] **had died** than Mary *thought that x-many trees had died*]

b. * Peter
thought [that [more **trees** than Mary *thought x-many trees had died*] **had died**]

(90b) violates the PF-constraint (60), while (90a) does not. (90a) also avoids antecedent-containment at LF, assuming that the extraposed elliptical *than*-clause is reconstructed into the comparative DP, yielding a constituent *more trees than Mary thought that x-many trees died* which then undergoes QR to a position above the matrix verb *thought*. Given that the same QR option is available for (90b), the ungrammaticality of this example can be taken as support for the hypothesis that I'-ellipsis is subject to (60).

Unlike VP-ellipsis, I'-ellipsis in comparative clauses is found in other languages, for example, German. The contrast in (88) can be reproduced for German, as illustrated in (91): [18]

(91)a. * weil Peter meinte, daß es öfter als ich geregnet hat
since Peter said that it more-often than I rained has

b. weil Peter meinte, daß es öfter geregnet hat als ich
"since Peter said that it rained more often than I (did)"

The German comparative morpheme selects the complementizer *als* which heads the comparative clause. [19] This clause may undergo I'-ellipsis, with the result that only the subject DP is overtly expressed. In simple clauses, in contrast to (91), extraposition of the *als*-clause (subject to the Right Roof constraint) is possible but not obligatory, regardless of whether it is elliptical or not, as shown in (92)-(93):

(92) a. ...weil er öfter als ich darüber nachgedacht hat
since he more-often than I there-over thought has
"since he has thought about this more often than I (have)"

b. ...weil er öfter darüber nachgedacht hat als ich

(93) a. ...weil er öfter, als ich darüber nachgedacht habe, darüber nachgedacht hat
"since he has thought about this more often than I have thought about it"

b. ...weil er öfter darüber nachgedacht hat, als ich darüber nachgedacht habe.

The examples (92) have the analysis (94). Neither (94a) nor (94b) violate (60):

(94) a. weil er [öfter als ich *x-oft darüber nachgedacht habe*] **darüber nachgedacht hat**
b. weil er [öfter _] **darüber nachgedacht hat** [als ich *x-oft darüber nachgedacht habe*]

By the same reasoning, the examples (91) are represented as (95):

(95)a. * weil Peter
meinte, daß es [**öfter** als ich *meinte daß es x-oft geregnet hat*] **geregnet hat**

b. weil Peter
meinte, daß es [**öfter** _] **geregnet hat** [als ich *meinte daß es x-oft geregnet hat*]

(95b) does not violate (60), but (95a) does. Assuming that in (95b), the extraposed clause reconstructs into the phrase headed by *öfter*, with the latter undergoing QR to a position above the matrix verb *meinte*, antecedent-containment is avoided at LF. Given that the second step, i.e. QR, is equally available to (95a), the ungrammaticality of (95a) provides support for the claim that (60) applies also in German. [20]

6. VP-ellipsis and string-containment

The account in the previous sections of the contrasts concerning wide scope ellipsis depended crucially on the PF-constraint (60). Now we return to the examples (96) (=65), which were put aside in section 3.2.

(96) a. John gave whoever he could _ two dollars
b. ? John finds everyone I do _ easy to work with.
c. ? John believes everyone I do _ to be a genius

The question is, given that constraint (60) excludes examples with an ellipsis site (properly) contained within the string corresponding to its antecedent in surface order, how do we account for good string-contained ACDs?

The idea I pursue here is that examples of this type are not ACD's at all. Rather, they involve two adjacent deletions, neither antecedent-contained in surface order, one due to Pseudogapping (ellipsis of subparts of VP), the other due to the independent process of Backward Deletion (BWD). The examples in (96) are therefore related to examples like (97), where part of VP in a relative clause undergoes deletion under identity with material belonging to the VP containing the object which that relative clause modifies:

(97) a. John gave [whoever Mary did _ two dollars] three dollars
b. John finds [everyone I do _ hard to work with] easy to work with.
c. John believes [everyone I do _ to be a fool] to be a genius

The claim is that (96) involves the same partial VP-Ellipsis as that in (97). In (96), this partial VPE interacts with Backward Deletion of the right-remnant of the embedded VP, under identity with material belonging to the matrix VP which follows the relative clause in surface order. The examples in (96) thus have abstract representations as in (98):

(98) a. John gave [whoever he could _ ~~two dollars~~] two dollars
b. John finds [everyone I do _ ~~easy to work with~~] easy to work with
c. John believes [everyone I do _ ~~to be a genius~~] to be a genius

Neither deletion site in (98) is properly contained within the string corresponding to its antecedent, in surface order.

Put another way, the claim is that the derivation of (96a) that involves VP-ellipsis (99a) is ill-formed, since it violates (60); while the derivation of the same string involving partial VP-ellipsis and BWD (99b) violates neither (60), nor any other constraint: [21]

(99) a. * John **gave** [whoever he could *give t two dollars*] **two dollars**
b. John **gave** [whoever he could *give t* ~~two dollars~~] two dollars

If this is correct, then the examples (96) have been mis-identified as ACDs arising from VP-ellipsis. What they actually indicate is that independent deletion processes may interact to give rise

to adjacent ellipsis sites. Hence, these examples have no direct bearing on the constraints governing antecedent-containment; instead, they relate to the question of how apparent VP-deletions are to be analyzed in the first place (cf. section 1.3).

To make this idea plausible, it is minimally necessary to show firstly, that in the good examples, the component deletions meet independently motivated constraints holding of each deletion process individually; and secondly, that individual violation of any such constraint leads to ungrammaticality, i.e. a bad derivation. The task is made more complex by the fact that for a given example, there may be more than one derivation to consider (cf. (99)).

In the next section, constraints on partial VPE are discussed. In section 6.2., constraints on BWD are introduced. Interactions between them are discussed in section 6.3.

6.1 Partial VPE (pseudo-gapping)

Attempts establish the nature of constraints on partial VPE are is hampered by the fact that "pseudogapping" examples are generally felt to substandard. This is especially true of examples where a verb is elided, stranding its direct object (the cases of interest here):

(100) ? John read the book and Mary did, the newspaper

Examples where the verb and its arguments (the traditional "V'-constituent") are elided, stranding peripheral adjuncts, tend to be far more acceptable:

(101)a. John read the book today and Mary did, yesterday
b. John put some money in the bank today and Mary did, yesterday.
c. ?* John put some money in the cashbox and Mary did, in the till.

However, the argument-adjunct contrast tends to disappear where remnants are finite clauses or PPs:

(102)a. John wrote that Bill was innocent and Mary did, that he was innocent.
b. John requested that Bill be sacked, and Mary did, that he be retained.

(103)a. If you'll explain this to the pupils, then I will, to their parents
b. John's staying with Mary and Bill is, with Sue

Occasionally, one encounters examples with stranded direct objects that are spotless ((104a) due to Sten Vikner), especially in comparatives:

(104)a. It is harder to persuade the Americans of this than it is, the English.
b. John started his paper at the same time as Mary did, her book.

Here, descriptive generalizations are proposed, drawing on comparative judgments on relevant example types - the aim being to establish where the line is to be drawn between grammatical examples of low acceptability, and genuinely impossible (ungrammatical) cases. The limits to pseudo-gapping so established can then be used in excluding otherwise possible VPE-BWD interactions. (I know of no theoretical account that accurately captures the distribution of generalizations given here).

There are two basic generalization at stake. The first is that no constituent of a VP may be deleted, unless the verb is itself deleted: the minimal VPE target consists of V alone (never of an object alone, for example). Hence the deletion of V is 'obligatory' (105a), the possibility for deletion of further constituents (105b) dependent on deletion of the verb. The second generalization is that pseudogapping may not target a string crossing a DP, PP or finite clause boundary (105c

(105) VPE targets:
a. V
b. optionally, DP, PP, finite CP
c. no subpart of DP, PP, finite CP

Hence, remnants tend to be complete arguments or adjuncts to the deleted verb. (106) illustrates ellipsis of a ditransitive V, together with its goal DP (106a), and its theme NP (106b):

(106)a. He gave them two dollars, and she did, three dollars (*give them...*)
b. ? He gave Peter two dollars, and she did, Mary (*give ... two dollars*)

In small clause and ECM constructions, the main verb may be deleted together with a subpart of its complement - the accusative subject (107a), or the accusative subject and the embedded verb (107b):

(107)a. ? He finds me stupid, and she does, smart (*finds me ...*)
b. ?? He believes them to like beans, and she does, spaghetti (*believe them to like...*)

The impossibility of deleting part of a finite clausal complement is illustrated in (108): [22]

(108)a. * I believe that John is a fool, and she does, is a genius (**believe that John...*)
b. * He believes that they like beans, and she does, spaghetti (**believe that they like..*)

A constraint which plays an important role in section 6.3 is that pseudogapping of a verb that is left-adjacent to a finite clause lacking a complementizer is impossible. This is so even in contexts where 'complementizer deletion' is otherwise possible: [23]

(109)a. * Mary believes (that) Peter is a fool and Bill does, he is a genius
b. Mary believes Peter is a fool and Bill does, that he is a genius

The ban on deletion of a complementizer following a gapped verb is independent of whether the complementizer of the complement of the antecedent verb is deleted (109b). A subcase of this generalization is illustrated in (110a). (110) shows that a verb whose complement contains a subject trace cannot undergo pseudogapping:

(110)a. * someone who he believes is a fool and (who) she does, is a genius.
b. * someone who he believes is a fool and (who) she does, that is a genius.

This fact follows from two conflicting requirements - the presence of the complementizer required by the constraint on pseudogapping induces a violation of whatever underlies the "that-trace filter". Where the latter is inoperative (object extraction), the complementizer may be overt, and pseudogapping is possible:

(111) someone who he believes is a fool and (who) she does, that we should hire.

Finally, recall from section 1.3 the question-mark raised, by the possibility for partial VPE, over the antecedent-contained status of ACDs. Given that discontinuous gaps can arise, it may be that apparent VP-gaps are actually a sequence of adjacent, smaller gaps. Then the antecedent for

each gap would be a correspondingly smaller constituent, possibly never containing its dependent gap (in surface structure).

An argument that genuine antecedent-containment arises at S-structure comes from considering wide scope ACDs. The fact that CP cannot be 'cut into' can be used to deny the possibility of analyzing the deleted string corresponding to CP as a collection of individual deletion sites. If the deletion site in (112a) is dismantled into its smallest independently motivated parts, the deleted CP is still a single deletion site (112b); the analysis (112c), for instance, not meeting the "major constituent" condition on VPE:

(112)a. John thinks everyone is intelligent that Mary does __
b. ... that Mary does [*think*] [*t is intelligent*]
c. * ... that Mary does [*think*] *t* [*is*] [*intelligent*]
d. John [$_{V2}$**thinks**]
[$_{CP2}$ everyone **is intelligent** that Mary does [$_{V1}$*think*] [$_{CP1}$ *t is intelligent*$_{CP1}$]$_{CP2}$]

Assuming the Right Roof Constraint to be valid, then that deletion site (CP) must be hierarchically contained in its antecedent (at S-structure) (112d).

6.2 Backward Deletion

The Backward Deletion (BWD) operation assumed here covers what is normally termed "Right Node Raising", most commonly found in coordinate structures (113a). Since we are dealing with noncoordinate structures, it is important to realize that the process also applies in noncoordinate structures, as in (113b) (from Williams 1990) and (113c):

(113)a. John talked to __) and really got to like my new friend from Kansas)
b. Anyone who talks to ___) really gets to like my new friend from Kansas)
c. ...go from the last town north __) to the first town south of that mosquito-infested river)

The conditions on BWD that are relevant below are: [24]

(114) a. the deletion site must be at right edge of its "domain"
b. the antecedent must be at right edge of its "domain"

There is no satisfactory account of what counts as a domain for BWD, that covers both coordinate and noncoordinate types. [25] Considering BWD in coordination, the domains mentioned in (114) must be conjuncts; moreover, these domains must be adjacent, with the conjunct containing the deletion site immediately preceding the conjunct containing its "antecedent" (hence the backward directionality). Assuming a view of phrase structure as proposed in Kayne (1994), including a Larsonian view of the internal structure of complex VPs, one can generalize from the coordinate to the noncoordinate cases by stating that the domain for the deletion site is a specifier, and the domain for its antecedent is the X'-sister of that specifier:

(115) ... [$_{XP}$ [**SPEC** ... ~~YP~~] [$_{X'}$ ... YP]] ...

Under the analysis of conjoined structures in which the conjunction is a head taking conjuncts as its complement and specifier, coordinate BWD configurations realize a specific variant of (115):

(116) [$_{\&P}$ [**SPEC=CP** ... ~~YP~~] [$_{\&'}$ and [$_{CP}$ ... YP]]

The non-coordinate cases also fit the scheme. (113b) involves the specifier of IP as the deletion domain, and I' as the antecedent domain (117a); (113c) might involve two PPs embedded in a Larsonian VP (117b):

(117)a. $[_{IP}\ [_{\textbf{SPEC=DP}}\ \ldots\ \sout{YP}\ ]\ [_{I'}\ \ldots\ \underline{YP}\ ]\ ]$

b. $\ldots\ [_{VP}\ [_{\textbf{SPEC=PP}}\ \ldots\ \sout{YP}\ ]\ [_{V'}\ \ldots\ [_{PP}\ \ldots\ \underline{YP}\ ]\ ]\ ]$

The cases we are interested in involve deletion in a relative clause modifiying a DP that is nonfinal in VP (118a). Under Larsonian assumptions, that relative clause forms the right edge of a complex DP in a specifier of an internal V'-constituent, as in (118b):

(118)a. John gave [whoever he could _ ~~two dollars~~] two dollars

$\ldots\ [_{VP}\ [_{\textbf{SPEC=DP}}\ \ldots\ [_{CP}\ \ldots\ \sout{YP}\ ]\ ]\ [_{V'}\ \ldots\ \underline{YP}\ ]\ ]\ \ldots$

In the following, whether (115) is correct or not, I assume that the right edge of the BWD domain must coincide with the right edge of the DP containing the relative clause. [26]

6.3 VPE / BWD interactions

In order for the account of "string contained" ACD's such as (96) as adjacent VPE and BWD deletions (98) to go through, it is necessary minimally to show independent word order rules interact with partial VPE and BWD to give the right results for good "containment" examples.

Beyond that, it is desirable to show that an independently motivated failure of partial VPE can be used to explain the ungrammaticality of a "string-contained" ACD; and likewise for BWD, to show that the ungrammaticality of a "string-contained" ACD can be attributed to solely to the violation of some constraint on BWD. The latter two points may of course be impossible to establish, insofar as the existence of a "bad" derivation may be obscured by a "good" derivation by another route, of the same string. Nonetheless, as will be shown below, all three goals can be secured: [27]

(I) For each "good" string-contained ACD, there is at least one derivation involving (partial) VPE, BWD, applying on independently well-formed surface word order
(II) If VPE fails, then the derivation fails
(III) If BWD fails, then the derivation fails

- *ad* (I):

All three examples (96a-c) have a derivation involving partial VPE and BWD and meeting requirements of both:

(119)a. John **gave** [whoever he could *give t* ~~two dollars~~] <u>two dollars</u>
b. John **finds** [everyone I do *find t* ~~easy to work with~~] <u>easy to work with</u>
c. John **believes** [everyone I do *believe t* ~~to be a genius~~] <u>to be a genius</u>

Partial VPE is independently established by the (relatively) acceptable examples (97) without BWD. BWD itself meets the requirements (114): the deletion target is right-peripheral in its domain (the DP containing the relative clause), and the antecedent to the deletion is right-peripheral in <u>its</u> domain - the matrix VP (strictly speaking, the V' sister of the DP containing the relative clause).

- *ad* (II):

For the original example (120) (=62), whose ungrammaticality was attributed to the PF condition (60) under the VPE-derivation (121a), it is now necessary also to exclude an alternative derivation, involving partial VPE interacting with BWD (121b):

(120) * John believes everyone I do, is a genius

(121)a. * John **believes** [everyone I do *believe t is a genius*] **is a genius**]
b. * John **believes** [everyone I do *believe* ~~is a genius~~] is a genius

In (121b), partial BWD targeting the string *is a genius* at the right edge of the relative clause, under identity with the same string at the right edge of the complement to the matrix *believe* should be possible.

In this derivation, however, partial VPE fails, for reasons discussed above: a verb whose complement clause contains a subject trace cannot undergo pseudogapping (109). Inserting a complementizer in the antecedent does nothing to save the example - the complementizer governed by *believe* in the relative clause cannot be included in a partial VPE, since VPE cannot delete "into" a finite clause: [28]

(122) * John **believes** [**that** [everyone I do *believe* [*that t* ~~is a genius~~]] is a genius]

Nor may the complementizer in the relative clause be left "undeleted" (to save pseudogapping) on pain of a *that*-trace violation:

(123)a. * John believes that everyone I do, that, is a genius
b. * John **believes that** [everyone I do *believe* that *t* ~~is a genius~~] is a genius

Hence, it can be concluded that an independently motivated constraint on partial VPE does real work in correctly excluding "string-contained" ACD's.

- *ad* (III):

To show that failure of BWD can lead to ungrammaticality in the realm of "string-contained" ACD's, it is sufficient to consider examples in which the part of the gap that is due to BWD (i.e. the part whose antecedent follows the gap) is itself not right-peripheral in the relative clause. Such examples are rather complex, and judgements subtle, but the contrasts are clear enough. Consider the paradigm (124):

(124)a. ? John believes [everyone I do _ to like spaghetti], to like beans
b. ? John believes [everyone I do_] to like beans
c. * John believes [everyone I do _ spaghetti], to like beans
d. ? John believes to like beans [everyone I do _ spaghetti]

(124a) involves partial VPE of *believe* in an ECM construction (cf. (97c) above). The same partial VPE is involved in (124b); the remainder of the gap (*to like beans*) is BWD-site, whose antecedent follows the gap, as in (124b'). The competing derivation (124b") which treats the whole gap as a VPE-site, violates string-containment (60):

(124)b'. John **believes** everyone I do *believe t* ~~to like beans~~, to like beans

b" * John **believes** everyone I do *believe t to like beans*, **to like beans** (**PF-containment*)

(124c) has no good derivation. Treating the deletion of *believes to like* as partial VPE as in (124c') - possible, given (107b) above - falls foul of the PF-condition (60). The alternative (124c") treats the gap as two deletions: pseudogapping of *believe*, as in (124a), and BWD of *to like*. The latter - crucially - violates the right-edge condition on BWD-sites (114a): the deletion is not right-peripheral in its domain (=the relative clause), since it is followed by non-deleted *spaghetti*:

(124)c'. * John **believes** everyone I do *believe to like* spaghetti, **to like** beans (**PF-containment*)
c". * John **believes** everyone I do *believe* ~~to like~~ spaghetti), <u>to like</u> beans) (**BWD*)

It is possible for the deletion site to be displaced to the right, by Heavy NP Shift of the accusative subject containing the relative clause, as in (124d). Crucially, now, the gap in the relative clause is no longer followed by part of its antecedent in surface order. Hence, there is a derivation involving partial VPE of the whole gap (possible, as in (107b)) which does not violate (60); and BWD is not necessary to derive any part of the gap:

(124) d'. John **believes to like** beans everyone I do *believe to like* spaghetti

Analogous paradigms are given in (125) and (126). The offending BWD-site in (125c) is *in the fridge*, the PP-complement of *put*; the element blocking right-peripherality of that BWD-site is the temproal adverb *at night*, distinct from the matrix *during the day* - hence necessarily part of the matrix. In (126c), BWD of *three dollars* is blocked by embedded *yesterday* - again, distinct from the matrix *today*, hence necessarily part of the relative clause. Reordering the gap to the right of the antecedent string in (125d) and (126d) again permits the gap to escape the conspiracy of the PF-constraint (60) and the BWD peripherality requirement (114a): [29]

(125)a. John puts [everything I do _ on the balcony at night] in the fridge during the day
b. John puts [everything I do _] in the fridge during the day
c. * John puts [everything I do _ at night] in the fridge during the day
d. John puts in the fridge during the day [everything I do _ at night]

(126)a. ? John gave [everyone that I did _ two dollars yesterday] three dollars today
b. ? John gave [everyone that I did _] three dollars today
c. * John gave [everyone that I did _ yesterday] three dollars today
d. John gave three dollars today [to everyone that I did _ yesterday]

From these paradigms, it can be concluded that constraints on BWD are independently needed to derive acceptability patterns among "string-contained" ACD's.

In sum, the contention that such examples do not constitute genuine ACD's finds subtle, hence strong, empirical support.

References

Baltin, M. 1987 Do antecedent-contained deletions exist? Linguistic Inquiry 18. 537-577.

Chomsky, N. 1993 A Minimalist Program for Linguistic Theory. In: K. Hale & S. Keyser (eds.) The View from Building 20. MIT Press, Cambridge MA. 1-53.

Chomsky, N. 1995 The Minimalist Program. MIT Press: Cambridge, Mass.

Chomsky, N. & H. Lasnik 1993 Principles and Parameters Theory. In: Jacobs, J., W. Sternefeld, A. von Stechow & T. Vennemann (eds) Syntax. An International Handbook., De Gruyter, Berlin. 506-569.

Dalrymple, M., S. Schieber & F. Pereira 1991 Ellipsis and Higher-Order Unification. Linguistics and Philosophy 14.
Diesing, M. 1992 Indefinites. MIT Press: Cambridge, Mass.
Fox, D. 1994 Quantifier Scope in VP Ellipsis. Ms. MIT
Guéron, J. & R. May 1984 Extraposition and Logical Form. Linguistic Inquiry 15, 1-31.
Haik, I. 1987 Bound VPs that Need to Be. Linguistics & Philosophy 10, 503-530.
Hornstein, N. 1994 An argument for Minimalism - the case of Antecedent-Contained Deletion. Linguistic Inquiry 25. 455-480.
Kayne, R. 1994 The Antisymmetry of Syntax. MIT Press: Cambridge, Mass.
Larson R. & R. May 1990 Antecedent containment or vacuous movement? Reply to Baltin. Linguistic Inquiry 21. 103-122.
Lasnik, H. 1995 Lecture Notes. Berlin.
May, R. 1985 Logical Form: Its Structure and Derivation. MIT Press: Cambridge, Mass.
Milsark, G. 1977 Towards an explanation of certain peculiarities of the existential construction in English. Linguistic Analysis 3. 1-29.
Sag, I. 1976 Deletion and Logical Form. PhD, MIT.
Stowell, T. 1981 The Origins of Phrase Structure. PhD, MIT.
Tancredi, C. 1992 Deletion, Deaccenting and Presupposition. PhD, MIT.
Vanden Wyngaerd, G. & J.-W. Zwart 1991 Reconstruction and Vehicle Change. In: A. van Kemenade (ed.) Linguistics in the Netherlands 1991.
Wilder, C. 1995a Some Properties of Ellipsis in Coordination. Geneva Generative Papers 2.
Wilder, C. 1995b Rightward Movement as Leftward Deletion. To appear in: Lütz, U. & J. Pafel (eds.) Extraposition and Extraction in German. John Benjamins: Amsterdam.
Williams, E. 1977 Discourse and Logical Form. Linguistic Inquiry 8, 101-139.
Williams, E. 1990 The ATB Theory of Parasitic Gaps. The Linguistic Review 6, 265-279.

[1] This is a preliminary written version of material that was presented in colloquia in Berlin (July and November 1995) and Potsdam (November 1995). Thanks to those audiences, and also to the following individuals, for helpful discussions and comment: Manfred Bierwisch, Daniel Büring, Marcel Den Dikken, Gisbert Fanselow, Hans-Martin Gärtner, Gerhard Jäger, Peter Ludlow, Anoop Mahajan, Robert May, Tom Roeper, Michal Starke, Markus Steinbach, Anatoli Strigin, Ralf Vogel, Sten Vikner, Jan-Wouter Zwart. The paper was indirectly inspired by Howard Lasnik's lectures on Ellipsis and QR in Berlin in June 1995.

[2] Even the simplest examples such as (2) demonstrate that phonological identity is not required - *meet* is not identical with *met*.

[3] I leave a third proposal - that VPE is some kind of pronominal element with no internal structure at any syntactic level (cf. Dalrymple et. al. 1991) - out of consideration here.

[4] Although *near* can be stranded under wh-movement (i), the combination with pseudogapping in a relative clause (ii) is not good, for unclear reasons. Pied-piping of the preposition (iii) is much better:
(i) ? John stood near Sue, who Mary stood near, too
(ii) * John stood near Sue, who Mary did near too
(iii) ? John stood near Sue, near whom Mary did too

[5] Baltin does not comment on whether in addition to (27), there is a filter banning antecedent containment at LF. Nor does he make explicit whether he endorses QR and LF-copying as the mechanism behind VPE-interpretation in these examples; but it is difficult to see what else could have been intended.

[6] An analysis of (31a) in which both complements have shifted rightwards, vacating VP, as in (i), is precluded by independent considerations:
(i) John [$_{VP}$ gave t_j t_k] [whoever he could _]$_j$ [two dollars]$_k$
While RCE may apply to a goal NP, splitting off the relative clause from the NP-head as in (ii), free relatives are never split by RCE (iii). In argument positions, free relatives behave like NPs; in particular, in the "goal" position in the double object construction, neither undergoes rightward displacement of the Heavy NP Shift type (iv,v).
(ii) John gave [everyone _] two dollars who he met at the party
(iii) * John gave [whoever_]two dollars he met at the party
(iv) * John gave _ two dollars whoever he met at the party
(v) * John gave _ two dollars the boy next door
Hence, under standard assumptions, the VPE in (31) is contained within its antecedent VP at S-structure.

[7] Given the VP-internal subject hypothesis, the VP that is copied in (34c) contains an A-trace left by the main clause subject *John*, which must function in the relative clause as the A-trace of the subject *Mary*. It must therefore be assumed that NP-traces count as identical for purposes of VPE. This assumption is necessary in any case for the simplest cases: cf. (i), where the trace left by movement of John in the first conjunct must be assumed to be identical with the NP-trace forming a chain with Mary in the second conjunct:
(i) John was [kissed t] and Mary was _ too

[8] One issue concerns the interaction of V-movement with VPE in the Minimalist framework, where it is assumed that finite verb raise to the highest Infl-head (AGs) in LF. If it is assumed (i) that the chain of the verb must be included in the constituent that matches (is copied into) the VPE, and (ii) copying / matching follows all LF-movement, then the account must give up the assumption that finite V raises to AGs. Otherwise, the smallest constituent in the antecedent that matches (can be copied into) the VPE would be AGs', which by hypothesis contains AGoP, hence also the VPE itself, even after object raising. Hornstein suggests that English V does not raise higher than AGRo at LF, and that the target of "VP-Ellipsis" is in fact AGo', including the raised verb but excluding the raised object.

[9] Chomsky (1995:Ch4) has since developed an account of covert A-movement (object raising) that does not involve pied-piping of the whole object phrase.

[10] Example (31a), with the VPE contained within the goal NP object of give in a double object construction, can plausibly be handled via A-movement.

[11] (39b) is slightly degraded in Lasnik's judgement. I find the example perfect.

[12] There is no narrow scope reading for the VPE in these examples: *be + XP* (*is a genius / were stupid*) cannot reconstruct under *do / did.*

[13] It must be assumed in addition that the extraposed relative reconstructs into the subject prior to QR. See section 3.2.

[14] It has been proposed that QR may extend the domain of extraposition, such that the extraposed clause may cross a sentence boundary rightwards, in violation of the Right Roof Constraint, just in case the head of the extraposed clause itself raises the same distance leftwards - cf. Guéron & May (1984). But nowhere has it been proposed that extraposition may violate the right-roof constraint in order to extend the domain of QR.

[15] This discussion presupposes that (60) belongs to UG. One may speculate that the facts (60) is intended to capture may in fact not be linguistic at all, but rather reflect an extralinguistic condition, perhaps relating to the parsing mechanism.

[16] Thanks to Anoop Mahajan and H-M. Gärtner for directing me to this reference. It is important to realize that the non-presuppositional reading is not to be equated with the narrow scope reading in sentences with an additional quantifier. Rather, only in its presuppositional reading does a weak quantifier undergo QR, in which case it may take narrow or wide scope with respect to the second quantifier.

[17] I ignore here the possiblity that the raised DP adjoins to the matrix V'.

[18] Some speakers find (91b) degraded, but the contrast with (91a) seems to be robust.

[19] The precise categorization of *than / als* (e.g. clause-governing preposition or clausal head C) is not crucial to the argument.

[20] This account of the contrast in (88) (English), and between (94) and (95) (German) depends crucially on the assumption that the complements of *than* and *als* in the comparative is <u>always</u> clausal, with the bare DP in these cases being left behind by ellipsis. If there is the possibility for *than / als* to govern a bare DP rather than a clause in the syntax, then there is the possibility for an abstract representation of (88) or (95) without ellipsis, and the constrast goes unexplained. This assumption has further consequences regarding the notion of identity required for ellipsis. In (i), the DP *Peter*, in the complement of *than*, is interpreted as the subject of a VP - more correctly, an I'-constituent - whose antecedent is nonfinite. Given that the DP *Peter* is overt, it must be in a Case position (in terms of Chomsky&Lasnik (1993), not a position in which "null Case" is licensed). Hence this DP cannot be the subject of a non-finite IP (ii). It must be concluded that a finite I' can be deleted under identity with a finite I' (iii). The same goes for German (iv-vi):

(i) to eat more apples than Peter is not advisable
(ii) * [PRO **to eat more apples** than Peter *to eat x-many apples*] , is not advisable
(iii) [PRO **to eat more apples** than Peter *eats x-many apples*] , is not advisable
(iv) mehr Äpfel zu essen als Peter, ist nicht ratsam
more apples to eat than Peter is not advisable
(v) * [PRO **mehr Äpfel zu essen** als Peter *x-viele Äpfel zu essen*] , ist nicht ratsam
(vi) [PRO **mehr Äpfel zu essen** als Peter *x-viele Äpfel ißt*] , ist nicht ratsam

[21] VP-ellipsis sites, including partial VPE, are given in italics; their antecedents in bold. BWD-dependencies are indicated by crossing-through of items deleted under identity with underlined items.

[22] Notice that two independent (partial) VPE operations can coexist in a complex structure, even a verb-complement construction such as (iii):
(i) John will leave, if you come, and Mary will, if I do
(ii) John hates anyone who works for Labour and Mary does, anyone who does, for the Tories.
(iii) John believes that Mary likes beans, and Bill does, that Sue does, spaghetti
The last example does not counterexemplify the generalization in the text. The deletion of the lower finite verb "begins" a new (partial) VPE domain.

[23] This fact is probably related to Stowell's (1981) observation that a deleted complementizer has to be adjacent to its governing verb: cf. *Mary believes sincerely *(that) Peter is a fool.*

[24] See Wilder (1995a, 1995b) for fuller discussion of BWD.

[25] Williams (1990) suggests that noncoordinate RNR contexts are really conjoined structures - subject conjoined with predicate, etc. I reject this view.

[26] In (118a), the antecedent string exhausts the overt terminals of the antecedent domain (=V'). The converse situation - where the deletion site exhausts the deletion domain (=Spec), must probably be excluded. Kayne (1994), Wilder (1995b) assume that a deletion site may not c-command its "antecedent".

[27] If (I)-(III) are not met, we are faced with a problem of undergeneration, avoidance of which is a *sine qua non.* Independently of this, questions of overgeneration also arise. For example, Robert May (p.c.) asks why BWD cannot apply independently of partial VPE in examples like (i), pointing to the (near)-unacceptable status of (ii):
(i) John gave whoever he could, two dollars.
(ii) *? John gave whoever he could give, two dollars.
(iii) John gave [whoever he could give *t* ~~two dollars~~] <u>two dollars</u>
(iv) * John gave [whoever he could give *t* two dollars]
The derivation in (iv), with overt *two dollars* inside the relative clause, has the status of a projection principle violation; there is no deletion rule that can yield deletion of the second object of the matrix verb *gave*, hence it cannot have been projected into the syntax. The BWD derivation (iii), with overt *two dollars* in the matrix VP, requires a prosodic boundary between *give* and *two dollars*, indicated by the comma in (i). One factor contributing to the deviant status of (ii) may be the identity of the verbs, which makes partial VPE in (i) possible in the first place. There is some evidence from coordination that there is an economy-like principle which prefers derivations with deletion over derivations without ("if a deletion can apply, it must"); consider the paradigm (v-vii):
(v) John switched the light off and Mary the TV (*switched+off*)
(vi) * John switched the light off and Mary the TV off (*switched*)
(vii) ? John switched the light off and Mary the TV on (*switched*)
If non-identical verbs are chosen in the construction (ii), the example seems to improve somewhat:
(viii)?? John sent whoever he couldn't give, two dollars.
Independently of these considerations, the problem of demarcating the applicability of (non-coordinate) RNR is in any case unresolved, cf. Williams (1990), Wilder (1995a).

[28] In fact, in the version of this example that avoids the PF constraint (60) by extraposition, it makes no difference whether or not the antecedent contains overt *that*, although the appearance of overt that in the relative clause would cause a *that*-trace violation:
(i) John believes that everyone is intelligent that I do _
(ii) * John believes that everyone is intelligent that I believe that is intelligent
From this example alone it is impossible to decide whether (i) the 'overt' *that* is reconstructed into the VPE site, with the *that*-trace filter only applicable to overt *that*; or (ii) the 'deleted' complementizer is treated as nondistinct from the overt complementizer for purposes of VPE-identity. Other examples indicate the former. Overt verbs of "manner-of-speaking" such as *mutter*, for example, do not license the 'deleted' complementizer:
(iii) * John muttered everyone is intelligent

However, wide scope VPE, involving reconstruction of *mutter* together with its complement CP containing a subject trace, is possible:

(iv) John muttered that everyone was intelligent that I did.

(v) * John muttered that everyone was intelligent that I did mutter that was intelligent

(vi) John [**muttered** [**that** [everyone _] **was intelligent** [that I did *mutter that t was intelligent*]]]

(vi) shows that the that-trace effect can be voided in VPE contexts for the verbs in question. The *that*-trace effect is a phenomenon sensitive solely to the PF-status (overtly realized or not) of the complementizer.

29 The bare DP goal object of *give* * does not undergo Heavy NP Shift; hence the effect of reordering the impossible deletion site of (126c) is shown using the variant with *to* in (126d).