Pseudoclefts and Ellipsis

Marcel den Dikken, André Meinunger & Chris Wilder
CLS/Tilburg University & ZAS/Berlin

1 Introduction

1.1 Connectivity effects in specificational pseudoclefts

A hotly debated issue in generative linguistic theory through the years has been the behaviour of specificational pseudocleft constructions (SPCs, for short) in the domain of 'connectivity' or 'connectedness' effects. Four such effects have figured prominently in the literature:

(1) *binding*
  a. what *he* is is [angry with {*himself*/**him*/**John*}][1]
  b. *he* is is [angry with {*himself*/**him*/**John*}]

(2) *opacity (de dicto/de re)*
  a. what John seeks is [a unicorn] [*de re* or *de dicto*]
  b. John seeks [a unicorn] [*de re* or *de dicto*]

(3) *bound variable anaphora*
  a. what *nobody* bought was [a picture of *his* house]
  b. *nobody* bought [a picture of *his* house]

(4) *negative polarity item (NPI) licensing*
  a. what *nobody* bought was [*any* wine]
  b. *nobody* bought [*any* wine]

In all of these cases it seems like the SPCs behave like their simple-clause counterparts in the b-examples, where the binder c-commands the bindee in (1), (3) and (4), and the intensional verb *seek* takes *a unicorn* as its complement. But in surface syntax the SPC certainly does not look like a simple clause, and no relationship seems establishable between the constituents that entertain a dependency relationship of some sort in the examples above. The question that SPCs pose, then, is how the apparent lack of a structural relationship between constituents of the *wh*-clause and the 'counter-weight' (as Heycock 1994 labels it) can be 'set straight' at some point in the derivation of specificational pseudoclefts.

The literature on SPCs has brought forth a variety of approaches to the problem, which can basically be grouped into three sets:

(i) *the semantic approach* (cf. Sharvit 1997 for a recent representative)
- seeks to derive the facts without syntactic c-command/constituency
- binding etc. is viewed as a side effect of semantic composition
- the semantic properties of *what* and *be* play a key role

(ii) *the syntactic reconstruction approach* (cf. Heycock & Kroch 1996, Bošković 1997)
- the 'counterweight' is 'moved into' the *wh*-clause at LF
- thus c-command/constituency is established at LF

(iii) *the ellipsis approach* (cf. Ross 1997)
- assumes that the 'counterweight' is a full IP before and after Spell-Out (cf. (5) below)
- thus has c-command/constituency at all levels of syntactic representation
- PF ellipsis reduces the 'counterweight' in cases in which it is smaller than IP

[1] Coreference is marked with italicisation here; *him* and *John* are fine in (1) if counterindexed with *he*.

This paper will not concern itself in much detail with (i); instead, it argues against approaches along the lines of (ii), siding with the ellipsis approach in (iii) for a *subset* of SPCs. Specifically, this paper will argue for:

(iv) *the 'two types of specificational pseudocleft' approach* (this paper)
- assumes that the 'counterweight' is a full IP before and after Spell-Out **only for a well-delineated subset of SPCs**
- establishes a link between this subset of SPCs and **Question-Answer pairs** (QAPs)
- reduces connectivity effects concerning **NPI** licensing to syntactic c-command
- but leaves the other connectivity effects to be treated in a different way

1.2 Two types of specificational pseudoclefts

The present paper thus aims to contribute to the discussion about pseudocleft constructions by presenting an extended argument for a distinction within the set of SPCs between two significantly different subtypes, Type A and Type B. Quintessential examples of Type A SPCs involve full-IP 'counterweights', as illustrated in (5):[2]

(5) a. what John bought was [he bought some wine]
b. what John didn't buy was [he didn't buy any wine]

We will argue that a subset of SPCs with apparently 'smaller-than-IP' counterweights should be analysed as **elliptical** counterparts of full-IP SPCs of the type in (5) — in particular, those featuring connectivity effects with Negative Polarity Items (NPIs) in SPCs, which we argue should be kept distinct from the other connectivity exemplified in (1)-(4), above.

(6) a. what John bought was [~~he bought~~ some wine]
(i.e. *what John bought was some wine*)
b. what John didn't buy was [~~he didn't buy~~ any wine]
(i.e. *what John didn't buy was any wine*)

Our prime cue to the distinction between NPI connectivity and other such effects comes from the fact that, while the connectedness effects shown in (1)-(3) are preserved when the word order of the pseudoclefts is turned around, as in (7a-c), the NPI connectivity case in (4) breaks down under 'inversion', as (7d) shows. It is this ban on 'inversion' which SPCs of the type in (4) share with the full-IP SPCs in (5), as seen in (8).[3]

(7) a. [angry with {*himself*/**him*/**John*}] is what *he* is
b. [a unicorn] is what John seeks [*de dicto* or *de re*]
c. [a picture of *his* house] is what *nobody* bought
d. *[*any* wine] was what *nobody* bought

(8) a. *[he bought some wine] was what John bought
b. *[he didn't buy any wine] is what John didn't buy

---

[2] cf. Clifton (1969:38), Ross (1972:89), Higgins (1979:47), Kayne (1998:26). Higgins (1979:86) points out that for him sentences of this type are ungrammatical, sounding 'irremediably anacoluthic'; he suggests that 'these sentences have arisen, historically, by analogy to question-answer pairs'. In his fn. 11 he compiles a wealth of interesting empirical evidence to underpin the links between SPCs of the type in (5) and Question-Answer pairs. We will capitalise on this connection in this paper.

[3] We are using the term 'inversion' in a pre-theoretical sense here, not making any claim with regard to the way in which it comes about (i.e., via Predicate Inversion in the technical sense of the term, or some other means). We will use 'invert' and 'reverse' interchangeably.

Irreversibility thus diagnoses Type A SPCs: on the basis of the fact that full-IP SPCs like (5) do not reverse, we will classify all SPCs with 'smaller-than-IP' counterweights that fail to reverse as Type A SPCs. All those SPCs which do exhibit reversibility, on the other hand, are Type B SPCs. Along with their irreversibility, Type B pseudoclefts never exhibit NPI connectivity.

There is a third component to the clustering of properties which tease apart Type A and Type B SPCs, which we will also use as a diagnostic tool in our exposition — the distribution of SPCs in the small-clausal complement of verbs like *call* and *describe as*[4]. While these small clause complements allow for SPCs with a *wh*<XP word order in principle, as shown in (9), they resist SPCs which are unequivocally of Type A, whether elliptical or 'undeleted', as shown in (10):

(9) a. I'd call [what John is {frustrated/a fool}]
b. I'd describe [what John is as {frustrated/a fool}]
(10) a. *I'd call [what John didn't buy [~~John didn't buy~~ any wine]]
b. *I'd describe [what John didn't buy as [~~John didn't buy~~ any wine]]

Whenever a pseudocleft with a nominal or adjectival counterweight fails to be embedded under verbs like *call* and *describe as*, we can be sure that we are dealing with a Type A SPC[5]. It is in this way that we will call upon the distribution of SPCs under *call/describe as* in the bulk of the paper: as a diagnostic for Type A status. In section 6 we will come back to the question of what the proper treatment of the examples in (9) should look like — with hindsight, after we have discussed the restrictions on the distribution of Type A SPCs.[6]

Throughout, what we see is that the following descriptive generalisation holds:

(11) NPIs are found in the counterweight of SPCs only where full-IP SPCs are licensed

Viewed this way, then, the NPI licensing effects seen in SPCs like the ones in (4) do not involve 'connectivity' in the strict sense of the term at all. Rather, the NPI is directly c-commanded by its licensor throughout the syntactic derivation, from Merge via Spell-Out through to LF. In section 2 we

---

[4] The small clauses in (9) count as SPCs by the selection test (see Higgins 1979:). It is not John's profession or role that is described as frustated, but John himself; likewise, John's profession is not described as being a fool (in the 'court jester' sense) in the most natural reading of (9).

[5] The verbs *call* and *describe as* impose categorial restrictions on the small clause they govern: in the frames *call XP YP/describe XP as YP*, YP may be adjectival or nominal (including gerundives, free relative, proper names), but not prepositional, verbal (participial) or clausal (finite or non-finite IP/CP):

(i) a. John called me frustrated/boring/clever
b. John called this his favorite/being silly/what he always wanted/Geronimo
c. * John called XP in the bath/for my birthday/with no sense
d. * John called XP (that) he had left/(for him) to have left
e. * John called XP bought some wine/run a mile

(ii) a. John described me as frustrated/boring/clever
b. John described this as his favorite/being silly/what he always wanted/Geronimo
c. * John described XP as in the bath/for my birthday/with no sense
d. * John described XP as (that) he had left/ (for him) to have left
e. * John described XP as bought some wine/run a mile

Thus, there are probably independent reasons why SPCs with PP, VP, IP or CP counterweights fail under these verbs. But SPCs with AP and DP counterweights do occur in this context, cf. (9). The striking failure of (10) thus supports our claim that counterweights containing NPIs are actually IPs, not DPs.

[6] The reason for this cautionary note with regard to the examples in (9) lies in the fact that embedding *wh*<XP SPCs under ECM (and raising) verbs is often claimed to yield ungrammatical results in all cases (cf. Higgins 1979). We will have more to say about this issue in section 6; electing not to burden the text discussion with questions concerning the analytical status of (9) at this point, we will put the question on reserve until the distributional restrictions on SPCs have been properly catalogued and accounted for.

will show that this is a decisive advantage of the ellipsis approach to NPI 'connectivity' in SPCs when compared to syntactic reconstruction accounts à la Heycock & Kroch (1996) and Bošković (1997).

There is evidence of various kinds (involving reversibility, NPI connectivity and distribution in small clause complements to verbs like *call*), then, to support a two-way split in the set of SPCs:

(A) a. Type A SPCs systematically involve full-IP counterweights
b. the fact that their counterweights may look smaller than IP is the result of **ellipsis** (specifically, Forward Deletion; cf. Wilder 1997)
c. NPI 'connectivity' involves regular c-command in the full-IP syntactic representation of the counterweight (cf. (2b) with (1b))

(B) a. Type B SPCs systematically feature 'smaller-than-IP' counterweights
b. there is no ellipsis in Type B SPCs
c. hence Type B SPCs do not feature NPI connectivity

The representation of Type A SPCs is our primary focus in this paper. We will liken their analysis to that of Question-Answer pairs (QAPs), which likewise show (optional) ellipsis and NPI connectivity, as seen in the parallel between (6) and (12):

(12) a. what did John buy? — [~~he bought~~ some wine]
b. what didn't John buy? — [~~he didn't buy~~ any wine]

The link with QAPs leads us to argue that the *wh*-clause in Type A SPCs should be analysed as an interrogative CP, *not* as a free relative.

The paper is organised as follows. In section 2 we will consider in detail the NPI connectivity effects exhibited by SPCs and QAPs, analysing them in terms of straightforward S-structure c-command obtaining in a counterweight/answer IP which is subject to optional ellipsis. Section 3 then addresses the restrictions on the ellipsis process operative in Type A SPCs and QAPs, against the background of Wilder's (1997) study of ellipsis in coordinate structures. A compendium of further evidence for the connection between QAPs and Type A SPCs is presented in section 4. Section 5 outlines the structural representation of Type A SPCs that we would like to propose. And finally, section 6 considers Type B SPCs and the connectedness effects that they exhibit, suggesting possible ways of coming to terms with these.

2 NPI connectivity in specificational pseudoclefts and Question-Answer pairs

2.1 NPIs and S-structure licensing

Negative Polarity Items are elements which must be licensed by a c-commanding negation (or affective operator; we will focus on the Neg-licensed cases); specifically, an inspection of the empirical facts shows that the licensing negation must c-command the NPI *at S-structure* in most cases. This is brought out by a comparison of the examples in (13):

(13) a. everybody didn't come [*every<not / not<every*]
b. many students didn't come [*many<not / not<many*]
c. *any student didn't come

While, as (13a,b) show, it is possible for the sentential negation to scope over a quantified subject at LF, (the *not<every/many* readings being grammatical), the lack of an S-structure c-command relationship must ostensibly be held responsible for the ungrammaticality of (13c), featuring an NPI in a subject position potentially c-commanded by the negation at LF. That the ungrammaticality of (13c) is not due to some ban on NPIs in subject positions is shown by the contrast between (13c) and the examples in (14), which all are like (13c) in featuring the NPI in subject position but differ from it in that the negation c-commands the NPI at S-structure.

(14) a. never did any students come
b. didn't any students come?
c. he doesn't think that any students came
d. there wasn't any students in the lecture hall
e. ?there didn't happen an accident of any kind

The fact that NPIs must appear in the scope of a negation at S-structure is further underscored by the ungrammaticality of the following sentences, which shows that A´-movement of an NPI or a constituent containing an NPI away from the licensing negation, into a position not c-commanded by the negation at S-structure, is not allowed, even though LF-reconstruction should be able to restore a c-command relationship between the NPI and its prospective licensor.[7]

(15) a. *any students, I didn't see [contrast *I didn't see any students*]
b. *see any students, I didn't

The sensitivity of NPI licensing to *S-structure* c-command in cases like the above allows us to distinguish between the various approaches to SPCs presented in the literature. In particular, syntactic reconstruction approaches to SPCs à la Heycock & Kroch (1996) and Bošković (1997) seem to establish the requisite c-command relationship between the NPI and its licensor at too late a point in the derivation: only at LF does the negation c-command the *any*-phrase. The ellipsis approach, by contrast, has the negation c-commanding the NPI throughout the derivation, in the IP counterweight:

(6b) what John didn't buy was [~~he didn't buy~~ any wine]

The same problem is posed by QAPs, which likewise allow for an NPI in the answer to apparently be licensed by a negation in the question. The ellipsis approach once again reduces the licensing problem to a simple case of S-structure c-command in an elliptical IP:

(12b) what didn't John buy? — [~~he didn't buy~~ any wine]

This is a decisive advantage of the ellipsis approach to (Type A) SPCs and QAPs with respect to (English) NPIs, which, as (15) showed, resist standard 'reconstruction'/LF licensing, requiring to be c-commanded by their licensor at S-structure in the bulk of cases.

2.2 *NPIs in preverbal subjects, and post-S-structure licensing*

As Uribe-Echevarria (1994) shows, however, there are certain conditions under which an NPI may occur within a preverbal subject, not c-commanded by a licensing negation at S-structure. Relevant examples are given in (16) and (17):

(16) a. [a doctor who knew anything about acupuncture] wasn't available
b. *[a doctor who knew anything about acupuncture] wasn't sitting on the floor
(17) a. [that anyone would leave the company] wasn't mentioned in the meeting
b. *[that anyone will leave the company] wasn't mentioned in the meeting

The conditions under which NPIs in preverbal subjects can be licensed are complex (see Uribe-Echevarria's work for detailed discussion). We need not be concerned with these here. What matters for our purposes in this paper is that precisely under the circumstances in which NPIs in preverbal subjects are licensed in (16) and (17), we also find them in both SPCs and QAPs:

---

[7] As a matter of fact, (15b) continues to be ungrammatical when turned into a biclausal construction with a negation in the upstairs clause and topicalisation downstairs, as seen in (i):
(i) *he doesn't think that [see any students], he will
This suggests that a 'topic island' effect constrains NPI licensing as well — something which follows (from subjacency) on an LF-movement approach to NPI licensing; see Moritz & Valois (1992) for relevant discussion.

(18) a. what wasn't available was [a doctor who knew anything about acupuncture]
b. *what wasn't sitting on the floor was [a doctor who knew anything 'bout acupuncture]
(19) a. what wasn't mentioned in the meeting was [that anyone would leave the company]
b. *what wasn't mentioned in the meeting was [that anyone will leave the company]
(20) a. ?what wasn't mentioned in the meeting? — [that anyone would leave the company]
b. *what wasn't mentioned in the meeting? — [that anyone will leave the company]

That we are dealing with Type A SPCs in (18) and (19) is evident from the fact that the grammatical examples resist inversion as well as embedding under verbs like *call*:

(21) a. *[a doctor who knew anything about acupuncture] was what wasn't available
b. *[that anyone would leave the company] was what wasn't mentioned in the meeting
(22) a. *I'd call what wasn't available [a doctor who knew anything about acupuncture]
b. *I'd call what wasn't mentioned in the meeting [that anyone would leave]

The contrast between the grammatical examples in (18a) and (19a) and their ungrammatical counterparts in (21)-(22) is sharp. Of the starred examples, the cases in (21) are perhaps the most striking ones, in the light of the grammaticality of the examples in (16a) and (17a), which have the same linear order as the ill-formed pseudoclefts in (21). The facts in (16)-(22) thus strongly confirm our claim that NPI connectivity effects are a property of Type A SPCs only.

2.3 NPIs in postverbal subjects of expletive constructions

2.3.1 there *expletive constructions*

Other NPI related facts do, too. Consider, first of all, the SPC in (23a) and the QAP in (23b):

(23) a. ?what didn't happen was [an accident of any kind]
b. ?what didn't happen? — [an accident of any kind]

These examples are acceptable (though perhaps slightly marginal) — surprisingly, it seems, given the ungrammaticality of (24). But in English indefinite subjects of (certain) intransitive sentences can occur postverbally, in *there* constructions. And as already shown in (10e), repeated below, NPI licensing in the *there* counterpart of (24) succeeds.

(24) *[an accident of any kind] didn't happen
(10e) ?there didn't happen an accident of any kind

What this means for (23) is that they can be derived as elliptical variants of the full-IP examples in (25), which have the same level of acceptability as (23), as expected.

(25) a. ?what didn't happen was [there didn't happen an accident of any kind]
b. ?what didn't happen? — [there didn't happen an accident of any kind]

In support of the claim that the SPC in (23a) is derived from a full-IP Type A SPC à la (25a), featuring expletive *there*, we point out NPI connectivity effects of the type in (23a) are found only in contexts which accept *there* expletives. To see this, contrast the example in (23a) with cases like (26), which are crashingly bad, on a par with the ungrammaticality of the *there* sentences in (27).

(26) a. *what didn't annoy Mary was an accident of any kind
b. *who didn't laugh was any of Bill's students
(27) a. *there didn't annoy Mary an accident of any kind
b. *there didn't laugh any of Bill's students

It is precisely in those contexts in which English allows postverbal subjects containing an NPI in *there* sentences that we find grammatical SPCs with NPI connectivity of the type in (23a).

Note furthermore that, just like (25a) is irreversible and unembeddable under verbs like *call* (as shown in (29)), so is (23a), as (28) illustrates:

(28) a. *[an accident of any kind] didn't happen
b. *I'd call what didn't happen [an accident of any kind]
(29) a. *[there didn't happen an accident of any kind] was what didn't happen
b. *I'd call what didn't happen [there didn't happen an accident of any kind]

So what we find, once again, is full parallelism (i) between SPCs and Question-Answer pairs, and (ii) between Type A SPCs with full-IP counterweights and SPCs featuring NPI connectivity.

2.3.2 it *expletive constructions*

A final case showing that SPCs with NPI connectivity parallel Type A SPCs with full-IP counterweights is introduced by the examples in (30):

(30) a. what happened next was [that he fell]
b. what happened next was [he fell]

These two examples seem superficially very similar. Yet there is quite a bit of evidence that they are in fact entirely different in structure and derivation. We will approach the matter from the perspective of the properties of the SPC in (30a), which — of the two cases in (30) — is the one which furnishes the argument for a link between NPI connectivity and Type A structures.

Consider the following facts about SPCs of the type in (30a):

(31) a. what happened next was [that he fell] (= (30a))
b. [that he fell] was what happened next
(32) a. ?what happened next was [it happened next [that he fell]]
b. *[it happened next [that he fell]] was what happened next
(33) a. what didn't happen next was [that anybody fell]
b. *[that anybody fell] was what didn't happen next
(34) a. *I'd call what didn't happen next [that anybody fell]
b. *I'd describe what didn't happen next as [that anybody fell]

While (31a) is perfectly reversible, (32a), the full Type A version of (31), is not; and on a par with (32a) but in contradistinction to (31a), the example in (33a), which features an NPI in the counterweight, to be licensed by the negation in the *wh*-clause, is irreversible as well. Moreover, the NPI case in (33a) also resists embedding under the verbs *call* and *describe as*, as seen in (34). This once again vindicates the link we have drawn between NPI licensing in the 'counterweight' of SPCs and a Type A structural analysis. And as before, these Type A SPCs behave like QAPs in all relevant respects:

(35) a. what happened next? — ?[that he fell] (cf. (30a))
b. what happened next? — ?[it happened next [that he fell]] (cf. (32a))
c. what didn't happen next? — ?[that anybody fell] (cf. (33a))

The facts in (31)-(35) can hence be analysed straightforwardly by saying that (30a) allows for two potential derivations, one built on the Type A structure in (36a), which is subject to optional ellipsis and which procures the NPI 'connectivity' effects, and the other based on the Type B representation in (36b), featuring a CP counterweight and no ellipsis (on Type B, see section 6 for more discussion):

(36) a. what happened next was [~~it happened next~~ [that he fell]] (Type A)
b. what happened next was [$_{CP}$ that he fell] (Type B)

As before, the Type B case allows for inversion while the Type A case does not. The example in (31b) hence unequivocally derives from (36b). The examples in (32)-(34), by contrast, can only be built on (36a), the ill-formedness of the unacceptable cases following from the general constraints on the distribution of Type A SPCs.

While this takes care of (30a), things remain to be said about the example in (30b), though. For notice that, in contrast to (30a), the example in (30b) cannot be reversed (cf. (37b)); in concert with this, it does not allow an NPI in the counterweight at all (cf. (38b)),[8] nor can it be embedded under the verbs *call* and *describe as* (cf. (39)):

(37) a. what happened next was [he fell] (= (30b))
b. *[he fell] was what happened next
(38) a. *what didn't happen next was [anybody fell]
b. *[anybody fell] was what didn't happen next
(39) a. *I'd call what happened next [he fell]
b. *I'd describe what happened next as [he fell]

The ungrammaticality of (37b) and (39) will be straightforward if we can ensure that (30b)/(37a) can only be a Type A SPC; put differently, the fact that (37b) and (39) are ill-formed will follow if an analysis of (30b) as in (40), the counterpart of (36b) with a bare IP counterweight, can be excluded on principled grounds:

(40) what happened next was [$_{IP}$ he fell] (* *qua* Type B SPC)

The question that (30b) hence poses is why a Type B scenario for this SPC, à la (40), is apparently unavailable. For (37b) an answer to this question in fact seems relatively easy to give: in general, IPs (such as *he fell* in (37)) cannot be root subjects, as shown by the ungrammaticality of the examples in (41):

(41) a. [*(that) he fell] was the next thing that happened
b. [*(that) he fell] was quite a happening

And though the ill-formedness of (39) on a Type B derivation seems less straightforward at first blush, we can make sense of it along essentially the same lines as (37b) on an analysis of Type B SPCs à la Heggie (1988), Heycock (1991), according to which these involve a small clause predication structure, the *wh*-clause functioning as the predicate and the IP as its subject:

(42) [$_{SC}$ [$_{Subject}$ counterweight] [$_{Predicate}$ *wh*-clause]] [Type B SPCs]

Since a Type B approach to (30b) would thus involve postulating a finite IP subject to a small clause, the general resistance of (finite) IPs to being placed in subject positions will rule out a Type B base for SPCs of the sort found in (30b).

While this plausibly answers the question of why the examples in (37b) and (39) are ill-formed (their ungrammaticality matching that of all Type A SPCs in these contexts), it still leaves (38a) unsolved. We will come back to this construction in sections 3.4 and 3.5, below (see also fn. 6, above, for some pertinent observations). But there is one particular aspect of the ill-formedness of (38a) that invites discussion right here: the question of why it apparently cannot be derived as in

---

[8] In point of fact, the problem with (38a) seems to be independent of NPI-licensing *per se*, for (ia) crashes as well, on a par with the QAP in (ib), while (iia,b) are (marginally) acceptable:

(i) a. *?what didn't happen next was [he fell]
b. *?what didn't happen next? — [he fell]
(ii) a. ?what didn't happen (at least) was {[it didn't rain]/[John didn't leave]}
b. ?what didn't happen (at least)? — {[it didn't rain]/[John didn't leave]}

From the perspective of the parallel account of Type A SPCs and QAPs, what the facts suggest is that a negated *wh*-interrogative is unable to take an indirect IP-answer that is not itself negated. Whether the deviance of (i) is a matter of grammatical ill-formedness or infelicitousness of some sort is a matter which we will leave open.

(43b), which parallels (43a), the structure that underlies the grammatical NPI case in (33a) (*what didn't happen next was that anybody fell*). This question prompts a discussion of the restrictions on ellipsis, which will furnish an answer to the question of why the ellipsis structure in (43b) is ill-formed. This will be the topic of the next section.

(43) a. what didn't happen next was [~~it didn't happen next~~ [that anybody fell]] (cf. (36a))
b. *what didn't happen next was [~~it didn't happen next [that~~ anybody fell]]

3 Ellipsis

This paper analyses a subset of SPCs — the ones we have labelled Type A SPCs — in terms of full-IP counterweights which optionally undergo ellipsis. At the end of the previous section we have come across a case where ellipsis in the full-IP counterweight of a Type A SPC fails (cf. (43b)). One of the things we will do in this section is to clear up the cause of the ill-formedness of (43b). We will start out by laying out the general constraints on ellipsis identified in Wilder (1997).

3.1 On the nature of ellipsis

In the discussion so far, with reference to a representation of the type in (43) we have informally referred to material missing at PF as 'deleted' material. There are various views on how deletion should be handled (as discussed in Wilder 1997). For our purposes, it is imperative that we take the view that deleted material is present and 'syntactically active' throughout the derivation (crucially including S-structure; cf. the remarks about NPI licensing in section 2.1) and in LF. We thus reject the view that missing material is absent in S-structure (hence in PF), and gets 'syntactically reconstructed' or 'copied in' in the LF-component.

Two perspectives on ellipsis then remain: either (i) the missing phonological material gets deleted in the PF-component; or — on the view that phonological forms of lexical entries are only inserted after spell-out ('late insertion'; cf. the Distributed Morphology framework of Halle & Marantz 1993) — (ii) the missing phonological material corresponds to terminals of the syntactic representation where form-insertion simply fails to apply. We will not choose between these two, either choice being compatible with what follows; we will generally use 'ellipsis' and 'deletion' interchangeably, without intending a particular bias towards the approach in (i).

Deletion in elliptical answers is argued in Wilder (1997) to be a case of the same operation responsible for 'Forward Deletion'/'Forward Conjunction Reduction' in coordination. We call this operation FWD.

3.2 General constraints on ellipsis (FWD)

FWD is 'syntactically governed'. The units of deletion are syntactic units (though more than one unit may be affected in a given structure); parallelism ('context identity') conditions governing the antecedent-deletion relation refer to syntactic terms such as grammatical function; and the 'content identity' condition must be stated in terms of identity at LF (or perhaps, some notion of 'identity of meaning'). This contrasts with Backward Deletion (BWD), which is operative in so-called *Right Node Raising* constructions. BWD is 'string-governed', with units determined prosodically, and identity determined phonologically.

The main properties of FWD that are relevant here are listed in (44) (see Wilder 1997 for discussion of these in relation to coordination):

(44) *properties of Forward Deletion (FWD)*
a. <u>directionality</u>
The antecedent constituent precedes the deletion site
b. <u>syntactic parallelism</u>

The antecedent constituent and the ellipsis site fulfill the same grammatical function in their respective clauses

c. abstract identity
The antecedent and the deleted constituent are identical at LF

With regard to directionality, FWD seems to be completely strict — the antecedent may not follow the deletion site (BWD underlies the converse requirement, equally strictly; while other ellipsis types, e.g. VP-ellipsis, seem to permit their antecedent to precede or follow the deletion site). No answer — elliptical or not — can precede its question in a well-formed discourse; so the facts concerning QAP ellipsis are consistent with (44a). So too are the facts concerning Type A SPCs. Pseudoclefts with the order XP<*wh* (i.e., reverse Type B SPCs) cannot involve FWD; as a result, licensing an NPI in the XP of a reverse Type B SPC is predicted to be impossible — in conformity with the facts, as we have seen in the above.

The syntactic parallelism requirement (44b) can be thought of as a 'context identity' condition. It is in one sense strict — e.g. a subject cannot license deletion of an object (cf. **John looked at Mary and then she kissed ∅*, contrasting with *John looked at Mary and then ∅ kissed her*) — but in another sense it appears somewhat 'loose', in that it permits the deleted item to be in a different surface position than its antecedent. Likewise, the content identity condition (44c) is both 'strict' and 'loose' — 'strict' in that the LF-content of the deleted constituent must be strictly identical to that of the the antecedent constituent; but to some extent 'loose', in that FWD does not require identity of phonological content.

Taken together, (44b,c) permit an abstract but accurate account of the ellipsis-antecedent relation in QAPs and in SPCs. First consider the examples in (45):

(45) a. what didn't John do? — [∅ buy any books]
b. what John didn't do was [∅ buy any books]
c. John didn't buy any books

Here ∅ corresponds to *John didn't* in the *wh*-clause; in the answer/counterweight, *John didn't* is elided under perfect identity with the antecedent. Notice that in (45a) the negated dummy auxiliary is in a syntactic position different from the one that it is in in the full version of the answer, (45c). This is in perfect agreement with the conditions on ellipsis summed up in (44) — though (44b) demands functional parallelism, there is no constraint which says that the antecedent and the elided element have to occupy identical syntactic positions.

Things get more tricky in the non-negated counterpart of (45), given in (46). It is this paradigm which allows us to illustrate the importance of (44c) for QAPs and SPCs.

(46) a. what did nobody do? — [∅ buy any book]
b. what nobody did was [∅ buy any book]
c. ≠*nobody did buy any book
d. ≠nobody DID buy any book
e. ≠nobody bought any book
f. nobody PAST [buy any book]

Assuming that — just as in (45), where this is ensured by the presence of an NPI in the answer/counterweight — the answer/counterweight in the examples in (46a,b) is a finite declarative IP, the question arises of what the content of the ellipsis site ∅ is. ∅ must contain two deleted constituents, a subject and an auxiliary specified for tense. The subject causes no problem — it is a DP (*John* (in (45), or possibly its pronominal counterpart; we do not decide this here). The case of the auxiliary is less trivial. It cannot be *did* given the ill-formedness of (46c). It could be emphatic *did*, as in (46c), if it matched a parallel emphasis in the *wh*-clause; but the *wh*-clause does not have the meaning of a sentence with emphatic *do* (nor, for that matter, would we expect that a focused auxiliary could be deleted). The ordinary form of the clause expressing the desired neutral declarative meaning is (46e), but in this case, tense is realised on the main verb — clearly in conflict with what we see in the

elliptical answer/counterweight in (46a,b). The only option left is to assume that ∅ contains the PAST morpheme, as in the abstract representation in (46f), before it merges with the verb. So dummy *did* in the *wh*-clause is coupled with PAST in the answer. This is perfectly consistent with (44c). A deleted constituent need not be a literal phonological copy of the antecedent, as long as the LF identity condition is met. And of course, since the dummy *do* has no semantic contribution to make, semantic/LF identity (44c) is respected in (46).[9]

In both (45a) and (46a), the past-tense dummy auxiliary *do* in the question, undergoing Subject-Aux inversion, licenses deletion of the auxiliary in the IP of the answer even though, as pointed out, antecedent and elliptee sit in different syntactic positions. The conditions on ellipsis listed in (44) allow this (most crucially, (44b) is not violated). That it is indeed important not to demand parallelism with respect to syntactic positions when it comes to the licensing of ellipsis is shown also by 'asymmetric' coordinations of the type in (47) (first discussed in Wilder 1997; see Höhle 1991 and Heycock and Kroch 1994 for different views on related examples), this time with reference to subjects rather than auxiliaries. In (47), the postverbal subject in the first conjunct licenses FWD of the subject of the second conjunct. The ungrammaticality of (48) is evidence that the deleted subject is not postverbal in its clause. Hence, a postverbal subject can antecede a preverbal subject — consistent with (44b) :[10]

(47) a. [there ran out the bushes a huge fearsome bear] and [attacked us]
b. [out of the bushes ran a huge fearsome bear] and [attacked us]

(48) a. *there attacked us a huge fearsome bear
b. *attacked us a huge fearsome bear

(49) [out of the bushes ran **a huge fearsome bear**] and
[~~*a huge fearsome bear*~~ attacked us] <FWD

3.3 Maximal ellipsis

---

[9] We stress that the token of *do* that we are referring to in (46a) is the dummy auxiliary undergoing Subject-Aux inversion, *not* the main verb *do* following *John*. The distribution of the latter raises questions of an entirely different nature, which we cannot begin to broach here. Eventivity seems to be the most significant factor in licensing the use of *do* in the *wh*-clause of a QAP or SPC — thus note the contrast between **what John did was know French* and *what John did was speak French*, the latter grammatical only on an eventive reading of *speak French* (i.e. being involved in actual conversation in French), not on the alternative interpretation of this phrase which essentially parallels that of *know French*. There seems to be a reasonably strong correlation between the possibility for a VP serve as the answer/counterweight to a *what X do* clause, and the possibility for that VP to be used in the progressive to refer to present events; the 'progressive' test fails only in precisely those environments in which eventive verbs do not occur in the progressive (e.g. infinitives): *what I'll try to do is be home by 6 pm*. Besides eventivity, agentivity of the subject also seems to play an important role. Exactly how the distribution of main verb *do* in QAPs and SPCs can be captured is a question we cannot answer at this time.

In connection with this, also note that *do* can introduce an argument of its own in the form of a PP in the *wh*-clause, the object of P corresponding to the (typically affected) object of the verb in the answer/ counterweight (cf. *what John did to the book was burn it*; see Higgins 1979:201 for other cases). The question that these cases pose is how the correspondence between the object of the PP in the *do* clause and the object of the main verb in the answer/counterweight is given shape. We can think of two suggestions to come to terms with this difficult question: either (i) one assumes that in constructions like these *what+do+to NP* forms one complex *wh*-phrase, matched (or replaced, as in Bošković 1997) by the VP in the answer/counterweight; or (ii) one assumes that the PP is actually present, abstractly, in the answer/counterweight as well. The latter approach would encompass a general perspective on the representation of 'affected' or 'experiencer' arguments — a sentence like *John burnt the book* would then have an underlying representation in which the object appears twice, once as the object of *burn* and once as the object of a covert preposition (cf. (*)*John burnt the book to itself*). We have neither the means nor the space to pursue suggestions along either of these lines; suffice it to say that, whatever the proper treatment of *what John did to the book was burn it* may be, it will not distinguish between the various overall approaches to SPCs reviewed in section 1. We refer ahead to fn. 25, below, for a possible alternative approach to the problem posed by *do+to* cases.

[10] The assumption of a syntactically represented indefinite in the second conjunct of (49) raises questions about how to guarantee the correct interpretation. A possible alternative is that the deleted subject such examples is a 'pronominal correlate' of its antecedent.

Pseudoclefts whose counterweight is a (surface) VP may not contain a finite verb, as seen in (50a). Similarly, a finite VP may not form an elliptical response in a QAP (cf. (50b)), though the corresponding IP-pseudocleft/IP-answer are both perfect (51):[11]

(50) a. *what John did was bought a book
b. what did John do? — *?bought a book
(51) a. what John did was [he bought a book]
b. what did John do? — he bought a book

What this indicates is that ellipsis of the subject alone is not possible in such examples, despite the fact that it is in principle recoverable. This can be made sense of if it is assumed that the appearance of the finite inflection on the main verb signals that Infl has not undergone ellipsis (cf. (52), where 'A' marks the counterweight, and 'XP' is the focused constituent in A).

(52) a. *[what John did ___] was [$_A$ ~~*John*~~ PAST [$_{VP=XP}$ buy a book]] ('bought a book')
b. [what John did ___] was [$_A$ ~~*John PAST*~~ [$_{VP=XP}$ buy a book]] ('buy a book')

The same pattern is found in German (53):

(53) a. *was Hans jetzt tut ist [kauft ein Buch]
what Hans now does is buys a book
b. was Hans jetzt tut ist [er kauft ein Buch]
what Hans now does is he buys a book
c. ?was Hans jetzt tut ist [ein Buch (zu) kaufen]
what Hans now does is a book to buy

The case illustrated in (50)-(53) falls under a wider generalisation. If *any* constituent of the answer is targeted by FWD under identity with antecedents in the question, then *every* such constituent must be deleted.

(54) *maximal ellipsis*
if A undergoes ellipsis, ellipsis must be *maximal* (down to XP)
[where 'XP' = the focused constituent in A; and 'A' = answer/counterweight]]

The effects of (54) are also apparent in the contrast between (52b) and (55)-(56). (52b) indicates that deletion of the answer up to VP (i.e. deletion of the subject and of Aux) is possible. However, this deletion pattern is not licit where the same clause answers the question in (55)-(56). This is due to (54); since the verb can be deleted in (55)-(56), either it must be deleted along with the subject and Aux, or else no deletion applies:

(55) what did John buy?
a. [John bought [$_{XP}$ a book]]
b. [he bought [$_{XP}$ a book]]
c *[buy [$_{XP}$ a book]] *ok in (52b)*
d. [$_{XP}$ a book]

(56) a. what John bought was [$_{IP=A}$ John bought [$_{XP}$ a book]]
b. what John bought was [$_{IP=A}$ he bought [$_{XP}$ a book]]
c. *what John bought was [buy [$_{XP}$ a book]] *ok in (52b)*
d. what John bought was [$_{XP}$ a book]

[11] The fact that the answer in (50b) is perhaps less deviant than the counterweight in (50a) may be attributable to the (marginal) possibility of a 'diary drop' analysis of the answer in (50b); cf. Haegeman (1990) and Wilder (1997).

The ill-formedness of (55c), (56c) and (50) is reminiscent of obligatory ellipsis effects found in gapping constructions (cf. Williams 1997). Neither the subject in (57a) nor the indirect object in (57b) may be overtly realised if it meets the LF-identity requirement for FWD:

(57) a. he_j bought a book and he_k ~~*bought*~~ a record
a'. *he_j bought a book and he_j ~~*bought*~~ a record
b. she gave him_j a book and ~~*gave*~~ him_k record
b'. *she gave him_j a book and ~~*gave*~~ him_j a record

It is plausibly the case that the ill-formedness of (57b) and (57d) on the one hand, and of (55c), (56c) and (50) on the other, is due to a single principle, a generalized version of (54) governing FWD.

3.4 Deletion up to but not into XP

Let us now return to the puzzle we were left with at the end of section 2 — the question of why (43b), repeated below as (58), is apparently an illegitimate representation.

(58) *[what *t* didn't happen next] was [$_A$ ~~*it didn't happen that*~~ anybody fell]

The reason for the impossibility of (58) is to be sought in the factors determining maximal extent of FWD. The remnant of ellipsis in (58) must include the complementiser.

Notice that it is not possible to maintain, in any general sense, that FWD of subconstituents of a finite complement CP is impossible. In (59), FWD deletes constituents of the main clause (including the negation) and of the complement clause, and the result is grammatical:

(59) ?what he didn't say he bought was any wine.
(60) [what he didn't say he bought] was [$_A$ ~~*he didn't say*~~ [~~*that he bought*~~ [$_{XP}$ any wine]]]

Rather, the maximal extent of FWD in the answer/counterweight is determined by the *focus-background structure* of that answer/counterweight, which in turn is determined by form of the *wh*-clause; in particular, of the moved *wh*-phrase. XP, the remnant of FWD in an answer/counterweight, is a *focus phrase*. In a given answer/counterweight, XP is determined with respect to the (surface) form of the corresponding *wh*-clause. XP must correspond to the overtly moved *wh*-phrase, including pied-piped material (even if pied-piped material is 'reconstructed' at LF).

Now notice, crucially, that FWD cannot remove subparts of a focus phrase. This fact is illustrated by the contrast in (61). Though the adjective *fast* may suffice 'semantically' to answer the question in (61a), it does not constitute a well-formed elliptical response to that question (though it does to the question with *how* in (61c)). The surface form of the moved *wh*-phrase determines that the object DP forms the focus constituent in the answer (62b). The deletion in (62a) is illicit.

(61) a. what kind (of car) does he drive? — *fast
b. what kind (of car) does he drive? — a fast car/a fast one
c. how does he drive? — fast
(62) a. *[$_A$ ~~*he drives*~~ [$_{XP}$ ~~*a*~~ [fast] ~~*car*~~]]
b. [$_A$ ~~*he drives*~~ [$_{XP}$ a fast car]]

The same point is illustrated by pied-piped possessors, as in (63). The form of the overtly raised *wh*-phrase determines that the object of the verb is the focus phrase of the response. The ellipsis necessary to generate (63b) represents an illicit deletion of a subpart of the focus XP:[12]

---

[12] The ellipsis of NP in the more natural response (i) is licensed by a different operation, independent of FWD; note the grammaticality of *my father's is a very nice car*, which cannot possibly be got via FWD (since the putative antecedent of the ellipsis site does not precede the ellipsis site here).
(i) [$_{DP=XP}$ my father's [$_{NP}$ Ø]]

(63) a. [whose car] did Mary borrow *t*?
b. *[$_{IP=A}$ ~~Mary PAST borrow~~ [$_{DP=XP}$ [my father]~~'s car~~]]
c. [$_{IP=A}$ ~~Mary PAST borrow~~ [$_{DP=XP}$ my father's (car)]]

Returning to the case of SPCs like *what happened next was (that) he fell*, we can observe that the focus phrase in the counterweight to the *wh*-clause is determined — by the form of the *wh*-clause — to be the subject of the verb *happened*. Hence, an elliptical response in which the subject of *happened* is realised by a clause must leave that clause *qua* focus phrase intact. Hence, there is no ellipsis derivation for 'IP-pseudoclefts' of the type **what didn't happen next was anyone fell* as in (58), or, for that matter, for *what happened next was he fell* as in (65):[13]

(65) *[what ___ happened next] was [$_{A}$ ~~*it happened*~~ [$_{XP}$ ~~*that*~~ he fell]]

Even for examples of the type (66) there are reasons to believe that there can be no ellipsis derivation, as in (67), though the missing complementiser is independently licensed, as seen in (68):

(66) what he says is he'll leave
(67) (*)what he says is [$_{IP=A}$ ~~he says~~ [$_{CP=XP}$ ∅ he'll leave]]
(68) he says [$_{CP}$ ∅ he'll leave]

In particular, the fact, illustrated in (69), that polarity items are not licensed by negation in the *wh*-CP if the complementiser is not present — reproducing the pattern just discussed for *wh*-CPs with *happen* — shows that a derivation à la (67) is unavailable here, too.[14] The reason why, in spite of the fact that

---

[13] Quirk *et al.* (1985:1062n.) mention an interesting example of an SPC featuring a *for-to* infinitival, reproduced here as (ia), noting that the infinitival complementiser can optionally be left out; similarly, (ib) (an example of our own, a case whose corresponding simple sentence features *for* obligatorily).

(i) a. what he didn't like was (*for*) me to be alone at night
b. what he likes best is ?(*for*) her to call often (cf. he likes best *(*for*) her to call often)

Facts like these may suggest that — unlike *that* — *for* can be a target for ellipsis. In this connection, note that, in parallel to examples like *John talked to Mary on Monday and Bill on Tuesday* (cf. Pesetsky 1995), we find coordinations in which *for* is left out in the second conjunct (cf. (ii)).

(ii) John liked very much for her to call often and him to keep silent

While (i) illustrates a case in which a *for* which is obligatory in the corresponding sentence can optionally be left out in the SPC, the converse also seems to exist: a case in which a *for* shows up obligatorily in an SPC which cannot show up in the corresponding simple sentence:

(iii) a. what I cannot believe is *(*for*) him to be top of the class
b. I cannot believe (**for*) him to be top of the class

The fact that (iiia) is ungrammatical without *for* follows straightforwardly from a Type A analysis, as a parallelism effect (cf. section 4.2): the complement of *believe* in the *wh*-clause is nominal (*what*) while the one in the counterweight is clausal (the ECM infinitival); cf. Bošković (1997) for a different account, based on an analysis of SPCs which we believe is false (see the discussion in sections 4 and 5). The fact that (iiia) is acceptable with *for* raises questions which can only be properly addressed once the broader questions that the *for-to* infinitival construction raises in general (e.g. concerning the proper analysis of ECM in these constructions, from the checking perspective of Chomsky 1995); we will put the specific questions raised by (i)-(iii) on reserve for the moment, pending the answers to the more general questions.

[14] *Anyone* in (69a) does not have an NPI-reading; a free choice reading is possible, but this is licensed by the modal in its own clause, hence irrrelevant.

Notice also that a pronoun in the postcopular constituent cannot be bound to a QNP in the *wh*-CP in these cases, *unless* the complementiser is present:

(i) a. what everyone$_j$ says is {*he$_j$/John} will leave
b. what everyone$_j$ says is that he$_j$ will leave
c. what no student$_j$ claims is {*he$_j$/John} can solve this problem
d. what no student$_j$ claims is that he$_j$ can solve this problem

Thus, the examples lacking the complementiser also fail to display a core connectedness effect, QNP-bound variable readings. This fact is significant, since bound variable anaphora is otherwise quite 'liberal' as far as connectedness effects go. It is also surprising, since in an obvious semantic sense, the content of the postcopular IP in (ia,c) is interpreted in the scope of the verb of saying in the *wh*-CP.

the verbs in the *wh*-clause in (66)-(69) license a Ø-complementiser, the a-examples cannot get an ellipsis derivation becomes clear when we realise that in gapping (70) and pseudogapping (71) constructions the Ø-complementiser normally licensed by verbs like *think* and *say* is not licensed when the verb itself is deleted (cf. Wilder 1997):

(69) a. *what John never says is anyone is allowed to leave
b. what John never says is that anyone is allowed to leave
(70) John said (that) Mary would win and Bill ~~said~~ *(that) Sue would win
(71) John said (that) Mary would win, while Bill did ~~say~~ *(that) Sue would win

So we have found a rationale for the ungrammaticality of (43b)/(58). What now remains to be said with respect to (30b), *what happened next is he fell*, is how it *can* be legitimately derived.

3.5 Direct vs indirect answers

What we were forced into concluding is that (30b) is effectively a Type A SPC — but not one involving ellipsis but instead one featuring an *indirect answer* as the counterweight:

(72) what happened next was [he fell] (Type A — 'indirect answer')

This is the only analysis of (30b) that is left to us. After all, a construal of (72) as a Type B SPC with an IP counterweight is unavailable for reasons discussed in section 2.3.2 (Type B SPCs *never* have IP counterweights, since IPs are impossible as subjects); and we just found an explanation for why a Type A *cum* ellipsis-into-CP analysis is illicit as well. Far from manoeuvring us into tight straights, the conclusion that (72) *qua* Type A SPC is the only possible approach to (30b) prompts a discussion — called for anyway — of the parallelism between QAPs and SPCs in the domain of direct *vs* indirect answers.

The ellipsis approach treats specificational pseudoclefts (of Type A) as 'self-answering questions'. The relation between the focus XP and the *wh*-CP is claimed to be the same as that between a (constituent) question and its answer. To maintain this view, it is necessary to show that there are pseudoclefts corresponding to the various types of QAPs. Where the two differ, an account is needed of how and why.

It turns out that the relation between *wh*-CP and XP is more tightly constrained in pseudoclefts than in QAPs. We suggest that this difference should be linked to an intuitively plausible distinction among possible question-answer relations. *Wh*-questions can receive 'direct answers' or a variety of answers of a more or less 'indirect' type; pseudoclefts, on the other hand, usually realise only 'direct' question-answer relations, with a couple of exceptions to the general rule — among which, of course, the example in (30b), now analysed as (72).

Consider (73). *Wh*-questions do not require elliptical answers; the answer can comprise a whole clause. To count as a direct answer, however, a response may differ in content from the question only with respect to material corresponding to the wh-phrase of the question:

(73) what did John buy?
a. John bought [a book]
b. he bought [a book]
c. I believe that John bought [a book]
d. I don't know (what John bought).
e. [Bill]$_F$ bought [a book] (... but I don't know what John bought)

This conjunction of properties is also observed in constructions involving sentence parentheticals (SPs) — the content of the main clause is interpreted in the scope of the verb of the SP, yet bound variable anaphora between the two clauses is not possible:

(ii) a. *Mary will, everyone$_j$ thinks, visit him$_j$
b. *Mary does, nobody$_j$ doubts, like him$_j$

The pronouns in (ii) cannot be construed as bound by the QNP subject of the SP verb.

(73a,b) correspond to direct answers to the question. (73c,d,e) count as indirect answers. (73c) directly answers a slightly different question: 'what do you believe that John bought?' (73d) can be thought of as directly answering the associated yes-no question 'do you know what John bought?', whose positive answer is a felicity condition on the illocutionary act of asking the question in (73). (73e) is only felicitous with a special intonation on the constituent marked F, which signals a change in the discourse topic; see Büring (1995) for a discussion of such 'contrastive topics'.

Well-formed pseudoclefts can be constructed only for (73a,b). Thus there is a tighter relation between the *wh*-CP and the IP in pseudoclefts than governs QAPs. In particular, the 'response' in a pseudocleft requires (semantic) identity of all constituents in A with parallel constituents in the question — except for XP, its focus constituent (corresponding to the *wh*-phrase of the 'question'):

(74) a. [what John bought ___] was [John bought [a book]]
b. [what John bought ___] was [he bought [a book]]
c. *[what John bought ___] was [I believe that John bought [a book]]
d. *[what John bought ___] was [[Bill] bought [a book]] ...
e. *[what John bought ___] was [I don't know ___]

The examples (74a,b) can be thought of as 'undeleted' pseudoclefts; i.e. self-answering questions whose answer is nonelliptical. They instantiate pseudoclefts where the postcopular constituent is an IP (in (74), IP is identical to A, but not identical to XP; rather, IP contains XP). As we pointed out at the start, the existence of IP-pseudoclefts provides one strong argument favouring the approach to Type A pseudoclefts as possibly elliptical self-answering questions.

Even though the SPCs in (74c-e) — corresponding to the perfect QAPs in (73c-e) — are ungrammatical for lack of 'directness', it would be wrong to say that pseudoclefts never instantiate indirect question-answer relations. We have seen that (30b), analysed as in (72), features an 'indirect answer' Type A SPC; and in fn. 6, and (66) and the accompanying discussion in fn. 12, above, we came across other examples of 'indirect answers' Type A SPC. The one from fn. 6 (once again involving the verb *happen*) is reproduced below, along with its QAP companion, in (75):

(66) what he says is he'll leave
(75) a. ?what didn't happen (at least) was {[it didn't rain]/[John didn't leave]}
b. ?what didn't happen (at least)? — {[it didn't rain]/[John didn't leave]}

So what we can conclude is that, under certain conditions, Type A SPCs can mimic QAPs very closely in even allowing for indirect 'answers' (i.e. counterweights not corresponding directly to any constituent of the *wh*-clause). Of course the discussion here begs a big question — the question of under what circumstances a pseudocleft permits an 'indirect answer' as its counterweight. This is a tricky one, which we will have to skirt at this time, for want of any particular insights to offer.

3.6 Conclusion

In this section we have reviewed the restrictions on ellipsis in QAPs and Type A SPCs, *en passant* noting that, to a limited extent, SPCs allow for 'indirect answer' counterweights. Such cases are instantiated by the example that prompted the discussion of conditions on ellipsis — (30b), *what happened next was he fell* — for which no Type B or Type A *cum* ellipsis analyses are available on principled grounds. The fact that Type A SPCs accept (IP) counterweights which do not directly correspond to anything in the *wh*-clause (though they 'answer' the subject of the *wh*-clause, they cannot actually fit in there) enhances the parallelism between these SPCs and QAPs, a connection which we will explore in further detail in the next section.

4 Specificational pseudoclefts and *wh*-questions

4.1 *The wh-constituent is an interrogative CP, not a free relative*

In the foregoing we have seen that there are strong connections between Type A SPCs and Question-Answer pairs. This suggests that the *wh*-constituent of a Type A SPC is a *wh*-interrogative, not a free relative. In this section we will compound arguments to further underpin this claim.[15]

4.1.1 *No wh+*ever *in specificational pseudoclefts*

One immediate indication that a *wh*-interrogative analysis is to be preferred to a free relative approach comes from the fact that the *wh*-constituent of English specificational pseudoclefts resists the addition of *-ever*, which is a convenient diagnostic to distinguish between *wh*-interrogatives and free relatives: undisputed cases of free relatives accept *-ever* but *wh*-interrogatives do not (cf. Iatridou & Varlokosta 1995):

(76)	[what(ever) he is] is immaterial	[FR]
(77)	I wonder [what(*ever) he is]	[*wh*]
(78)	[what(*ever) he is] is proud	[SPC]

For English at least, the facts in (76)-(78) cast doubt on an analysis of the *wh*-clause of a specificational pseudocleft as a free relative.[16]

4.1.2 *Topicalisation across the wh-phrase*

Further such doubt is prompted by the fact (not previously observed in the literature, to our knowledge) that English specificational pseudoclefts pattern with *wh*-interrogatives and differ from free relatives in allowing topicalisation across the *wh*-phrase. Consider the facts in (79)-(81):

(79)	*[to Mary, what he gave] caused a scandal
(80)	?[to Mary, what will he give]?
(81)	?[to Mary, what he won't give] is any wine

The acceptability of the *wh*-interrogative in (80) (observed by Emonds 1976, Pesetsky 1989 and others) is matched by that of the SPC in (81), which differs markedly from the free relative in (79).

---

15 See Hankamer (1974) and Bošković (1997) for additional evidence against a free relative analysis of the *wh*-clause of English specificational pseudoclefts; also note that in the literature on SPCs in American Sign Language a standard claim is that these have the syntax of Question-Answer pairs (but cf. Wilbur 199x for a different view).

16 Iatridou & Varlokosta (1995) explain the deviance of (78) by appealing to an analysis of all SPCs (not just our Type B) à la Heggie (1988), Heycock (1991), according to which the *wh*-clause is a predicate of the counterweight. The ban on *-ever* in SPCs then follows on the assumption that *-ever* turns the *wh*-clause into a quantificational phrase; as a consequence, the *wh*-clause can no longer act as a predicate (on the assumption that QPs do not qualify as predicates) once *-ever* is added to the *wh*-phrase in SpecCP.

Dayal (1997) claims that FR *-ever* is possible in SPCs in case the main-clause copula is negated, as in:

(i) ?whatever she bought was not *Barriers*

It seems to us, however, that we are not dealing with a specificational pseudocleft in (i). An immediate indication that (i) is not an SPC is the very fact that it features a matrix negation — something which, as Higgins (1979) observed, a genuine SPC can never do in English (cf. **what John is isn't proud of himself*). Moreover, two properties of (i) which would be altogether mysterious if it was an SPC (of Type B) are the fact that that it is irreversible (cf. (ii)) and entirely fails to be embeddable under ECM-verbs taking *to*-infinitival complements (compare (iii) with the appreciably better *?I consider what John is to be important to himself*) . We suggest that what underlies (i) is the construction in (iv), where *whatever she bought* functions as an adjunct topic. From this analysis the irreversibility and unembeddability of (i) follows straightforwardly: movement across topics is impossible, and topicalisation in ECM-infinitivals is excluded.

(ii) **Barriers* was not whatever she bought

(iii) *I consider whatever she bought not to be *Barriers*

(iv) whatever she bought, it was not *Barriers*

Notice also that (81) seems to be a hybrid *wh*-clause, mixing properties of root and embedded questions. Though the *wh*-clause lacks the Subject-Aux inversion effect typical of root questions, it does exhibit the word-order of root *wh*-questions featuring topicalisation when it comes to the relative placement of topic and *wh*-phrase, differing in this regard from embedded *wh*-interrogatives with topicalisation, in which the topic linearly follows the *wh*-phrase, as shown in (82):

(82) a. ?I wonder [what to Mary, he will give]
b. *I wonder [to Mary, what he will give]

The lack of Subject-Aux inversion in (81) indicates that the *wh*-clause is not a root sentence; but the fact that the topic attaches outside the *wh*-phrase suggests — on the assumption that the attachment site of topics in *wh*-questions reveals something crucial about the clause's thematic function (argument or predicate); see Chomsky (1986, 1995) — that the *wh*-clause is not an argument or predicate of the matrix clause either. We will come back to the implications of this in section 5.

4.1.3 Pied-piping

The previous arguments for our claim that the *wh*-clause of an SPC behaves like an interrogative clause, not as a free relative, came from English. In this subsection we will present an interesting argument to the same effect based on the distribution of pied-piping in German and Dutch (cf. also Higgins 1979:41 on what look like similar facts from Spanish).

Let us set the stage for the discussion of Dutch and German to follow by considering pied-piping in English. As (83) shows, English SPCs robustly resist pied-piping. In this regard, the English SPC at first blush seems to pattern with free relatives (cf. (84)), not with *wh*-interrogatives (cf. (85)):

(83) a. *with whom he went to school was with Mary
b. who he went to school with was Mary
(84) a. *with whom he went to school was stupid/has just entered the room
b. who we went to school with was stupid/has just entered the room
(85) a. ?with whom did he go to school?
b. who did he go to school with?

Upon closer inspection, though, the facts in (83)-(85) do not overturn the *wh*-interrogative analysis of the *wh*-constituent of specificational pseudoclefts. For as Kayne (1994:25) points out, while acceptable in more formal registers, *wh*-interrogatives with pied-piping of the type shown in (85a) are not possible in colloquial English. The examples used to support his claim are reproduced here as (86):

(86) a. *we want to know about what you are thinking
b. *tell me at whom you were looking

Coupled with the fact that specificational pseudoclefts are a typical feature of the colloquial language rather than of the more formal registers, this defuses a potential argument against a *wh*-interrogative approach to the *wh*-constituent of SPCs.

So the English pied-piping facts are at least *compatible* with the analysis of SPCs that we are advocating here. But we can do much better than this, by considering the distribution of PP pied-piping in Dutch and German. In these languages, unlike in English, PP pied-piping is not restricted to formal registers in questions; it *is*, however, completely impossible in free relatives (barring possible matching contexts, which are irrelevant). This is shown in (87) and (88):

(87) a. met wie heeft Maria gesproken? [Dutch]
b. mit wem hat Maria gesprochen? [German]
with whom has Maria spoken
(88) a. *met wie Maria gesproken had kwam zojuist de kamer binnen [Dutch]

b. *mit wem Maria gesprochen hatte kam gerade ins Zimmer hinein [German]
with whom Maria spoken had came just (into) the room inside

Interestingly, now, specificational pseudoclefts featuring PP pied-piping are grammatical in both Dutch and German, as shown in (89):

(89) a. met wie Maria gesproken had was met Peter [Dutch]
b. mit wem Maria gesprochen hatte war mit Peter [German]
with whom Maria spoken had was with Peter

So again we see that the *wh*-constituent of SPCs behaves like a *wh*-clause, not like a free relative.

4.1.4 Multiplicity

A spectacular set of facts underscoring the same point involves multiple *wh* SPCs. Consider the examples in (90) (from Ross 1997) and (91) (from Meinunger 1997):

(90) [who ordered what] was [Tom (ordered) a beer and Jim a watermelon flip]
(91) a. [wer hier wem geholfen hat] war [die Hilde dem Heinz] [German]
who here whom helped has was the Hilde the Heinz
b. *wer hier wem geholfen hat scheint die Hilde dem Heinz zu sein
who here whom helped has seems the Hilde the Heinz to be
c. *die Hilde dem Heinz war wer hier wem geholfen hat
the Hilde the Heinz was who here whom helped has

These SPCs, which — as the deviance of the examples in (91b,c) shows — are unequivocally of Type A, further the cause of the *wh*-interrogative approach to the *wh*-clause of SPCs. After all, the alternative free relative approach here has to contend with the fact that neither in English nor in German do we otherwise come across cases of multiple relativisation, while multiple *wh*-questions are perfectly common in these languages.

4.1.5 Case connectedness

German furnishes a fifth argument for an analysis of the *wh*-constituent of SPCs as *wh*-interrogatives, coming from the domain of case. In German an object *wh*-question such as (92a) is answered as in (92b), with an accusative-marked DP, not as in (92c); a predicate *wh*-question such as (93a), by contrast, can only receive a nominative-marked answer.

(92) a. was hat er schon immer kaufen wollen?
what has he PRT always buy want
'what has he always wanted to buy?'
b. einen Audi
a-ACC Audi [make of car]
c. *ein Audi
a (NOM) Audi
(93) a. was ist er?
what is he
'what is he?'
b. ein Arzt
a (NOM) doctor
c. *einen Arzt
a-ACC doctor

The interest of this for the discussion of SPCs lies in the fact that a pseudocleft like (94) can feature an accusative-marked counterweight alongside a nominative one, while (95) can only get nominative marking on the counterweight:[17]

(94) was er schon immer kaufen wollte, ist ein/einen Audi
(95) was er ist, ist ein Arzt/*einen Arzt

Moreover, while the nominative variant of (94) allows for inversion, embedding under the raising verb *scheinen*, or the addition of modal auxiliaries to the matrix clause, the accusative one does not, as seen in (96):

(96) a. was er schon immer kaufen wollte scheint ein Audi/*einen Audi zu sein
what he PRT always buy wanted seems a(NOM)/*a-ACC Audi to be
b. was er schon immer kaufen wollte hätte ein Audi/*einen Audi sein können
what he PRT always buy wanted would-have a(NOM)/*a-ACC Audi be can
c. ein/*einen Audi ist was er schon immer kaufen wollte
a(NOM)/*a-ACC Audi is what he PRT always buy wanted

By contrast, it is precisely the accusative contender that wins out in SPCs involving NPI connectivity, as shown in (97) (intuitions on (97) are less robust than they are on (96)):

(97) a. ?was er niemals kaufen würde ist [auch nur irgendeinen japanischen Wagen]
what he never buy would is also only any-ACC Japanese car
b. *was er niemals kaufen würde ist [auch nur irgendein japanischer Wagen]
what he never buy would is also only any(NOM) Japanese car

This said, the account of 'case connectedness' falls readily into place: the accusative variants of the pseudocleft examples presented above are are all Type A SPCs, hence they are grammatical on the *wh*<XP order only, and they resist embedding under raising verbs or addition of modal auxiliaries (like all Type A SPCs); by contrast, it is exactly these accusative variants that, by virtue of their Type A status, cater for the licensing of NPIs in the counterweight. The 'undeleted' source for the accusative variant of (94) thus reads as in (98), which is in effect grammatical in German (cf. Weinert 1995 for attested examples):

(98) was er schon immer kaufen wollte ist er wollte einen Audi kaufen

For the QAP in (92) an essentially similar story presents itself, the example in (90b) involving an elliptical full-IP answer, and the alternative in (92c) being ungrammatical because, apparently, questions always receive a clausal answer (in German at least). And of course no accusative case will ever be made available for the answer in (93c) or the counterweight in (93) since, unlike in the case of (92) and (94), there is simply no verb in the elliptical IP here which can assign this case to the DPs involved.

So there is a perfect match between (the Type A versions of) the SPCs in (94) and (95) and the QAPs in (92) and (93). And moreover, there is a contrast here between SPCs and relative clause

---

[17] Iatridou & Varlokosta (1995) argue that the two variants of (92) differ in interpretation — the nominative variant is a predicational pseudocleft while the accusative one is an SPC. It can be shown, however, that (94) with a nominative counterweight *can* be specificational (see the BT-A and QP/bound-vbl. connectivity effects in (i) and (ii); cf. also Sharvit 1997 on superficially similar cases from Hebrew), but that, whenever it is, it can only be of Type B (see below, main text); the accusative variant of (92) is a Type A SPC.

(i) was *er* schon immer lesen wollte ist ein Artikel über *sich*
what he PRT always read wanted is a(NOM) article about REFL
(ii) was *niemand* lesen will ist *sein* erster Artikel
what nobody read wants is his(NOM) first article

constructions. For as (99) shows, relative constructions do not exhibit 'case connectedness' of the sort found in (92) and (94).[18]

(99) das, was er schon immer kaufen wollte, ist ein/*einen Audi
that which he PRT always buy wanted is a(NOM)/*a-ACC Audi

The Case connectedness effects reviewed in this section confirm not just the correlation between QAPs and Type A SPCs but also the ellipsis approach to the latter — both central tenets of the approach to SPCs taken in this paper.[19]

4.1.6 Participial connectedness

There is another type of connectivity effect in SPCs and QAPs — one rarely discussed in the literature (or at least not in terms of 'connectivity') — emanating from an interesting observation made with respect to SPCs in Quirk *et al.* (1985:1388; we have adapted their examples slightly, to

---

[18] Things highly similar to those reported for German in this subsection are true for Russian pseudoclefts (which are structured in the following way: *wh*-clause – *tak eto* ('then this-is') – counterweight). If the *wh*-constituent does not originate in a subject position, nominative or instrumental (which is also found in ordinary copular sentences) is impossible. The counterweight must show the same case as the *wh*-constituent:

(i) cto on vypil tak eto vodku / *vodkoj / *vodka (Izvorski 1997)
what he has-drunk 'tak eto' vodka-ACC / vodka-INST / vodka-NOM
'what he drank was (the) vodka'

'Undeletion' is possible:

(ii) cto on vypil tak eto on vypil vodku

And although Russian is very liberal when it comes to word order, reversing the *wh*-clause and the counterweight is impossible in these pseudoclefts:

(iii) *vodku tak eto cto on vypil

[19] Apparent breakdown of Case connectedness is found in the following examples of SPCs and QAPs, due to Sharvit (1997) (italicisation marks intended variable binding). In (i)-(vi), we are concerned with the reading in which the pronominal is understood as being bound by the italicised QNP in the whCP.

(i) who does *every professor* think *t* should get tenure? — *himself/*him/*he*

(ii) who *every professor* thinks *t* should get tenure is *himself/*him/*he*

These cases pose two problems: besides the Case form of the answer/counterweight (accusative rather than the expected nominative), the fact that it can only be an anaphor, not a pronoun (not, at least, if it is to be bound by *every professor*), also seems surprising, in the light of the fact that in the 'corresponding' simple sentence in (iii) only *he* is possible:

(iii) *every professor* thinks *he* should get tenure

There is evidence, however, to indicate that the form that surfaces in (i)-(ii) is in fact the *emphatic reflexive*, as in (iv); as is well known, the emphatic reflexive of English surfaces in the accusative Case form only (**he heself*), hence the facts in (i)-(ii) are less dramatic than they seem at first blush.

(iv) *every professor* thinks that *he himself* should get tenure

The evidence that we are dealing with emphatic reflexives in (i)-(ii) is twofold. First, in German (v), the reflexive anaphor *sich* is impossible; the only option in this context is the nominative pronoun *er* accompanied by the emphatic reflexive particle *selbst*. Secondly, Sharvit herself notes an example provided to her by A. Kroch that indicates that reflexive and pronominal elements used in elliptical responses are governed by the Binding Theory in a manner consistent with the ellipsis theory (but irreconcilable with her own, semantic approach) — see (via), where the deviance of the reflexive can be attributed to the fact that emphatic reflexives are subject to the Specified Subject Condition, as confirmed by (vib).

(v) a. wer denkt *jeder Professor t* soll eine unbefristete Stelle bekommen?
who thinks every professor should a tenured position get

b. **sich/*ihn/??er/er selbst*
refl/he(acc)/he(nom)/he(nom)+emph.refl

(vi) a. ?who *every professor* likes [Mary's picture of *t*] is *him/*himself*

b. *every professor* likes [Mary's picture of *him/*himself/*him himself*]

As for the impossibility of the *nominative* pronoun in the English examples in (i)-(ii), it suffices to note that it can never be used in isolation; cf. Cardinaletti & Starke (1994), and the contrast in (vii):

(vii) a. who came? — John (did)

b. who came? — he *(did)

make the point clearer). They point out that both (100a) and (100b) are grammatical as counterweights of the *wh*-clause in (100). To this observation we add that (100c) is ungrammatical, and that while the 'perfect match' full-IP counterweight in (100e) is impeccable, the imperfective full-IP counterweight in (100d) is marginal at best. These observations can be duplicated for QAPs, as shown in (101). Of course, this is what our approach to (Type A) SPCs and QAPs leads us to expect.

(100) what he has done is
- a. [taken some pictures]
- b. [take some pictures]
- c. *[took some pictures]
- d. ?[he took some pictures]
- e. [he has taken some pictures]

(101) what has he done?
- a. [taken some pictures]
- b. ?[take some pictures]
- c. *[took some pictures]
- d. (?)[he took some pictures]
- e. [he has taken some pictures]

And note that, for both SPCs and QAPs alike, no past participle can be found in the counterweight or answer when the *wh*-clause is imperfective, as shown in (102) and (103):

(102) what he did was
- a. *[taken some pictures]
- b [take some pictures]
- c. *[took some pictures]
- d. [he took some pictures]
- e. ??[he has taken some pictures]

(103) what did he do?
- a. *[taken some pictures]
- b. [take some pictures]
- c. *[took some pictures]
- d. [he took some pictures]
- e. ?[he has taken some pictures]

The ungrammaticality of the c-examples in the above was discussed in section 3 in connection with maximal ellipsis; but the other examples, when taken together, raise a novel point, which we will briefly comment upon in the context of the ellipsis approach to SPCs and QAPs.

Taken at face value, it would seem that, of the a- and b-examples in (100) and (101), the former are straightforward cases of Type A SPCs/QAPs, with the participle in the counterweight/ answer licensed under ellipsis, while the b-cases, exhibiting no participial connectivity, instantiate Type B. But what flies in the face of such a classification of the facts is that the a- and b-examples in (100) each allow inversion and NPI licensing — particularly the fact that, as (104) and (105) show, inversion and NPI licensing seem just as acceptable in the a-examples of (104) and (105) as they are in the corresponding b-examples indicates, from the perspective we have developed on the analysis of SPCs, that neither the a-case nor the b-case instantiates Type A.

(104) a. [taken some pictures] is what he has done
- b. [take some pictures] is what he has done

(105) a. what he has not done is [taken any pictures]
- b. what he has not done is [take any pictures]

So a squish between ellipsis and non-ellipsis (Type A and Type B) is not what we are dealing with. Instead, to get the optional participial connectivity effect under control, what we will capitalise on is the alternation between (106a) and (106b) (the latter noted in Emonds 1976):

(106) a. John said that he would take some pictures, and [taken some pictures], he has
- b. John said that he would take some pictures, and [take some pictures], he has

The pair in (106) shows that, when detached from the auxiliary of the perfect, the main verb phrase can show up in either of two morphological forms — a participial or a plain infinitival form. This of course reminds us of the alternation between (100/101a) and (100/101b), a link which is enhanced further by the fact that, at least in British English, the auxiliary of the perfect in (106) can be followed by *done*, which also surfaces in the *wh*-clause of the SPCs in (100) and the QAPs in (101):

(107) a. John said that he would take some pictures, and [taken some pictures], he has done

b. John said that he would take some pictures, and [take some pictures], he has done

A movement approach to VP topicalisation would confront us with what looks like an insurmountable problem: how to account for the non-occurrence of participial morphology on the clause-initial VP in the b-examples in (106) and (107)? These examples thus suggest that VP topicalisation involves base-generation, the initial VP never actually being in the complement of the participle at any point in the derivation, and linked instead to a *pro*-VP which in (106) is null and in (107) gets morphologically realised as *done*. The connection between the VP-topic and the *pro*-VP proxy is apparently allowed to be lax — while the pro-VP must be participial, the VP-topic does not have to agree with it in features.[20]

(108) [take(n) some pictures], he has {*pro*-VP/done}

This said, the examples of SPCs and QAPs with participial connectedness are straightforwardly compatible with a non-ellipsis approach to these examples; it is actually the b-examples which are of greater interest, since it is these which, when coupled with the VP topicalisation cases, present an argument against the syntactic reconstruction approaches to SPCs mentioned in the introduction (Heycock & Kroch 1996, Bošković 1997) — for these approaches the lack of participial morphology on the counterweight in the example in (100b) seems very difficult to accommodate.

4.1.7 On the breadth of the generalisation

Of the arguments against a free relative analysis of the *wh*-constituent of specificational pseudoclefts, and in favour of a link between SPCs and QAPs, that we reviewed in the preceding subsections, four involve irreversible and unembeddable SPCs (cf. (109)-(110), below, and also (91), and (96), above) — hence the topicalisation, multiplicity, pied-piping and case connectedness arguments strictly speaking apply to our Type A SPCs only.

(109) a. *this wine is [to Mary, what he will/won't give]
b. *I'd call [to Mary, what he will/won't give] this/any wine

(110) a. *met Peter was met wie hij gesproken had [Dutch]
b. *mit Peter war mit wem er gesprochen hatte [German]
with Peter was with whom he spoken had

Though our conclusion that the *wh*-clause of SPCs is not a free relative thus holds firmly in the case of Type A constructions, one might wonder whether a free relative approach to Type B SPCs could still be viable. As a matter of fact, for languages like Bulgarian, Greek and Hebrew, a case has been made in the literature that the *wh*-constituent of their SPCs really is a free relative.

Evidence to this effect comes, among other things, from the morphological make-up of the *wh*-forms opening the *wh*-constituent of SPCs in these languages. In Bulgarian and Greek, the *wh*-pronoun heading a free relative obligatorily has a definite determiner affixed to it (e.g. Bulgarian *koj-to* and Greek *o-pjos* 'who+DEF'), while the *wh*-pronoun of questions does not; it is the former, affixed form of the *wh*-pronoun that shows up in Bulgarian and Greek SPCs:

(111) kakvo-*to* kaza bese ce Maria e umna [Bulgarian; Izvorski (1997)]
what-DEF said was that Maria is smart
'what he said was that Maria is smart'

Notice that SPCs of the sort in (111) can be of Type B. In fact, it seems that in the languages for which evidence has been presented to the effect that the *wh*-clause of the SPC is a free relative,

[20] Notice that the *done* found in the examples in (107) is an instance of main-verb *do*; as such it introduces questions of a type similar to the ones we posed in fn. 7, above. It will be useful, in future research, to consider these questions in tandem. We have not conducted a detailed investigation at this time.

Type B is the *only* type of SPC. As Anastasia Giannakidou (p.c.) has pointed out to us, Greek SPCs exhibit a cluster of properties which conspires to the conclusion that Type A SPCs are absent from this language: Greek SPCs (i) do not license any NPIs in the 'counterweight'; (ii) do not allow 'undeleted', full-IP counterweights (only *na*-clauses being possible, these being larger than IP, possibly as large as CP); and (iii) feature *wh*-clauses which unambiguously qualify as free relatives, Far from defeating our conclusion that in a well-defined class of cases the *wh*-clause of SPCs exhibits the behaviour of a *wh*-interrogative rather than that of a free relative, this particular clustering of properties peculiar to Greek SPCs (and presumably shared with their counterparts in Bulgarian and Hebrew) confirms once again the distinction that should be made between SPCs of Type A and Type B.

In particular, while the links with the *wh*-clauses of Question-Answer pairs that we have drawn attention to hold for unequivocal Type A SPCs, an analysis of the *wh*-clause as a free relative seems plausible for the case of Type B SPCs, not just in the light of the Bulgarian, Greek and Hebrew facts but also in view of the analysis of Type B SPCs that we were led to at the end of section 2 of this paper, repeated below.

(42) [SC [Subject counterweight] [Predicate *wh*-clause]] [Type B SPCs]

In this small clause structure, the *wh*-constituent functions as the predicate of the counterweight. Now, while free relatives make perfect predicates, *wh*-questions typically do not. There is reason to believe, therefore, that Type A and Type B SPCs differ with respect to the way they analyse the *wh*-clause: as a *wh*-interrogative in the former case, and as a free relative in the latter.[21] [22]

---

[21] A potentially serious objection to our claim that the *wh*-clause of Type A SPCs is a *wh*-interrogative, not a relative clause construction of sorts, comes from the fact that many SPCs with *what* can be paraphrased with constructions featuring headed relatives. And what is particularly interesting is that these headed relative constructions can even preserve the NPI connectivity effects of their *wh*-counterparts. To see this, consider the examples in (i), which are both fine in elliptical as well as 'undeleted' form:

(i) a. what John didn't do was [~~he didn't~~ buy any linguistics book]
b. the (one/only) thing John didn't do was [~~he didn't~~ buy any linguistics book]

One thing is clear: headed relatives do not serve as questions; they are unanswerable. Nonetheless, they do show up in 'echoes' introducing an answer to a question like (iia), as seen in (iib):

(ii) a. A: what is the (one/only) thing John didn't do?
b. B: the (one/only) thing John didn't do? — buy any linguistics books

What this suggests is that, even though these relatives do not look like questions, they can be underlyingly represented as (non-root) *wh*-interrogatives à la *what the (one/only) thing that John didn't do is*. There is some evidence to suggest that relatives like these can indeed be embedded in an elliptical *wh*-question of this sort: the emergence of spurious *is* in examples like (iii) (discussed in January 1992 on *Linguist List*; constructions like these are found in regions throughout the United States, but to British English speakers they sound very awkward; we have found them attested in spontaneous speech in Dutch as well):

(iii) a. the reason <u>is is</u> we have no handle on this construction
b. all it is saying <u>is is</u> that you are being paid out of the grant [attested sentence]
c. the (one/only) thing John does to linguistics books <u>is is</u> burn them

Of the two tokens of *is* in (iii) one is apparently spurious; but this is merely apparent if we look upon the first of the two *is*-es as the lexicalisation of the copular head of the elliptical *wh*-interrogative which *reason/all/thing* is the predicate. (Note that, except in echoes like (iib), headed relatives do not suffice as root *wh*-questions; thus, eliding the copula *is* seems possible only when it does not find itself in Comp.)

Apart from the NPI effects, constructions of the type in (ib) also share with Type A SPCs their peculiar restrictions on auxiliaries in the root clause:

(iv) a. the (one/only) thing John may have claimed was that she had given you the book
b. *the (one/only) thing John claimed may have been that she had given you the book
(v) a. what John may have claimed was that she had given you the book
b. *what John claimed may have been that she had given you the book

In both these regards, relatives with *contentful* heads (i.e. heads other than dummy's like *thing* and *all*) exhibit parallel behaviour as well (cf. Higgins 1973:343 on the auxiliary cases):

(vi) the only claim that John didn't make was that she had given you any books
(vii) a. the only claim that John may have made was that she had given you the book
b. *the only claim that John made may have been that she had given you the book

4.2 *The link with Question-Answer pairs: Parallelism effects*

Our arguments to the effect that the *wh*-constituent of a Type A SPC has the syntax of a question establish a connection between SPCs and QAPs, as we pointed out earlier on. Not only can this link be profitably exploited in the discussion of ellipsis and the conditions thereon, as we showed in section 3, it also allows us to make sense of some of the most mysterious quirks of SPCs, all involving some sort of PARALLELISM EFFECT between the *wh*-clause and the counterweight/answer. In this section we will compound a variety of facts underscoring the role of parallelism.

4.2.1 *Multiple wh-constructions*

Several popular approaches to SPCs 'recreate' a simple sentence LF structure out of the surface SPC, via a variety of operations that we need not be concerned with in this context (see e.g. Heycock & Kroch 1996 and Bošković 1997, and references cited there). For all accounts of SPCs that assume that their LF is that of a simple clause, a question that arises is why multiple embedded questions such as (112), which are well formed as simple, non-cleft clauses, do not have a specificational pseudocleft counterpart (the deviance of (113) is absolute, i.e. it cannot be attributed to a wh-island violation incurred by movement of *what* as in *??what did John wonder who read?* and *??what John wondered who read was the Bible*):

(112) John wondered who read a book by which linguist
(113) *what John wondered who read was [a book by which linguist]

The deviance of (113), while mysterious from the perspective of 'simple clause LF' approaches to SPCs, is readily expected on the approach to SCPs taken here: it is of a kind with the ungrammaticality of the full-IP SPC in (114) and the infelicitousness of the QAPs in (115):

(114) *what John wondered who read was
[he wondered who read a book by which linguist]
(115) a. what did John wonder who read? — *a book by which linguist
b. what did John wonder who read? — *he wondered who read a book by which linguist

---

Our suggestion that headed relatives can occur in elliptical (non-root) *wh*-interrogatives and serve as the *wh*-clause of SPCs in this fashion readily carries over to the examples in (vi)-(vii). Constructions such as the ones in (ib), (iv) and (vi)-(vii) are seriously problematic for an approach to SPCs à la Bošković (1997), according to which the counterweight is moved into the *wh*-clause at LF to replace the 'surface anaphor' *what* in examples like (va) — since in (ib), (iv) and (vi)-(vii) there seem to be no 'surface anaphors' involved, it is unclear what the counterweight could possibly replace at LF in these examples.

Sharvit (1997) rejects a parallel treatment of 'run-of-the-mill' SPCs and cases like (ib), on account of the observation that the 'pseudo-pseudocleft' in (viiib) is ungrammatical:

(viii) a. what John didn't buy was any linguistics book
b. *the (one/only) thing that John didn't buy was any linguistics book

The fact that (ib) is grammatical suggests, however, that failure to license an NPI in the counterweight is not a systematic property of the 'pseudo-pseudocleft'. What appears to be wrong with the example in (viiib), in comparison with (ib), is that *the (one/only) thing* in the former two examples functions as an argument of the verb heading the relative clause, while in (ib) *thing* corresponds to a predicate (the VP *buy any linguistics book*) in the counterweight. We can try to rationalise this restriction on the distribution of *the (one/only) thing* type 'pseudo-pseudoclefts' as a parallelism effect: on the (admittedly *ad hoc*) assumption that *the (one/only) thing* in 'pseudo-pseudoclefts' can only be a predicate inside the relative clause, and that there must be matching across the copula with respect to predicativity (cf. section 4.2.8, below), the ungrammaticality of (viiib) may follow.

[22] The observation (see section 4.1.1) that English SPCs never allow the addition of *-ever*, which is a hallmark of free relatives, does not contradict the claim that the *wh*-clause of Type B SPCs is a free relative even in English, if it is assumed (as in Iatridou & Varlokosta 1995) that the addition of *-ever* makes it impossible for the free relative to function as a predicate — all and only those FRs that can be predicates can figure in Type B SPCs, given the analysis in (42).

The obvious reason why the examples in (114) do not work is that one cannot answer a *wh*-question with a *wh*-answer of this type. And since our approach to SPCs likens them to QAPs, it does not come as a surprise that the examples in (113) and (114) are parallel to those in (115).

The previous examples involved ungrammatical SPCs which, when 'reduced' to simple clauses at LF via some sort of (post)syntactic operation (à la Heycock & Kroch 1996 or Bošković 1997), would correspond to grammatical multiple *wh*-constructions — making a case against syntactic reconstruction approaches of these types. It is worth reminding the reader, at the end of this subsection, that the deviance of the examples discussed here cannot possibly be blamed on some general problem with multiple *wh*'s in SPCs. For recall that, as we noted in section 4.1.4, one does in fact come across multiple *wh*-SPCs.

4.2.2 *Special NPIs (I):* not ... until

A strong current in our approach to SPCs is the fact that Type A SPCs exhibit NPI connectivity. Section 2 presented a variety of cases in which an NPI in the counterweight is licensed by what seems to be a constituent of the *wh*-clause, something which turns out to be illusory on the account we have developed, in which the NPI is in effect licensed entirely within the counterweight, by a negation that is part of the elided material. But it turns out that not just *any* NPI in the counterweight can be so licensed. One conspicuous exception is formed by *not ... until*, as seen in (116a) (cf. Clifton 1969, Higgins 1979:45, Sternefeld 1997). While we do not profess to have an account of this surprising breakdown of NPI connectivity, we would like to draw attention to the fact that (116a) behaves just like the QAP in (116b) in this regard:

(116) a. *what John didn't do was leave until 6pm (cf. John didn't leave until 6pm)
b. what didn't John do? — *leave until 6pm

That we are really dealing with a property of QAPs, not with a restriction on NPI licensing of sorts, is shown by the fact that the examples in (116) do not improve at all when the elliptical counterweight/answer is replaced with its full-IP counterpart, as in (117):

(117) a. *what John didn't do was [he didn't leave until 6pm]
b. what didn't John do? — *he didn't leave until 6pm

4.2.3 *Special NPIs (II): negatively polar idiom chunks*

Similar observations can be made with regard to the licensing of negatively polar idiom chunks in SPCs and QAPs. Consider the following examples:

(118) a. *what John didn't have was a red cent (cf. John didn't have a red cent)
b. what didn't John have? — *a red cent

Once again, the 'covertness' of the NPI licensor in these elliptical examples does not seem to matter, for 'undeleting' the counterweight/answer, as in (119), does not lead to a significant improvement:[23]

(119) a. *what John didn't have was [he didn't have a red cent]
b. what didn't John have? — *he didn't have a red cent

4.2.4 *Idiom chunks in general*

Idiom chunks lead to crashes even when there is no NPI licensing involved. That is, both SPCs and QAPs with idioms split between *wh*-clause and counterweight/answer are no good:

[23] The examples in (119), to the extent that they are acceptable at all, have the effect of a (linguistic) joke — the effect typically procured by a *zeugma*. At best, then, (119) might work as an *indirect answer* construction; see the discussion at the end of section 3 on the limited availability of 'indirect answer' counterweights in SPCs.

(120) a. what John took was a picture [not idiomatic]
b. what did John take? — a picture [not idiomatic]

While grammatical, the examples in (120) both lack an idiomatic interpretation for *take a picture* (i.e. 'photograph'). And even here, the full-IP alternatives fail to procure an idiomatic reading as well:

(121) a. what John took was [he took a picture] [not idiomatic]
b. what did John take? — he took a picture [not idiomatic]

It seems as though the *wh*-clause is a 'wrong-footer': the verb in the *wh*-clause is not interpretable idiomatically; a parallelism constraint on the *content* of the *wh*-clause and the counterweight/answer seems responsible for the unavailability of an idiomatic reading throughout.[24]

4.2.5 Multiple quantifiers and (lack of) scope ambiguities

Parallelism, or rather, the lack thereof, also gives us the key to an account of the fact (noted in Williams 1994:62) that, even though the simple sentence in (122) is scopally ambiguous, the corresponding (elliptical and 'undeleted') SPC in (123) does not allow the quantifier in the counterweight to scope over the QP in the *wh*-clause.

(122) every article that appeared bothered a friend of mine
(123) what bothered a friend of mine was [every article that appeared ~~bothered a friend of mine~~]

Here again, QAPs are similar to SPCs, as the lack of scope ambiguity for the example in (124) shows:

(124) what bothered a friend of yours? — [every article that appeared ~~bothered a friend of mine~~]

The cause of the unavailability of the wide-scope readings for the universal quantifiers in (123) and (124) is straightforward: since in the *wh*-clause there can be no scopal ambiguity, parallelism ensures that there cannot be any in the counterweight/answer either.

4.2.6 Negative Raising

Similar such parallelism constraints rear their heads in a closely related domain as well: Negative Raising. Consider the following ill-formed SPCs (cf. Higgins 1979; also discussed in Bošković 1997):

(125) John does not believe that she will graduate
a. = John does not hold the belief that she will graduate
b. = John holds the belief that she will not graduate
(126) what John does not believe is that she will graduate
a. = John does not hold the belief that she will graduate
b. ≠ John holds the belief that she will not graduate

While (125) is ambiguous between a reading in which negation scopes over the matrix clause and one in which it confines its scope to the embedded clause (the latter being a case of Neg Raising), in the SPC in (126) only the matrix reading for the negation is available. Notice that the SPC in (126) shares

[24] Here as in (119) the SPCs and QAPs might work as *zeugmas* — i.e., cases in which there is *form* identity of question and answer but no *content* (or LF) identity. The LF-identity constraint can be flouted (not just in QAPs and SPCs but in coordinations as well), with the result of the typical effects of *zeugma*.

this interpretive property with both the full-IP Type A SPC in (127) and the QAP in (128), thereby once again confirming the link between SPCs and QAPs.

(127) what John does not believe is [John does not believe that she will graduate]
a. = John does not hold the belief that she will graduate
b. ≠ John holds the belief that she will not graduate
(128) what does John not believe? — that she will graduate
a. = John does not hold the belief that she will graduate
b. ≠ John holds the belief that she will not graduate

We can understand the unavailability of the Neg Raising reading in (126)-(128) as a reflex of a parallelism requirement on the *wh*-clause and the counterweight — since in the *wh*-clause of (126)-(128) the negation has no choice but to take matrix scope, it must take matrix scope in the counterweight/answer as well.

4.2.7 Argument structure

Parallelism also gives us a handle on the ungrammaticality of SPCs of the type in (129), discussed in Bošković (1997):

(129) a. *what John gave was Mary a book
b. *what John gave was a book to Mary

Bošković proposes to analyse the ungrammaticality of these examples with an appeal to a breakdown of connectivity for the head trace present in the bracketed parts of the examples — on a Larsonian approach to the structure of triadic constructions (Larson 1988), these bracketed constituents would be VPs with a V-trace as their head. We believe that an analysis of (129) along such lines is not worth pursuing, in view of the fact that even the full-IP counterparts of (129), given in (130), as well as the QAPs corresponding to (129), shown in (131), are ungrammatical.

(130) a. *what John gave was [John gave Mary a book]
b. *what John gave was [John gave a book to Mary]
(131) a. what did John give? — *Mary a book
b. what did John give? — *a book to Mary

The problem with all these examples seems to be lack of parallelism between the *wh*-clause and the counterweight/answer: the counterweight/answer 'harks back to' a *wh*-clause with triadic *give*, but the *wh*-clause features dyadic *give* (the beneficiary being unrealised). Notice that the counterweight in (129b) and the answer in (131b) are actually fine in multiple *wh* SPCs/QAPs (see 4.1.4 on these): *what John gave to whom was a book to Mary* and *what did John give to whom? — a book to Mary*. This further supports our perspective, and gives an additional indication that Bošković is not on the right track when it comes to explaining the ungrammaticality of (129).

4.2.8 Predicativity

There is evidence that the *wh*-clause and the counterweight/answer in SPCs and QAPs must match with respect to the nature of predicates as well. To see this, consider the following examples. In (132) we present a copular predication which is ambiguous in principle between a stage-level/existential interpretation (paraphrased in (132a)) and an individual-level/generic reading (given in (132b)):

(132) a fireman is available

a. = there is a fireman available
b. = a fireman has the intrinsic (IL) property of being available

Interestingly, now, the SPC counterpart of (132), presented in (133), is unambiguous — only the b-reading survives:

(133) what a fireman is is available
a. ≠ there is a fireman available
b. = a fireman has the intrinsic (IL) property of being available

That this is not a peculiar property of SPCs only is shown by the fact that QAPs behave the same way. The question in (134) can only be answered by mentioning some intrinsic (IL) property which the replier considers a fireman to possess. We believe that this is due to the fact that the question itself features an IL predication: *what*, the question word, is a *nominal* predicate (*contra* Jacobson 1995, Sharvit 1997 and others, who claim that the *what* of SPCs is a cross-categorial element); nominal predicates are IL predicates (cf. Kratzer 1989). The answer must parallel the question with respect to the IL/SL distinction; if it does not, the result is infelicitous. And what is true for QAPs is true for SPCs as well.

(134) what is a fireman?

4.2.9 Particle placement

Now that we have touched upon the behaviour of predicates in SPCs and QAPs, it is but a small step to considering the placement of one particular type of secondary predicate, the verbal particle (cf. Den Dikken 1995a and references cited there for detailed discussion of its status as a predicate), in relation to the verb and the object.

As the examples in (135) and (136) show, there is a strong tendency for a particle to be placed on the same side of the object in the *wh*-clause and the counterweight; whenever particle placement in the two major constituents of an SPC varies, a degradation results.[25]

(135) a. who looked the words up was [John looked the words up]
b. ??who looked the words up was [John looked up the words]
c. ??who looked up the words was [John looked the words up]
(136) a. ?where he looked the words up was [he looked the words up in Webster's]
b. *where he looked the words up was [he looked up the words in Webster's]
c. *where he looked up the words was [he looked the words up in Webster's]

As before, the facts of Type A SPCs find a parallel in the domain of QAPs:

(137) a. who looked the words up? — [John looked the words up]
b. ??who looked the words up? — [John looked up thewords]
c. ??who looked up thewords? — [John looked the words up]
(138) a. where did he look the words up? — [he looked the words up in Webster's]
b. ??where did he look the words up? — [he looked up the words in Webster's]
c. ??where did he look up thewords? — [he looked the words up in Webster's]

By now the reader will be able to guess what is at issue here — the facts in (135)-(138) show that the *wh*-clause and the counterweight/answer have to be parallel. this time with respect to the question of

[25] (136a) is not perfect to begin with; speakers of English are usually not very comfortable with SPCs whose *wh*-clause features a *wh*-word other than *what* or *who*, for reasons which are immaterial (i.e., it does not differentiate in any direct way between the various approaches to SPCs taken in the literature; in particular, it does not favour an analysis of the *wh*-constituent of SPCs as a free relative, since free relatives with *where*, *how* and *why* are fine: *where he lives is a nice place, how/why he did it is a mystery to me*).

whether Verb-Particle Reanalysis (abstract incorporation; cf. Den Dikken 1995a) takes place or not: applying reanalysis in the *wh*-clause (and thereby deriving an 'inner particle' construction with V-Prt-Object order) is legitimate iff it is also applied in the counterweight/answer.

This said, now consider the following data, taken from Kayne (1998:26):

(139) a. what he looked up was [he looked up a linguistics term]
b. ??what he looked up was [he looked a linguistics term up]
(140) a. what is he looking up? — [he's looking up a linguistics term]
b. ?what is he looking up? — [he's looking a linguistics term up]

Once again we find a link between the Type A SPCs in (139) and the QAPs in (140), which confirms the analytical connection we have drawn between the two. The fact that the particle cannot comfortably surface in the outer position in these examples suggests, on the analogy of what we found for (135)-(138), that Verb-Particle Reanalysis takes place in the *wh*-clause in these examples — and *must* take place here. This can be thought of as a reflex of the overt-syntactic extraction of the object in these sentences. On the assumption that extraction of the object in a V-Prt construction is contingent on reanalysis of verb and particle, the facts in (139) and (140) are entirely on a par with those found in the c-examples in (135)-(138).

Verb-particle reanalysis thus presents yet another case in which *wh*-clause and counterweight/answer in SPCs and QAPs have to respect parallelism.[26] One last case of parallelism that we want to address in this section — even though it presents no further evidence for the link between SPCs and QAPs — concerns the distribution of so-called emotive *should.*

4.2.10 Emotive should *connectivity*

Higgins (1979), and Heycock & Kroch (1996) and Bošković (1997) in his wake, notice that the distribution of emotive *should* — a form of *should* licensed by certain 'emotive' predicates in the upstairs clause — shows what looks at first like a surprising pattern in SPCs. To set the stage for the SPC examples, consider first the simple case of emotive *should* in (141a):

(141) a. it is unfortunate that he should be proud of himself
b. it is apparent/clear that he {is/*should be} proud of himself

Here the presence of *should* in the embedded clause is licensed by the adjective *unfortunate* in the matrix. In the absence of a licensor of the *unfortunate* type, no *should* can be included in the embedded clause, as (141b) shows.

Now notice the contrast between the following two sentences:

(142) a. it is unfortunate that [what he should be] is proud of himself
b. *it is unfortunate that [what he is] should be proud of himself

That (142b) is ungrammatical is not too much of a surprise — after all, as is well known, SPCs in general do not allow for the addition of modals to the root: while *what John is is proud of himself* is grammatical, **what John is can be proud of himself* is not. (We will turn to this property of SPCs in section 5 of the paper.) But what is interesting is that (142a) *is* acceptable. Hence *unfortunate* is apparently capable of licensing emotive *should* in a clause which it does not govern. That this really is

[26] The idea that (139b) and (140b) are awkward because V-Prt reanalysis obligatorily applies in the *wh*-question (while it has failed to take place in the counterweight/answer) because of the fact that the object is extracted is not entirely unproblematic. Den Dikken (1995a) argues that V-Prt reanalysis is incompatible with modification of the particle by elements like *right*; and he also shows that V-Prt reanalysis fails, on principled grounds, in complex particle constructions like *they made John out a liar.* Yet extraction of the object in a complex verb-particle construction succeeds (*who did they make out a liar?*) and premodification of the particle in simple V-Prt constructions is not blocked when the object is extracted (*what did they look right up?*). An alternative rationalisation of the drive for inner particle placement in (139) and (140), in terms of *focus*, is offered in Kayne (1998:26).

an extraordinary property of SPCs is shown by the fact that emotive *should* cannot be licensed in a finite CP functioning as the subject of a clause embedded under *unfortunate*:

(143) it is unfortunate that [that he {is/*should be} proud of himself] caused a scandal

Even though subject sentences seem to behave like the *wh*-clause of SPCs in some other respects (see fn. 32, below), they do not follow the pattern of (143) at all.

For the 'syntactic reconstructionists' (Heycock & Kroch, Bošković), the grammaticality of (142a) of course constitutes no surprise. For them the embedded clause behaves like the simple clause in (141a) at LF; (141a) and (142a) are fully parallel. But our approach fares no worse in this domain; for on our analysis, too, there is the possibility of analysing the counterweight of the embedded SPC in (142a) as a full IP. Since, as will be evident in section 5, this IP is the root of the SPC, the adjective *unfortunate* can license emotive *should* in the counterweight IP. Parallelism between the modal content of the counterweight and that of the *wh*-clause then ensures that, whenever *unfortunate* does indeed select emotive *should* in the elliptical counterweight IP, *should* will also surface in the *wh*-CP.

Contrary to the Heycock & Kroch (1996) and Bošković (1997) approaches to the representation of SPCs, our analysis of the facts in (142) makes the prediction that the licensing of emotive *should* should fail when the linear order of *wh*-clause and counterweight is reversed — after all, SPCs with XP<*wh* order cannot be Type A SPCs; and since our account of the facts in (142) relies on a Type A analysis of the SPC in (142a), reversal of the major constituents of the embedded clause in this example should lead to ungrammaticality. This prediction is borne out:

(144) it is unfortunate that proud of himself is [what he {is/*should be}]

The fact that (144), while grammatical with *is* in the *wh*-clause, crashes with emotive *should* vindicates our analysis of the connectivity effects involving emotive *should* in terms of Type A SPCs and parallelism between the *wh*-clause and the (elliptical) counterweight.

4.3 Lack of parallelism in syntax: Expletives

While the foregoing discussion has carried the point home that the lack of parallelism between a *wh*-clause (or *thing*-type headed relative) and the counterweight/answer can be held responsible for the ungrammaticality of a variety of putative SPCs and QAPs, it turns out, on the other hand, that the drive for parallelism sometimes gets relaxed in particular ways.

For instance, consider again the example in (23) from section 2, repeated here as (145):

(145) ?what didn't happen was [an accident of any kind] (= (23), above)

We pointed out in section 2 that the fact that this SPC is acceptable, while surprising at first blush in the light of the ungrammaticality of **an accident of any kind didn't happen*, can be explained by assuming that (130) is the elliptical counterpart of the full-IP SPC in (146):

(146) ?what didn't happen was [~~there didn't happen~~ an accident of any kind] (= (25a))

This analysis gives us a handle on (145), but it does of course introduce a case of non-parallelism between the *wh*-clause and the counterweight of the SPC. But this does not turn out to be an SPC-specific quirk. For notice that the behaviour of QAPs is entirely on a par with that of SPCs, something which once again emphasises the connection between SPCs and QAPs.

(147) what didn't happen? — ?there didn't happen an accident of any kind

Mismatches in 'expletivity' are found with both expletives of English — not just with *there* but with *it* as well. We have already come across an example of this type: the example in (33a) from section 2, whose structure, on the Type A analysis, looks as in (43a), repeated here as (148):

(148) what didn't happen next was [~~it didn't happen next~~ [that anybody fell]] (= (43a))

Once again, QAPs exhibit a similar behaviour:

(149) ?what didn't happen next? — it didn't happen next that anybody fell

The theory of matching or parallelism for QAPs (and, by extension, SPCs (of Type A) as well) should be permissive enough to allow an expletive construction to serve as the answer/counterweight to a non-expletive construction in the *wh*-clause.

So what we have seen is that in the vast majority of cases we have reviewed, there is a strict parallelism constraint on the *wh*-clause and the counterweight/answer in SPCs and Question-Answer pairs. But in the case of expletive constructions, the drive for parallelism seems to be relaxed.[27]

Let us approach the problem from the perspective of the elliptical example in (148), and ask whether it is legitimate at all to delete an expletive *it* independently of its associate — something which, if the approach to (33a) that we developed in section 2 is to go through, must in fact be legitimate. The examples in (150), instantiating FWD in coordination constructions, show that this is indeed the case:[28]

(150) a. it is certain [that John is the culprit], and
~~it is~~ imperative [that he be caught]

b. it is assumed by some [that John is the culprit], and
~~it is assumed~~ by others [that Bill is]

c. to me, it is clear [that John is the culprit], and
to you, ~~it is clear~~ [that he is not]

d. by whom is it assumed [that John is the culprit], and
by whom ~~is it assumed~~ [that he is not]?

Hence, from the point of view of the deletion site itself (*it* being deleted alone, without its associate), the analysis of SPCs like (33a) as in (43a)=(148) is unproblematic. The same is true for similar examples, like the one in (151a), analysed as in (151b):

---

[27] Another interesting case of lack of parallelism in SPCs, of a rather different sort, is the following. Higgins (1979:85) notes that the following is a grammatical SPC of English:

(i) what John does that we disapprove of is shave himself with a copper strip

What is interesting about this example is that it correlates with a simple clause counterpart in which what is a restrictive relative in (i) comes out as a non-restrictive relative:

(ii) John shaves himself with a copper strip, which we disapprove of

For approaches to SPCs along the lines of Heycock & Kroch (1996) and Bošković (1997), the pair in (i)-(ii) presents a difficult challenge: how to turn a restrictive relative into a non-restrictive one? From the perspective of our analysis of SPCs, on the other hand, the problem posed by (i) is similar, in a way, to the one we came across in the last paragraph of fn. 7 of section 3.2, addressing cases like *what John did to the book was burn it*. In both, what we see is that the *wh*-interrogative contains more information than the counterweight. We predicated the discussion in fn. 7 on the premise that all material present in the *wh*-clause gets represented in the counterweight. Suppose, though, that information contained in the *wh*-clause can be left uncopied. Then no problems are incurred by the *do+to* cases, the *to*-PP being represented in the *wh*-clause but not in the counter-weight; and (i), above, will be unproblematic as well. Thus, a relaxation of the semantic identity of *wh*-clause and counterweight will allow us to steer clear of any problems with regard to (i) and the *do+to* cases of fn. 7.

[28] Notice that, when we transform the second conjunct of (150a) into a SPC, as in (i), we discover yet another connectivity effect: subjunctive connectivity.

(i) what is imperative is [~~it is imperative~~ [that he *be* caught]]

Subjunctive connectivity behaves just like NPI connectivity in that it distinguishes between (30a) and (30b) — that is, while (i) is fine, (iia) is not (instead, (iib) must be used whenever no complementiser is realised); similarly, while (33a) is grammatical, (38a) is out.

(ii) a. *what is imperative is [he *be* caught]
b. what is imperative is [he *must be* caught]

For discussion of the distinction between (30a) and (30b) with respect to connectivity effects, see section 2.3.2.

(151) a. what is not possible is [that any students were there]
b. what is not possible is [~~it is not possible~~ [that any students were there]]

Examples of this type parallel those in (150b-d) with respect to the target of ellipsis; they pose no special problems in this domain, therefore.

Things are different, as already pointed out above, when it comes to the parallelism constraint to which ellipsis is subject (see section 3). While in the coordination examples in (150) the elliptical expletive in the second conjunct has a parallel, identical antecedent in the first conjunct, there does not seem to be an antecedent for the elided expletives in the structures in (148) and (151b). The question is, then, how the deleted expletive is licensed with respect to *recoverability* conditions. The structurally parallel antecedent for *it* in the *wh*-clause is *what* (or its trace). But deleted *it* cannot literally be identical with *what*, or its variable — it is not interpreted as a question word or as a variable. So to claim that *it* takes *what* as its antecedent would require tampering in an undesirable way with the notion of content identity.

Also, if *it* is parallel with the *wh*-phrase in (148) and (151b), then it should belong to the focus of the answer. But the pronoun *it* alone (where it clearly is an argument expression) can never form a focus; its strong pronoun counterpart *that* must be chosen, ellipsis or no ellipsis (cf. (152)). The pronoun *that*, on the other hand, cannot associate with an extraposed CP.

(152) a. *what is not possible is *it* (is not possible)
b. what is not possible is *that* (is not possible)

The solution for the parallelism problem will ultimately depend on the analysis of *it* in constructions with CP-associates. Is *it* in (148) and (151b) genuinely expletive (the theta-role going to the CP; this is the mainstream position, see Chomsky 1995), or does *it* count as an argument or predicate (entering some kind of cataphoric or predication relationship with the CP; cf. Bennis 1986 on *it* as an argument, and Moro 1997 on *it* as a predicate of the associate CP)? If the former, we might reasonably assume that *it* is not visible at LF in any case, thus satisfying conditions on FWD vacuously. If the latter, something extra needs to be said in connection with (148)/(151b) and the parallelism constraints on ellipsis.

From our present perspective, then, the optimal (or at least, the least cumbersome) approach is to treat *it* in (148) and (151b) as a genuine expletive, invisible at LF, hence not in need of an antecedent. The case of *there* (cf. (146)) will then be parallel, again on the assumption (taken in most of the literature on *there* sentences; but see Hoekstra & Mulder 1990, Moro 1997, Den Dikken 1995c for a different view) that *there* is an expletive. There is no antecedent for *there* in (146) any more than there is for *it* in (148)/(151b).

Concretely, then, we assume (by way of the easiest — but not necessarily the only — way out of the problem posed by (146), (148) and (151b)) that *it* and *there* are semantically vacuous placeholders, invisible from the LF representations of the constructions in which they occur. As such, they will assist the licensing of NPIs in examples like (145) and (151a), since at the point in the derivation at which the NPIs in these sentences are licensed (S-structure), the expletives are still present (i.e., no expletive replacement has yet taken place); but they do not cause any trouble with respect to parallelism, since by the time questions of content identity come into play (LF), the expletives are no longer there.

One might ask, since the expletives in (146), (148) and (151b) never receive a phonetic matrix at PF and do not survive at LF either, why they are present at all. Part of the answer is, of course, that, if they were not present, the NPIs would fail to be licensed (since we would not be dealing with post-verbal subjects in that case, but with preverbal ones instead). The other part of the answer is that these expletives must be present in SpecIP in the pre-Spell-Out representation of the elliptical clauses in (146), (148) and (151b) as a reflex of a purely formal requirement imposed by the I-head of the clauses that they appear in — the Extended Projection Principle (i.e., there is a strong D-feature on I which must be checked, and in the absence of movement of the logical subject to SpecIP expletives fill this position to meet the EPP).

To accommodate the non-realisation of the expletives in these elliptical cases, some mechanism must be assumed which suspends the requirement of PF-realisation, which is otherwise inviolable. The problem that then comes up is how to prevent such a mechanism from (wrongly) applying in non-elliptical IPs — for otherwise, given that it is not constrained by any requirement for an LF-antecedent, the mechanism allowing expletives to disappear in PF could in principle apply anywhere.

We have come across no case in which an overt finite Infl has a covert subject (expletive or not) in the elliptical IPs under discussion (QAPs and SPCs alike). It may suffice to assume, then, that PF-realisation of an expletive is forced only if its Infl is PF-overt. Then, FWD of finite Infl means that the PF-realisation of its EPP-checker is not necessary. No recoverability issue arises, as we have shown; so deletion of the expletive can then be viewed as being forced by the principle of maximal ellipsis discussed in section 3.3.

5 The analysis of Type A specificational pseudoclefts

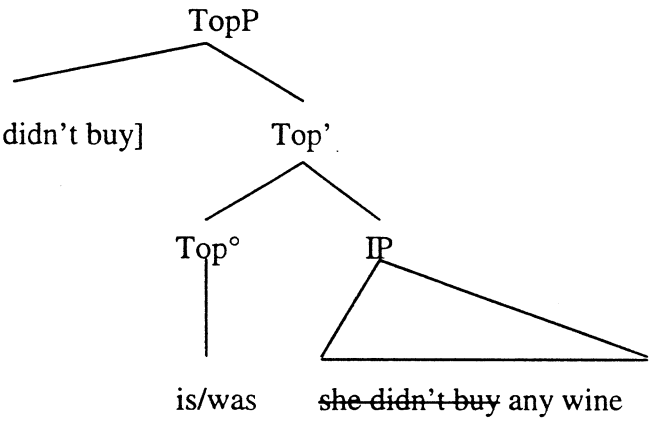
After this extensive review of the most important properties of specificational pseudoclefts of Type A, the (optionally) elliptical type, let us proceed to presenting an analysis of these SPCs which allows us to make sense of these and other characteristics of these constructions.

5.1 Topic-Comment

As our starting point we take the observation (cf. Hankamer 1974, Drubig 1996) that SPCs are 'topic-comment' constructions, the *wh*-clause being the topic and the counterweight the comment. We would like to structurally represent this observation by proposing the following analysis of Type A SPCs:

(153) (i) $[_{TopP}$ [*wh*-clause] $[_{Top°}$ is/was/were] [IP = counterweight]]

(ii)



The top structure of a Type A SPC is a TopP whose specifier position is filled by a base-inserted topic, the *wh*-clause; the head of TopP is filled by a form of the copula, and the complement of Top is the counterweight, a full IP (to which ellipsis optionally applies).

In the following sections we comment on the various ingredients of this structure of Type A SPCs, and discuss the predictions that this structure makes.

5.2 The topic

The *wh*-clause of a Type A SPC is an interrogative CP base-inserted into the specifier of the TopP. In this position, the *wh*-clause is neither a root question nor an embedded one. This gives the *wh*-clause of Type A SPCs a unique mixture of properties. Since the *wh*-clause is not a root question, it displays no Aux-to-Comp. But since it is not an argument of any predicate or a predicate of any argument

either, it does not behave strictly like non-root *wh*-questions with respect to topic placement. As we pointed out in section 4.1.2, topics can be placed in a position to the right of the *wh*-phrase in root questions but not in embedded questions, in English:

(154) ?to Mary, what should he give?
(155) ?I don't know [what to Mary, he should give]

And we also noted, in support of our *wh*-interrogative approach to the *wh*-clause of Type A SPCs, that these pseudoclefts allow topics to hop across the *wh*-phrase, as shown in (156).

(156) ?to Mary, what he will never give is any books

With regard to the placement of the topic in (156), the *wh*-clause of the SPC seems to behave like a root question. But interestingly, it seems that — in contrast to all other *wh*-constructions of English — the *wh*-clause of an SPC allows the topic to dock on either side of the *wh*-phrase in SpecCP; that is, (157) is acceptable to the same degree that (156) is:
(157) ?what to Mary, he will never give is any books

The *wh*-clause of SPCs can exhibit this hybrid behaviour thanks to is unique position in the tree: being base-generated in topic position, it is not theta-related to any constituent of the sentence, so adjunction to the maximal CP dominating the *wh*-clause is legitimate (cf. Chomsky 1986, 1995 on the restrictions on adjunction in connection with theta-theory; and see Motapanyane 1994 for evidence from Romanian showing that adjunction to CP is possible when CP is not an argument but barred when it is argumental); and since it is not a root question, featuring no Subject-Aux inversion, adjunction of the topic *inside* the *wh*-clause is also an option (on the assumption that the ungrammaticality of **what did to Mary, he give?* is due to the interference of the topic with Aux-to-Comp raising; cf. Kayne 1984:Chapter 9 on French inversion).

5.3 The comment

The IP counterweight in the complement of Top is the comment of the structure. It is a root IP; the TopP dominating it is an extension of its projection. And since the IP counterweight is a root clause, it can be coordinated with another root IP, in sentences like (158):

(158) what Bill is is [$_{IP}$ Bill is overbearing] and [$_{IP}$ Sue is timid]

Just as in regular cases of IP coordination, the second conjunct is subject to optional *gapping* affecting the copula *is* following *Sue*. The result of gapping applied to the second IP of (158) is the example in (159) (where '___' marks the position of the gapped copula), which is likewise grammatical:

(159) what Bill is is [$_{IP}$ Bill is overbearing] and [$_{IP}$ Sue ___ timid]

Finally, the IP in the first conjunct can undergo optional *ellipsis*, as in all cases of Type A SPCs. The result of ellipsis applied to the first IP in (159) is (160):

(160) what Bill is is [$_{IP}$ ~~Bill is~~ overbearing] and [$_{IP}$ Sue ___ timid]

Heycock & Kroch (1996:34) present examples of this type, which, though awkward, they judge to be grammatical. We concur with their judgement, and note that this type of construction is readily accommodated by the approach to SPCs taken in this paper.[29]

---

[29] Heycock & Kroch (1996) accommodate examples of this type in a not dissimilar way: they reduce the SPC in the first conjunct to a simple IP at LF and perform linear processing on the resultant structure; by the time the gapping clause (the second conjunct) is processed (i.e. after the first conjunct has been processed), we are

5.4 *The Top-head*

The examples in (158)-(160) are derived from a structure involving coordination of two IPs — the IP in the complement of Top is coordinated with another IP, and regular gapping and ellipsis operations can be performed on the result. Of course, we may also have coordination target the higher TopP node in the structure of the SPC *what Bill is is (Bill is) overbearing*, and construct a coordination of two SPCs of Type A. As expected, the output is grammatical:

(161) [$_{TopP}$ [what Bill is] is [$_{IP}$ Bill is overbearing]] and
[$_{TopP}$ [what Sue is] is [$_{IP}$ Sue is timid]]

Ellipsis straightforwardly applies to the IP counterweights of each of the two SPCs, reducing (161) to (162), which is, again, fully grammatical:

(162) [$_{TopP}$ [what Bill is] is [$_{IP}$ ~~Bill is~~ overbearing]] and
[$_{TopP}$ [what Sue is] is [$_{IP}$ ~~Sue is~~ timid]]
(i.e. *what Bill is is overbearing, and what Sue is is timid*)

But interestingly, (162) cannot be further reduced via gapping of the *is* between the *wh*-clause and the counterweight in the second conjunct; that is, (163) is ungrammatical, as noted in Higgins (1979:305) and Williams (1983:249) (cf. also Heycock & Kroch 1996:32):

(163) [$_{TopP}$ [what Bill is] is [$_{IP}$ ~~Bill is~~ overbearing]] and
*[$_{TopP}$ [what Sue is] ___ [$_{IP}$ ~~Sue is~~ timid]]
(i.e. **what Bill is is overbearing, and what Sue is, timid*)

This shows us something important about the nature of the copula in SPCs: while run-of-the-mill copulas are perfectly gappable (even in inverse copular sentences like *the best candidate is John and the runner-up, Bill*), the copula in SPCs *must* be realised. This suggests that the copula in SPCs is not a run-of-the-mill copula: the *is* mediating between the *wh*-clause and the counterweight in an SPC such as *what Bill is is overbearing* has a function different from that of other instances of *be*; it is not a support vehicle for tense/agreement morphology, nor is it a signal of Predicate Inversion (cf. Den Dikken 1995b), but it marks the presence of the TopP structure typical of SPCs.

The form of *be* intervening between the *wh*-clause and the counterweight of a Type A SPC, then, is the overt realisation of the head Top in the structure of these constructions. In a non-coordinate SPC, Top must always be overtly realised — i.e., **what Bill is* [*Bill is overbearing*] is ungrammatical. We can make sense of the obligatoriness of the lexical spell-out of Top once we realise that TopP is always the highest functional projection in its clause; hence its head can never be 'properly governed' by any higher head, so that leaving the Top-head empty contravenes the ECP (or some modern incarnation thereof; cf. Stowell 1981 on the ECP as a restriction on non-trace empty functional heads).

Since in a simple, non-coordinate SPC the Top-head can never be left unrealised, it now follows straightforwardly that in a case of coordination of two TopPs, each of the Top-heads *must* be spelled out overtly. The root of the matter, then, is that Top must be overt; as a consequence, gapping the Top-head *is* in the second conjunct of (163) leads to ungrammaticality. The significance of this discussion of gapping in SPCs with respect to the analysis of SPCs presented in this paper is that it vindicates the TopP approach to the structure of Type A SPCs by showing that the mediating copula is a spell-out of the head Top, rather than a run-of-the-mill copula.

Before closing this section, let us consider other options. Heycock & Kroch (1996) have a different account of (163), building on their LF reduction approach to SPCs: since by the time the gapping clause is processed, the antecedent is no longer parallel to it (since iota-reduction has

---

dealing with a straightforward case of coordination of two copular IPs with adjectival predicates: *Bill is overbearing and Sue is timid.*

'reshaped' the first conjunct into a simple IP), no gapping can apply to the relevant token of *is* in (163), for want of parallelism. There is evidence that distinguishes between Heycock & Kroch's approach to gapping in SPCs and ours, arguing in favour of the latter. Let us briefly review it here.

So far we confined our attention to SPCs featuring a *wh*<XP order. We know from the above that these can be of Type A. Reverse SPCs, with XP<*wh* order, by contrast, can never be of Type A, for reasons which we will turn to presently. This said, consider the behaviour of reverse SPCs in the domain of gapping, shown in (164)-(166). As the contrast between (165) and (163), and between (166) and (160) shows, reverse SPCs behave as the exact opposites of their *wh*<XP counterparts: the copula mediating between the XP and the *wh*-clause, while absolutely indelible in (163), is gappable in (165); but gapping fails in (166) while it succeeded in (160).

(164) overbearing is what Bill is, and timid is what Sue is

(165) (?)overbearing is what Bill is, and timid, what Sue is (cf. (163))

(166) *overbearing is what Bill is, and {timid, Sue/Sue, timid} (cf. (160))

From Heycock & Kroch's (1996) perspective on SPCs, the behaviour of the reverse cases in (165) and (166) with respect to gapping seems difficult to account for. They have iota-reduction 'transform' all SPCs, regardless of their surface word order, into simple clauses at LF; and as before, the feasibility of gapping in the second conjunct will depend on a parallelism restriction, the structure of the second conjunct being compared to that of the first, which by that time has been reduced to a simple IP in all cases of coordination of SPCs.

Concretely, then, by the time gapping in the second conjunct is processed in the examples in (165) and (166), what the processor is presented with on the Heycock & Kroch approach to SPCs are the following strings:

(167) Bill is overbearing, and [second conjunct]

a. ... [timid ___ what Sue is] (cf. **Bill is overbearing, and timid, what Sue is*)

b. ... [timid ___ Sue/Sue ___ timid] (cf. **Bill is overbearing, and timid, Sue*)

In neither of these cases is gapping expected to be permissible — correctly so in the case of (167b) (cf. the ungrammaticality of (166)), but contrary to fact in the case of (167a), given the well-formedness of (165).

While unexpected from Heycock & Kroch's (1996) perspective on SPCs, our analysis readily predicts the facts in (165) and (166). The thing to bear in mind is that these SPCs, surfacing as they do in a reverse, XP<*wh* order, can *only* be analysed as Type B pseudoclefts. Type B SPCs, as we pointed out in section 2.3.2, have a 'simple' small clausal structure, *not* a TopP structure. In fact, the TopP structure of Type A SPCs could never serve as the input to a derivation of the reverse, XP<*wh* order, for the simple reason that movement of the counterweight across the *wh*-clause in SpecTopP is out of the question — if such movement did take place, it would incur a violation of Relativised Minimality (Rizzi 1990) since the SpecTopP position would be crossed; and besides, there does not seem to be a landing-site available for XP-movement in any case. So the examples in (164)-(166) must be cases of Type B SPCs.

Type B SPCs differ in two crucial respects from their Type A pendants: (i) the counterweight of Type B SPCs is never an IP (for reasons discussed in section 2.3.2: IP is unacceptable as the subject of the small clause of which the *wh*-clause is the predicate); and (ii) the status of the copula mediating between the XP and the *wh*-clause in Type B SPCs is *not* that of spell-out of a Top-head, but rather that of a run-of-the-mill copula.

It is this latter difference between Type A and Type B SPCs that gives us the key to the account of (165). Since the copula sandwiched between *timid* and *what Sue is* in the second conjunct is a regular copula (cf. (ii)), and since we know that regular copulas are gappable, the grammaticality of (165) is as expected. The deviance of (166) is also straightforwardly accounted for. This particular output could only be got on the basis of the input in (164), *overbearing is what Bill is, and timid is*

what Sue is, by (a) gapping of the copula following *timid* (which is legitimate, as we just found out; cf. (165)) plus (b) ellipsis in the *wh*-clause, reducing *what Sue is* to plain *Sue*. But the latter will fail for obvious reasons, the conditions on ellipsis laid out in section 3 (in particular the need of an antecedent which makes the ellipsis site recoverable) being flouted. The analysis of SPCs laid out in this paper — in particular the distinction that it draws between Type A and Type B SPCs, reverse cases qualifying as unequivocal Type B specimens — thus captures all the gapping facts of SPCs reviewed in this section without further ado.

5.5 Restrictions on the root of specificational pseudoclefts

Let us take stock. We have proposed a structure for Type A SPCs involving a TopP, the specifier of which is occupied by the *wh*-clause (a base-inserted topic); the head Top is obligatorily filled by a form of the copula (which on this approach comes out as a kind of topic marker similar to Japanese *-wa*); Top takes as its complement the full-IP counterweight, which is subject to optional ellipsis. We have seen so far that this approach to Type A SPCs accounts in a straightforward fashion for the following batch of properties of these constructions:

(168) *properties of Type A SPCs explained so far*

a. NPI connectivity (strictly correlated with the distribution of 'undeleted' full-IP SPCs)
b. topic placement in the *wh*-clause
c. multiplicity
d. pied-piping effects
e. case connectedness
f. participial connectedness
g. restrictions on ellipsis (parallel to Forward Deletion)
h. parallelism effects
i. irreversibility
j. restrictions on gapping

In our account of (168j) we crucially appealed to the claim, emanating directly from our TopP approach to the structure of Type A SPCs, that the copula linking the *wh*-clause and the counterweight IP is an indelible lexicalisation of the Top-head. In what follows, we will start out by considering some other consequences of this claim, after which we will address a number of further properties of the root of Type A SPCs, which will be seen to fall out naturally from the account proposed.

5.5.1 Tense/Modality/Aspect/Modification/Inversion restrictions

One of the most 'celebrated' properties of SPCs is the fact that their root is subject to very tight restrictions on tense, modality, aspect, adverbial modification (including sentential negation) and Subject-Aux inversion, not shared by any other construction of English (or at least, not to the same extent).[30]

---

[30] The discussion in this section is based largely on Higgins' (1979) example material; also cf. Bošković (1997) for a recent treatment of the TMA restrictions on SPCs.

Note that English Locative Inversion (cf. (i)) has been claimed in the literature (cf. Aissen 1975; Coopmans 1989) to be subject to a subset of the contraints which govern SPCs — for instance, Aissen observes that negative (ib) is ungrammatical, and Coopmans claims that Locative Inversion in English does not allow auxiliaries (cf. (ic)). But as Breckenridge (1975) has pointed out, (id) is grammatical, which suggests that the ban on negation and auxiliaries does not hold categorically of Locative Inversion constructions.

(i) a. on this wall hung a picture of US Grant
b. *on this wall never hung a picture of US Grant
c. *on this wall can hang a picture of US Grant
d. on this wall has never hung a picture of US Grant

In any case, whatever the precise extent of the restrictions on Locative Inversion, they will always be just a subset of the constraints ruling SPCs with *wh*<XP orders.

The TMA, modification and inversion restrictions that Type A SPCs evince can be captured by the following generalisations:

(169) a. the copula agrees in tense with the *wh*-clause
b. there can be no modal auxiliaries present in the root of a *wh*<XP SPC
c. there can be no aspectual auxiliaries present in the root of a *wh*<XP SPC
d. there can be no adverbial modifier/negation present in the root of a *wh*<XP SPC
e. there can be no Subject-Aux inversion performed to the root of a *wh*<XP SPC

Illustrative examples of each of these five generalisations are given in the example pairs below, where (169a) is exemplified by (170), and so forth:

(170) a. what John is {*is*/**was*} angry with himself
b. what John was {*was*/**is*} angry with himself
(171) a. what John could be *is* angry with himself
b. *what John is *could be* angry with himself
(172) a. what John has been *is* angry with himself
b. *what John is *has been* angry with himself
(173) a. what John isn't is angry with himself
b. *what John is {is*n't*/is *probably*} angry with himself
(174) a. ?I wonder whether what John is is angry with himself
b. *is what John is angry with himself?

What we will do in this subsection is show how these restrictions are explained by the TopP approach to Type A SPCs taken in this paper.

5.5.1.1 Tense harmony as a reflex of Spec-Head agreement in TopP

The fact that the copula linking the *wh*-clause and the counterweight IP of a Type A SPC has to agree in tense marking with the tense of the *wh*-clause can be viewed, on the present analysis, as a reflex of the Spec-Head agreement relationship obtaining between the *wh*-clause in SpecTopP and the copula in Top. The copula in Top is unique in comparison with the various other tokens of the copula in that its projection is not embedded in a TP; hence the copula linking the two major constituents of a specificational pseudocleft cannot receive an independent value for tense. It is entirely dependent, when it comes to tense, on the value for tense borne by the T-head of the *wh*-clause in SpecTopP. Via the general feature-sharing process of Spec-Head agreement, the copula in Top receives whatever value for tense is present in the *wh*-clause.[31] This takes care of the generalisation in (169a), illustrated by the examples in (170).

5.5.1.2 There is only one head (Top) between wh*-clause and counterweight*

The generalisations in (169b) and (169c), exemplified by (171)-(172), also fall our readily from our analysis — quite simply from the fact that there is one and only one head position between the *wh*-clause and the counterweight IP in the TopP structure of Type A SPCs: the Top-head. General principles of X-bar structure (reducible to more fundamental principles à la Kayne 1994 or Chomsky 1995) thus ensure that only bare *is* or *was* may intervene between the two major constituents of SPCs of Type A.

5.5.1.3 Restrictions on adverb placement

The fact that there can be no sentential negation or any other adverbial modifier attached to the root of a Type A SPC follows from general restrictions in adverbial placement, we would like to claim. First,

[31] Note that, even though T is not the head of the *wh*-clause in SpecTopP (which bears the category label CP), the tense features of T are visible on CP as a result of LF movement of T to C.

consider the ban on negation. On the assumption (argued for in work by Pollock 1989, Zanuttini 1991, Laka 1990, Haegeman 1994 and many others) that negation heads a functional projection of its own, and that this NegP finds itself below the highest inflectional projection of the clause (AgrS or T; we have conflated these to Infl since nothing turns on the split-IP hypothesis here), the fact that there can be no negation between the copula in Top and the counterweight IP follows straightforwardly: no NegP can be built on top of the IP in the complement of Top.

The ban on adverbial modification of the root of SPCs other than negation is a little less straightforward, both empirically and analytically. While the ungrammaticality of (173b) with *probably* following the copula is a well-known and undisputed fact about English SPCs (cf. Higgins 1979), Bošković (1997:268, fn. 35) observes that its counterpart featuring the adverb to the left of the copula (??*what John is probably is angry with himself*) is 'slightly better'. Let us see how we can make sense of this.[32]

The ungrammaticality of **what John is is probably angry with himself* cannot be explained with a simple-minded appeal to a general ban on adverbial adjunction to IP, for we know that adverbs like *probably* do readily precede IP in other contexts. Since we are treating the IP counterweight of a Type A SPC as a regular root clause, the ungrammaticality of the above example thus seems to raise a serious problem. We can make it follow, however, if we can ensure that adverbial adjunction to the IP in the complement of *be*-filled Top is not allowed (or, alternatively, if each adverb comes with its own functional projection, as in Kayne 1994, Alexiadou 1997, and Cinque, in press, that generating an AdvP on top of the IP in the complement of *be*-filled Top is impossible). Whatever the precise explanation of such a restriction (we have no particular insights to offer at this time), notice that such a constraint, generalised in such a way that it forbids adverbial modification of an IP in the complement of a functional head filled by a finite auxiliary, seems operative outside SPCs as well — it is responsible for the ill-formedness of sentences like (175):

(175) *why is *probably* John angry with himself?

The structure of (175) is highly similar to that of **what John is is probably angry with himself*: in both, there is an adverbially modified IP immediately dominated by a functional projection (CP in (175) and TopP in the SPC case) whose head is filled by a finite auxiliary. In both cases, ungrammaticality is the result.

The fact that ??*what John is probably is angry with himself* seems to be slightly better than its counterpart with *probably* to the right of Top suggests that adverbial adjunction to Top′ is marginally possible, presumably on the analogy of *John probably is angry with himself*, which involves adverbial adjunction to T′. Though adverbial adjunction to an X′ constituent is possible in principle (in the phrase structure model of Chomsky 1995), adjoining an adverb to the projection of Top is an unlikely move from a semantic point of view: there is nothing in Top′ which the adverb could directly modify. This explains the marginality of the example.

5.5.1.4 No CP above TopP in root clauses

While the discussion of adverbial placement in Type A SPCs has involved a certain amount of hand-waving, we can end the discussion of the generalisations in (169) on a much brighter note. The ban on Subject-Aux inversion in Type A SPCs, codified in (169e) and illustrated in (174), follows given that, as is well known on independent grounds, no CP can be built on top of TopP in matrix clauses. That is, the ungrammaticality of (174b) is on a par with that of (176a,b), the latter of which we have had occasion to discuss at various points in the foregoing discussion:

---

[32] In addition to the text proposal, we can think of two alternative ways of capturing the improvement of the example provided by Bošković over the example in (173b): (i) the adverb in the former is inside the *wh*-clause, or (ii) the adverb is a parenthetical. The former would make the Bošković example parallel to ?**I wonder what John is probably t*; the latter assimiliates it with *John probably is angry with himself*. Neither seems to make exactly the right prediction regarding the status of the SPC case, the former making it worse than it actually is and the latter making it too good. We will set these two options aside at this time.

(176) a. *will to Mary, John give the book?
b. *what will to Mary, John give?

Since there can be no CP built on top of a root TopP, there is no head position above Top to which the copula in an SPC like *what John is is angry with himself* could possibly raise, to cross the *wh*-clause.

In embedded clauses, as we also know independently, from the grammaticality of (177), topicalisation *can* take place within (*wh*-interrogative) CPs:

(177) a. ?I wonder whether to Mary, John will give the book
b. ?I wonder what to Mary, John will give

The grammaticality of the SPC in (174a), *?I wonder whether what John is is angry with himself*, thus matches that of the examples in (177), as expected.[33]

Note that what was said above also immediately ensures that Type A SPCs are irreversible: an IP<*wh* order could be created on the basis of the TopP input structure, with the *wh*-clause in SpecTopP, only via raising of the counterweight and the head Top to the specifier and head positions of some functional projection on top of TopP. But we have just argued no such functional projection can exist. The irreversibility of Type A SPCs — one of our diagnostics for Type A status — thus follows.[34]

5.5.1.5 Where Type B is different

---

[33] Bošković (1997) gives an account of the generalisations in (169) whose central tenet is that SPCs with *wh*<XP order lack an IP; that is, the copula mediating between the *wh*-clause and the counterweight projects a bare VP of which the *wh*-clause and the counterweight are constituents, and no IP can be projected on top of this VP (the reason being that, if an IP *was* so projected, the *wh*-clause would have to raise to SpecIP for EPP reasons; but such raising fails given that, in general, the *wh*-clause of an SPC is immune to raising). While Bošković's hypothesis covers the facts of root clauses, we wonder how it would accommodate *embedded* SPCs with *wh*<XP order: if such SPCs radically lack an IP, how can they be embedded under a *finite* complementiser like *that* (which is generally taken to categorially select an IP complement)? Note that Bošković is a representative of the 'syntactic reconstructionists'; i.e., at LF he transforms the SPC into a simple clause by raising the counterweight into the *wh*-clause, obliterating *what*. At LF, then, there will be an IP in the complement of the finite complementiser in sentences like *I think that what John is is angry with himself*; but selectional restrictions are standardly held to wield their powers at the outset of the derivation, not by the very end of it.

With regard to Bošković's initial representation of SPCs with *wh*<XP order, note also that he explicitly assumes that the *wh*-clause and the XP entertain no relationship whatsoever; they just happen to be constituents of a bare VP projected by *be*. What Bošković does not address is how (and why) Merge would ever bring together two constituents that are claimed to have no relationship to each other in a *lexical* projection headed by a semantically empty copula. Put differently, though everything in Bošković's analysis of SPCs works reasonably well *from* the point at which the VP of *be* has been put together, the crucial question for Bošković to answer is how that initial point in the derivation ever comes into being. Note that our approach faces no such questions: while we agree with Bošković in claiming that the *wh*-clause and the counterweight IP in Type A SPCs entertain no thematic relationship, the subconstituents of TopP are not unrelated — one can think of the connection between the base-topic *wh*-clause and the counterweight as being similar to the relation between an *as for* topic and the clause following it (cf. *as for cars, John likes {expensive ones/Buicks and Chevrolets}*); see Meinunger (1997) for further discussion.

[34] All that was said in this subsection with respect to Type A SPCs seems to carry over to constructions featuring subject sentences:

(i) a. that John is angry with himself is unfortunate
b. *is that John is angry with himself unfortunate?
c. ??I wonder whether that John is angry with himself is unfortunate

This is not surprising, if Koster (1978) is right that subject sentences do not exist; in actual fact, what looks like a clause in subject position is a clause in topic position (SpecTopP), linked to a silent filler of SpecIP. (The example in (ic) seems somewhat worse than the corresponding pseudocleft in the main text, for reasons that are unclear to us.)

Now that we have an account of the generalisations in (169) from the perspective of a Type A approach to the SPCs in question, let us point out that unequivocal cases of Type B SPCs do *not* exibit most of these restrictions (cf. Bošković 1997):

(178) a. angry with himself {*is*/**was*} what John is
b. angry with himself {*was*/?*is*} what John was
(179) a. angry with himself *is* what John could be
b. angry with himself *could be* what John is
(180) a. angry with himself *is* what John has been
b. *angry with himself *has been* what John is
(181) a. angry with himself is what John isn't
b. angry with himself {is*n't*/is *probably*} what John is
(182) a. I wonder whether what angry with himself is what John is
b. is angry with himself what John is?

Where Type B is the same as Type A is in the domain of tense and aspectuality restrictions (cf. (178) to (170) and (180) to (172)).[35] These are different from the various other restrictions on SPCs in that they involve a relationship of *feature agreement* between the *wh*-clause and the copula of the main clause — a relationship which is established both in Type A SPCs (via Spec-Head agreement in TopP) and in Type B SPCs (via complement selection: the copula takes the small clause of which the *wh*-clause is the predicate as its complement, and imposes selectional restrictions on this small clause predicate). Thus, in (178) the *wh*-clause has to agree in tense to the copula of the root. And (180b) crashes because the perfective root fails to match the imperfective *wh*-clause which it selects; interestingly, as soon as we perfectivise the *wh*-clause as well, we find an improvement (though it is tough to find a context in which (180b') would actually be meaningfully used).

(180b') ?angry with himself *has been* what John *has been*

The selectional restrictions between the root clause and the tense and aspectual features of the *wh*-clause are imposed by L-related heads of the clause immediately dominating the *wh*-clause (which, recall, is the predicate of the small clause in the complement of *be*). Non-L-related heads do not entertain a relationship of selection with the predicate of the small clause; hence no agreement with respect to modality or negation (Mod and Neg both arguably being non-L-related or A'-heads) is imposed in the examples in (179) and (181). And even though we have extended the range of possible selectors of small clause predicates beyond the immediate governor (V) up to the maximal L-related extended projection of V, we do still keep selectional restrictions very local — in particular, we expect them to be clause-bound. That this is the right result is shown by the fact that the past tense of *seem* in the example in (183) does not cause the *wh*-clause to bear past tense morphology (unlike what we see in (178b), where there is a local selectional relationship between the matrix tense and the tense of the small clause predicate under the copula, i.e. that of the *wh*-clause).[36]

(183) angry with himself seem*s* to be what John *was*

So, as desired, all but two of the restrictions to which we have seen Type A SPCs to be subject evaporate for the Type B cases: Type B SPCs involve a regular small clause structure

---

[35] (178b) with *is* actually seems better than the reverse, while (178a) remains bad with *was*; we have no account for the surprising case of (178b), but to our knowledge no extant analysis of SPCs does.

[36] The reverse of (183) is ungrammatical regardless of the tense of the matrix and *wh*-interrogative clauses (cf. (i)), for reasons that will be discussed in the next section.

(i) a. *what John is seems to be angry with himself
b. *what John is seemed to be angry with himself
c. *what John was seemed to be angry with himself

embedded in a normal copular clause; there are no restrictions on modality, adverbial modification or Subject-Aux inversion, any more than there are in other copular sentences.[37]

5.5.2 *Why Type A SPCs are unembeddable under ECM and raising verbs*

Our TopP top structure of Type A SPCs provides an immediate explanation as well for the fact that these SPCs cannot be embedded under ECM and raising verbs. Recall from section 1 that the following examples of embedding an unequivocal Type A SPC under verbs like *call* and *describe as* are ungrammatical:

(184) a. *I'd call what John didn't buy any wine
b. *I'd describe what John didn't buy as any wine

The same is true for the examples in (185) and (186), involving attempts at embedding Type A SPCs in the infinitival complements of *believe* (ECM) and *seem* (raising), respectively:

(185) *I believe what John didn't buy to be any wine
(186) *what John didn't buy seems to be any wine

Of course accounting for the deviance of (184)-(186) on the TopP analysis of Type A SPCs is entirely straightforward. For we know independently, from the ungrammaticality of such sentences as (187a,b), that topicalisation in the complement of ECM verbs is out of the question:

(187) a. *I'd call [as for opera singers] Pavarotti the most successful of all time
b. *I believe [as for opera singers] Pavarotti to be the most successful of all time

And we also know that topics, base-generated as they are in A´-positions, do not undergo A-movement (which would involve improper movement). That is why there is no grammatical derivation for (186).

5.5.3 *Extraction restrictions*

One final property of SPCs left to be explained concerns the opacity effects that they incur. These are illustrated in (188)-(191):

(188) a. *what* do you think that John doesn't have [any pictures of *t*]?
b. **what* do you think that [what John doesn't have] is [any pictures of *t*]?
(189) a. [pictures of Berlin] I think that John doesn't have *t*
b. *[pictures of Berlin] I think that [what John doesn't have] is *t*
(190) a. who thinks that John has [which picture of Berlin]?
b. *who thinks that [what John has] is [which picture of Berlin]?
(191) a. I don't think that John has [any pictures of Berlin]
b. *I don't think that [what John has] is [any pictures of Berlin]

The deviance of the b-examples in the above sentence pairs leads us back to one of our central claims: the idea that SPCs are similar to QAPs, involving self-answering questions. What all the ungrammatical examples in (188)-(191) involve is a counterweight with an unbound variable in it — a trace of overt movement (in the first two cases) or LF extraction (in the latter two; we are adopting an LF-movement analysis of *wh-in-situ* and NPIs; cf. Moritz & Valois 1992, Den Dikken, to appear, and others for an LF-movement approach to NPI licensing). And answering a question with a constituent

[37] Bošković (1997) also assumes a regular IP structure for reverse (i.e. unequivocally Type B) SPCs; he hence makes largely the same predictions that we do, but note that since he does not assume there to be *any* structural relationship between the *wh*-clause and the counterweight in SPCs (regardless of their surface order), he fails to explain the persistence of tense and aspect restrictions in reverse Type B SPCs.

containing an unbound variable is no good — such constituents are not propositions, hence do not qualify as licit answers or counterweights.[38]

Notice that Bošković's (1997) analysis of SPCs also manages to account for the deviance of (188b) and (189b). For him these are ungrammatical as a result of the fact that their LF-derivation involves a violation of the constraint which prevents Move from applying to a trace or a constituent containing a trace (cf. Chomsky 1995). After all, Bošković šargues that at LF, *what* in the *wh*-clause of SPCs is replaced with the counterweight, via an LF instantiation of Move. Now, in (188b) and (189b) had Move already applied to the counterweight prior to Spell-Out — removing either a subpart of or the entire counterweight and raising it to the front of the root clause. The traces left by these overt applications of Move subsequently make it impossible for Move to target the counterweight again, at LF, to replace *what*. Since *what* has to be replaced by the counterweight, but since such replacement fails in the b-examples in (188) and (189), their ungrammaticality follows.

Elegant though this account may seem, it is fatally incomplete since it covers only the overt-movement cases in (188) and (189), and does not extend to the LF-movement examples in (190b) and (191b). In the latter two, no violation of any theoretical principle is expected to arise on Bošković's assumptions — after all, by LF the structure of an SPC is 'transformed' into a simple IP, basically parallel to the a-sentences in (190) and (191); and in a simple IP licensing a *wh-in-situ* or NPI should of course be entirely unproblematic, which (contrary to what Bošković leads us to expect) it is not.

5.5.4 Summary: The properties of Type A specificational pseudoclefts

With these results in mind, we can now go back to the list of properties of Type A SPCs that we had managed to get under control before we embarked on the discussion of the restrictions on the root of SPCs. That list was given in (168), repeated here:

(168) *properties of Type A SPCs explained up to section 5.5*

- a. NPI connectivity (strictly correlated with the distribution of 'undeleted' full-IP SPCs)
- b. topic placement in the *wh*-clause
- c. multiplicity
- d. pied-piping effects
- e. case connectedness
- f. participial connectedness
- g. restrictions on ellipsis (parallel to Forward Deletion)
- h. parallelism effects
- i. irreversibility
- j. restrictions on gapping

To this list we can now add the properties discussed in the preceding subsections:

(168) *properties of Type A SPCs explained in section 5.5*

- k. restrictions on the tense of the root clause
- l. the ban on modality in the root clause
- m. the ban on aspectual auxiliaries in the root clause

---

[38] There is another potential account of the facts in (188)-(191) — the account of SPCs in terms of Predicate Inversion presented in Heycock (1991), which assimilates the ungrammatical cases to their counterparts in copular inversion constructions of the type discussed extensively in Moro (1997) and Den Dikken (to appear).

(i)
- a. I think that the cause of the riot is a picture of the wall
- b. *which wall do you think that the cause of the riot is a picture of *t*?
- c. *which picture of the wall do you think that the cause of the riot is *t*?
- d. *I don't think that the cause of the riot is any picture of the wall
- e. *who thinks that the cause of the riot is which picture of the wall?

We find these parallels important, but will stick to the TopP approach to SPCs here since it allows us to explain, besides these extraction restrictions, the constraints in the domain of tense, modality, aspect, modification/negation, and Subject-Aux inversion as well — constraints which, as the reader may verify, are not mimicked by copular inversion constructions of the type in (ia) (see also fn. 28, above, on Locative Inversion; also cf. fn. 37).

n. the ban on adverbial modification (including negation) of the root clause
o. the ban on Subject-Aux inversion in the root clause
p. the ban on extraction of and from the counterweight, both before and after Spell-Out
q. the ban on embedding under ECM and raising verbs

To our knowledge, this basically exhausts the inventory of characteristics of the (English) specificational pseudocleft construction that can be pieced together from the literature on the subject.

5.5.5 A left-over

Let us return to the examples in (185) and (186). The literature on SPCs assigns examples of these types a star categorically, regardless of whether we are dealing with cases involving NPI connectivity (i.e., unequivocal Type A cases such as the ones in (185) and (186)) or with plain SPCs like those in (192) and (193). In our judgement, the examples in (192) and (193) do improve slightly in comparison with the examples in (185) and (186); but we concur with the standard view that the former are not impeccable.

(192) *I believe what John is to be proud
(193) *what John is seems to be proud

In this respect, the *call* and *describe as* cases seem to be genuinely different, as we already noted in section 1 — while (184a,b) are out, their plain, NPI-less cousins in (194) are fine:

(194) a. I'd call what John is proud
b. I'd describe what John is as proud

The persistence of ungrammaticality in (192)-(193) *versus* the contrast between (184) and (194) now raises a delicate question. We have an account for the ungrammaticality of the examples in (184)-(186) on a Type A approach, built on the TopP structure in (153). We can extend this account to (192) and (193) on the tacit assumption that these also feature a structure of Type A SPCs. But the problem that arises is that there seems to be no particular reason why these examples should necessarily involve a Type A structure — since these sentences feature neither an undeleted, full-IP counterweight nor an NPI inside the counterweight, they might just as well be given a Type B analysis, it seems; and since Type B SPCs involve a regular small clause structure embedded under a run-of-the-mill copula, there would then appear to be no reason why embedding under ECM and raising verbs should fail in (192) and (193). The question is even subtler than this; for the grammaticality of (194) seems to suggest that in some contexts a Type B analysis should indeed be assumed.

All in all, while our story for Type A SPCs is solid and complete, our insistence on a two-way split in the realm of SPCs now seems to cause us trouble in the account of (192)-(194). While the main body of our paper has zoomed in on Type A SPCs, we are now forced to consider the question:

- *what about Type B SPCs?*

Note that the context of (184)-(186) and (192)-(194) is not the only one that leads to this question; in the discussion throughout section 4 we have come across a number of cases in which what was said there with respect to unequivocal Type A SPCs carries over to *all* SPCs with the order *wh*<XP. Let us address this question in the closing section of the paper (which is still tentative and open-ended at this point).

6 What about Type B SPCs?

6.1 Are all SPCs with wh<*XP order of Type A?*

Life with respect to the examples in (184)-(186) and (192)-(193) (and a variety of other cases in which the behaviour of unequivocal cases of Type A SPCs is exemplary of the entire set of SPCs with an order in which the *wh*-clause precedes the counterweight) would be simple if *all* SPCs with *wh*<XP order were instances of our Type A — hence would involve a TopP structure of the type in (153). What would then be left for Type B are cases in which the counterweight precedes the *wh*-clause. Suppose we hypothesise this:

(195) a. all and only those SPCs with *wh*<XP order are of Type A
b. all and only those SPCs with XP<*wh* order are of Type B

The only empirical fact that would seem to stand in the way of (195) is constituted by examples of the type illustrated in (194). We seem to be dealing with *specificational* pseudoclefts here; they have a *wh*<XP order; but, contrary to unequivocal cases of Type A SPCs, they do allow embedding under *call* and *describe as*.

So if we take the route defined by (195) we have to analyse (194) in some other way — in particular, we then have to deny that the pseudoclefts embedded under *call* and *describe as* in these examples are *specificational* pseudoclefts. Two options then come to mind:

(196) a. an analysis of (194) in terms of *predicational* pseudoclefts (PPCs)
b. an analysis of (194) in some hitherto unidentified third way

What flies in the face of an approach along the lines of (196) is the fact that, even though they do not exhibit NPI connectivity, constructions of the type in (194) do show all other connectivity effects typical of SPCs (and atypical of PPCs; cf. (198)). In (197) we illustrate this with reference to BT-A connectivity (but the other connectedness effects in (1)-(3), above, can be reproduced here as well):

(197) a. I'd call what *John* is proud of *himself*
b. I'd describe what *John* is as proud of *himself*
(198) what *John* does is important to {**himself*/*him*} [PPC]

It seems, then, that the counterweight in (194)/(197) cannot be classified as a predicate of the *wh*-clause, at least not in the same way that *important to him* is a predicate of *what John does* in (198), the latter being entirely parallel to the relationship between *important to him* and *his work* in a sentence like *his work is important to him.*

Hence, if (194) is not to be an SPC, the only thing we are left with is the claim that *tertium datur* — (194) instantiates a third type of construction, different from both SPCs and PPCs. Though we will not pursue this possible hypothesis at any depth here, it does seem to us that it would have something to recommend it. What (196b) would presumably come down to is the idea that *proud (of himself)* in (194)/(197) is a label of sorts — a label assigned to the *wh*-clause, paraphrasable in terms of a noun phrase headed by nouns like *qualification* or *label*, as in the example in (199):

(199) I'd give what *John* is the qualification/label "proud of *himself*"

Here, *proud of himself* seems to behave neither like a predicate nor like an argument. Possibly, sentences of this type instantiate *equative constructions*, whose properties are far from clear (see Heycock & Kroch 1996 and references there for some discussion of the problems posed by equatives).

(200) possibly, pseudoclefts of the type in (194)/(197) are *equative pseudoclefts* (EPCs)

Whatever the outcome of the discussion, though, the least we can conclude with respect to sentences like (194) and (197) is that their status as specificational pseudoclefts deserves is less than crystal

clear, and that only if one plays the *tertium non datur* gambit do these constructions seriously undermine the pair of hypotheses in (195) — hypotheses which open up an interesting, restrictive perspective on the dichotomy between Type A and Type B SPCs.[39]

6.2 *Connectivity effects in Type B SPCs*

The discussion of (197), featuring BT-A connectivity, also leads us to the broader question of how to deal with connectivity effects other than the ones which we have accounted for with the aid of our ellipsis approach to SPCs of Type A — in particular, the effects listed in (1)-(3) in the introduction (involving the Binding Theory, opacity and bound variable anaphora). As we know, these effects, as opposed to the NPI connectivity case which served as our prime diagnostic throughout the paper, show up in SPCs in either order: *wh*<XP and XP<*wh* alike:

(7) a. [angry with {*himself*/*him*/**John*}] is what *he* is
b. [a unicorn] is what John seeks [*de dicto* or *de re*]
c. [a picture of *his* house] is what *nobody* bought
d. *[*any* wine] was what *nobody* bought

For the examples in (7a-c) we can be sure that, since undeleted full-IP counterweights are not possible in clause-initial position in SPCs (cf. (8), repeated below), we are not dealing with Type A SPCs; hence we cannot reduce the connectivity effects exhibited by these examples to straight c-command in an elliptical counterweight IP. So some other way has to be found to accommodate the connectivity effects in SPCs other than those involving NPI licensing (or, for that matter, case connectivity which we discussed in section 4.1.5).

(8) a. *[he bought some wine] was what John bought
b. *[he didn't buy any wine] is what John didn't buy

We may wonder which way to turn to find a suitable approach to SPCs that can give us the recalcitrant connectivity effects. In the introduction, we listed the three major types of approach to SPCs that the literature has brought forth: (i) the semantic approach, (ii) the syntactic reconstruction approach, and (iii) the ellipsis approach. Our analysis is a sophisticated specimen of the third variety; and we know that it will not deliver in the case of (7a-c). So we are left with (i), (ii) or some entirely novel approach. Of the two extant possibilities, the syntactic reconstruction approach seems out of the question. After all, we emphasised in the above that the connectivity effects exhibited by SPCs split

---

[39] They do of course raise the question as to *why* these linearity statements should hold. The fact that Type A SPCs can only feature a *wh*<XP order follows from the TopP analysis in (153); but the converse claim, inherent in (195), that Type B SPCs have a rigid XP<*wh* order does not follow from anything we have said so far. The structure we have proposed for Type B SPCs (following Heggie 1988, Heycock 1991) — given in (42) in the main text, repeated below as (i) — is a predication structure in which the *wh*-clause functions as the predicate of the counterweight XP.

(i) [$_{SC}$ [$_{Subject}$ countweight] [$_{Predicate}$ *wh*-clause]]

Deriving the XP<*wh* order from this base structure is of course entirely straightforward; the inverse, *wh*<XP order could only be derived from (i) by applying some leftward movement process to the predicate of the small clause — presumably an instance of Predicate Inversion. This said, we can ensure a rigid XP<*wh* order for Type B SPCs on the hypothesis that Predicate Inversion fails to apply to the predicate of the small clause in (i), because the predicate is of the wrong type, in a sense to be made precise. Notice that not just any predicate can undergo Predicate Inversion in English — thus, while (iia,b) are fine, (iic) is ungrammatical. The restrictions on the types of predicate that can be affected by Predicate Inversion are still largely mysterious; but the possibility certainly presents itself to group the free relative in (i) together with predicates like *a doctor* in (iic), and to thus rule out *wh*<XP orders for Type B SPCs as a reflex of general restrictions on Predicate Inversion. We leave this possibility, and the further questions that it raises, as a topic for future investigation.

(ii) a. the best doctor in town is John
b. the doctor is John
c. *a doctor is John

out into two separate classes — those involving NPI and case connectivity *versus* all the others. Now, precisely because the syntactic reconstruction analyses proposed by Heycock & Kroch (1996) and Bošković (1997) by their very nature generalise over *all* possible cases of connectedness (since they reshape the S-structure pseudo-cleft into a simple IP at LF), they have no obvious tools to make the desired split.

More promising would seem to be a particular development of a semantic approach to a subset of connectivity effects, along the lines of Sharvit (1997). Such an approach can, in principle, be tailored precisely to the needs of the empirical lie of the land — those connectivity effects that persist in all SPCs regardless of their word order and regardless of the syntactic environment in which the SPC shows up will come under the umbrella of *semantic* connectivity, while those whose distribution is exactly that of Type A SPCs are of a fundamentally *syntactic* nature (as we have shown at length in these pages). Ultimately, then, the facts of pseudoclefts will teach us precisely which connectedness effects belong to the syntax and which do not. It will then be up to semantic theory to appropriately delineate the class of semantic connectivity effects (as distinct from their syntactic cousins). Obviously, undertaking a project which will yield an appropriately explicit semantic theory of connectivity is well beyond the scope of the tasks we had set ourselves at the outset of this paper. We will therefore leave this issue as a topic for future research.

References

Aissen, J. 1975. Presentational-there insertion: a cyclic root transformation. Chicago Linguistic Society.

Alexiadou, A. 1997. *Adverb Placement. A Case Study in Antisymmetric Syntax.* Linguistik Aktuell 18. John Benjamins.

Bennis, H. 1986. *Gaps and dummies.* Dortrecht: Foris Publications.

Bošković, Ž. 1997. Pseudoclefts. *Studia Linguistica* 51 (3), 235-277.

Breckenridge, J. 1975 The Postcyclicity of *es*-insertion in German. *Chicago Linguistic Society.*

Büring D. 1995. *The 59th Street Bridge Accent.* PhD. University of Tübingen.

Cardinaletti, A. & M. Starke 1994. The typology of structural deficiency: on the three grammatical classes. *ZASPIL* 1, 1-55.

Cinque, G. (in press) *Adverbs and Functional Heads.* New York: Oxford University Press.

Chomsky, N. 1986. *Barriers.* MIT Press, Cambridge, Massachusetts.

Chomsky, N. 1995. *The Minimalist Program.* Cambridge, Mass.: MIT Press.

Clifton, E. 1969. The English pseudo-cleft. Ms.

Coopmanns, P. 1998.

Dayal, V. 1997. Free Relatives and –ever: Free Choice and Identity Readings, paper delivered at SALT VII.

Den Dikken, M. 1995a. Extraposition as intersposition, and the syntax of English tag questions. Ms. Vrije Universiteit Amsterdam/HIL.

Den Dikken, M. 1995b. *Particles. On the syntax of verb-particle, triadic, and causative constructions.* New York: Oxford University Press.

Den Dikken, M. 1995c. Copulas. Ms. Vrije Universiteit Amsterdam/HIL.

Den Dikken, M. (to appear) Predicate Inversion in DP. In: Alexiadou, A. & C. Wilder (eds.) *Possessors, Predicates and Movement in the DP.*

Drubig, B. 1996. Fokusstruktur und Fokuskonstruktion im Englischen. (Sprachtheoretische Grundlagen für die Computerlinguistik: Arbeitspapiere des SFB.)

Emonds, J. 1976. *A Transformational Approach to English Syntax.* New York: Academic Press.

Haegeman, L. 1990. Non-overt subjects in diary contexts. In Mascaro and Nespor (eds.) *Grammar in Progress*, 167-79.

Haegeman, L. 1994. *The Syntax of Negation.* Cambridge: Cambridge University Press.

Halle, M. & A. Marantz 1993. 'Distributed morphology and the pieces of inflection'. In K. Hale & S. Keyser (eds.) *The view from building 20*, 111-176.

Hankamer, J. 1974. On the noncyclic nature of wh-clefting. *Proceedings of CLS 10*, 221-286, Chicago, IL: CLS.

Heggie, L. 1988. *The syntax of copular structures*. PhD. University of Southern California.
Heycock, C. 1991. *Layers of Predication: The Non-Lexical Syntax of Clauses*. PhD UPenn.
Heycock, C. 1994. The internal structure of small clauses: New evidence from inversion. In: *Proceedings of NELS* 25.
Heycock, C. & A. Kroch. 1994. Verb Movement and Coordination in a Dynamic Theory of Licensing. The Linguistic Review 11. 257-283.
Heycock, C. and A. Kroch 1996. Pseudo-cleft connectivity: implications for the LF interface. Ms. Univ. of Edinburgh, Pennsylvania.
Higgins, F. R. 1979. *The pseudocleft construction in English*. Garland, New York.
Höhle, T. 1991. On Reconstruction and Coordination. In H. Haider & K. Netter (eds), *Representation and Derivation in the Theory of Grammar*. Dordrecht, Kluwer. 139-197.
Hoekstra, T. & R. Mulder 1990. Unergatives as copular verbs; Locational and existential predication. *The Linguistic Review* 7, 1-79.
Iatridou, S. and S. Varlokosta 1995. Pseudoclefts Crosslinguistically. NELS Hand out.
Izvorski, R. 1997. On the Type of *Be* and the Nature of the *Wh*-Clause in Specificational Pseudoclefts, paper given at the ,'Workshop on the Syntax and Semantics of (Pseudo-) Clefts, ZAS.
Jacobson, Pauline. 1995. On the quantificational force of English free relatives, in E. Bach, E. Jelinek, A. Kratzer & B. Partee, eds., *Quantification in Natural Languages* [Vol. 2]. Dordrecht: Kluwer.
Kayne, R. 1984. *Connectedness and Binary Branching*. Dortrecht: Foris.
Kayne, R. 1994. *The Antisymmetry of Syntax*. LI Monograph 25. MIT Press.
Kayne, R. 1998. Overt vs. Covert Movement, Ms. New York University.
Koster, J. 1978. Why subject sentences don't exist. In S. J. Keyser (ed.) *Recent Transformational Studies in European Languages*. MIT Press, Cambridge.
Kratzer, A. 1989. Stage-Level and and Individual-Level Predicates. Ms. University of Amherst.
Laka, I. 1990. *Negation in syntax: on the nature of functional categories and projections*. PhD MIT.
Larson, R. 1988. On the Double Object Construction. In: *Linguistic Inquiry* 19, 235-289.
Meinunger, A. 1997. 'The Structure of Cleft and Pseudo-Cleft Sentences'. In: Blight, R. C. & M.J. Moosally (eds.) *The Syntax and Semantics of Predication*. Texas Linguistic Forum 38, 235-246.
Moritz, L. & D. Valois 1992. French sentential negation and LF pied-piping. *Proceedings of NELS 22*, 319-334.
Moro, A. 1997. *The raising of predicates. Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.
Motapanyane, V. 1994. An A-Position for Romanian Subjects. *Linguistic Inquiry* 25, 729-734.
Pollock, J. Y. 1989. Verb movement, Universal Grammar and the Structure of IP. *Linguistic Inquiry* 13, 297-13.
Pesetsky, D. 1989. Language-particular processes and the Earliness Principle. GLOW newsletter 22. 48-49.
Pesetsky, D. 1995. *Zero Syntax*. Cambridge , Mass.: MIT Press.
Quirk, R; Greenbaum, S. ; Leech, G & J. Sharvit 1985. *A Comprehensive Grammar of English Language*. Longman: London.
Rizzi, L. 1990. *Relativized Minimality*. LI Monographs, Cambridge, MA: MIT Press.
Ross, J. R, 1997. The Source of / The Remote Structure of ... Pseudo-cleft Sentences. Handout of Talk given at UPenn.
Sharvit, Y. 1997. A semantic approach to connectedness in specificational pseudoclefts. Ms. UPenn.
Sternefeld, W. 1997. The Semantics of Reconstruction and Connectivity. Arbeitsberichte des SFB 340, Universities Stuttgart and Tübingen.
Stowell, T. 1981 *The Origins of Phrase Structure*. PhD MIT.
Uribe-Echevarria, M. 1994. *Interface licensing conditions on negative polarity items: A theory of polarity and tense interactions*. PhD., UConn.
Weinert, R. 1995. Focussing constructions in spoken language. Clefts, Y-movement, thematization and deixis in English and German. *Linguistische Berichte 159*. 34-369.

Wilder, C. 1997. Some Properties of Ellipsis in Coordination. In A. Alexiadou & T. Hall (eds.) *Studies in Universal Grammar and Typological Variation*. Benjamins. 59-107.
Williams, E. 1983. Semantic vs. Syntactic Categories. *Linguistics and Philosophy* 6. 423-446.
Williams, E. 1994. *Thematic Structure in Syntax*. LI monograph 23. MIT press.
Williams, E. 1997. Blocking and Anaphora. *Linguistic Inquiry* 28, 577-628.
Zanuttini, R. 1991. *Syntactic properties of sentential negation: A comparative study of Romance languages*. Ph.D. UPenn.