

Dependency and Directionality

Marcel den Dikken

*Department of English Linguistics • SEAS • Eötvös Loránd University
Research Institute for Linguistics • Hungarian Academy of Sciences*

Chapter 4

A syntactic typology of long \bar{A} -dependencies

4 A syntactic typology of long \bar{A} -dependencies

4.1 Introduction

If an English speaker wants to know how many girls, in his or her interlocutor's opinion, are coming to the party, he or she has two biclausal surface forms at his or her disposal to ask a question: the ones in (1a) and (1b).

- (1) a. how many girls do you think (*that) *ec* are coming to the party?
b. of how many girls do you think that they are coming to the party?

The question in (1a) represents the familiar pattern of a long-distance filler–gap dependency: *who* is associated with a gap in the subordinate clause. In standard English, this gap is not allowed to be right next to a complementiser: no *that* can be used in (1a). This is something I will talk about in chapter 5. The alternative strategy, in (1b), avoids the building of a long-distance filler–gap dependency, and avails itself instead of a combination of prolepsis (*of* introduces the proleptic object *how many girls*) and resumption (*they* is a resumptive pronoun linked to the proleptic object *how many girls*). Salzmann (2006) calls this strategy ‘resumptive prolepsis’, a useful label that I will adopt here. In a resumptive prolepsis construction, the *wh*-constituent binds a gap in the matrix clause as a proleptic object of *think* (which, in English, can accommodate this object only in an *of*-PP); the subject position of the embedded clause is occupied by a resumptive pronoun. The two strategies are not entirely semantically equivalent: the *wh*-phrase in (1b) has a strong tendency to be interpreted specifically, while there is no specificity bias in (1a). But they both serve the native speaker well in enquiring about the number of girls that his or her interlocutor thinks are coming to the party.

The long-distance dependency exemplified by (1a) is not particularly marked in English. But in parts of the Germanic-speaking world, this strategy sometimes delivers rather unnatural results — especially for subject dependencies (such as the one under scrutiny in (1a)), and particularly in relative clauses. The following example will serve to illustrate this. On the occasion of his ascension to the throne in 2013, the new head of state of the Kingdom of the Netherlands (King Willem-Alexander) was offered a song (the so-called *Koningslied* ‘King’s Song’) written on the basis of text contributions sent in by Dutch citizens. One of the lines of the text was (2a) — a long-distance filler–gap dependency between the relative pronoun *die* and the subject position of the clause embedded under *wist* ‘knew’. Immediately upon the release of the *Koningslied*, a furore about it broke out, with much of the venom being directed specifically at the text line in (2a), which many disgruntled observers found to be ungrammatical. Those who cared to offer improvements typically volunteered (2b), a resumptive prolepsis construction, with the PP *waarvan*, dependent on *wist*, containing the relative pronoun, which is resumed in the subordinate clause by the subject pronoun *'ie*.

- (2) a. de dag die je wist dat zou komen is eindelijk hier (Dutch)
the day RELPRON you knew that would come is finally here
‘the day that you knew would come is finally here’
b. de dag waarvan je wist dat 'ie zou komen is eindelijk hier
the day where.of you knew that it would come is finally here
‘the day of which you knew that it would come is finally here’

The fact that the matrix verb is factive certainly contributes to the marked nature of (2a) — but with a bridge verb such as *dacht* ‘thought’ or *zei* ‘said’ substituted for factive *wist* ‘knew’, the ‘feel’ of the sentence does not change significantly: it remains awkward. To be sure, the grammar of Dutch allows long-distance dependencies of this type: (2a) is grammatical. But (2b) unquestionably feels better. Under certain circumstances, long-distance dependencies are dispreferred — more so with subjects than with non-subjects (*de dag die je wist dat hij nooit zou vergeten* ‘the day that you knew that he would never forget’ sounds much less jarring).

In Hungarian, too, long-distance filler–gap dependencies are dispreferred to alternative strategies, again more so with subjects than with non-subjects.¹ Consider first the empirical picture for long-distance dependencies involving the subject of a finite clause. It is laid out in (3), which presents variation along four separate but partly correlated parameters: (a) the placement of the *wh*-expression (upstairs, as in (3a–d), or downstairs, as in (3e)); (b) the case form of the *wh*-expression (nominative in (3a,e), delative in (3b), accusative in (3c,d)); (c) (in)definiteness agreement on the matrix verb (DEF agreement with the embedded clause in (3a,b), INDEF agreement with the *wh*-phrase in (3c–e)); and (d) number inflection on the embedded verb (singular only in (3a,e), optional plural inflection in (3b), and speaker variation with respect to downstairs number inflection in (3c,d)).

- (3) a. ^(?)*hány lány gondol-od, hogy jön(*-nek) a buliba?* (Hungarian)
 how.many girl(NOM) think-2SG.DEF that come-*3PL the party.to
- b. *hány lány-ról gondol-od, hogy jön(-nek) a buliba?*
 how.many girl-DEL think-2SG.DEF that come-3PL the party.to
- c. ^{✓(?)}*hány lány-t gondol-sz, hogy jön-nek a buliba?*
 how.many girl-ACC think-2SG.INDEF that come-3PL the party.to
- d. *hány lány-t gondol-sz, hogy jön a buliba?*
 how.many girl-ACC think-2SG.INDEF that come.3SG the party.to
- e. *mi-t gondol-sz, hogy hány lány jön(*-nek) a buliba?*
 what-ACC think-2SG.INDEF that how.many girl(NOM) come-*3PL the party.to
 all: ‘how many girls do you think are coming to the party?’

Of the five construction types in (3), (3a) is the closest Hungarian counterpart to English (1a): the extracted subject has (unmarked) nominative case and binds a gap in the lower clause, whose verb is in the third-person singular form, in conformity with the fact that *hány lány*, while notionally plural, is formally singular in Hungarian. The sentence in (3a) generally receives rather lukewarm reactions from Hungarian native speakers: though deemed to fall within the realm of the grammatical, for no Hungarian speaker is (3a) ever the preferred way of expressing what English (1a) conveys. The most popular way of rendering (1a) is with the aid of a so-called ‘*wh*-scope marking’ or ‘partial *wh*-movement’ construction, as illustrated in (3e). Here, *hány lány* is pronounced within the subordinate clause, its matrix scope being marked by the element *mit* ‘what.ACC’. This surface pattern does not involve a long-distance filler–gap dependency at all. Plural *jönnek* ‘come.3PL’ is strictly impossible in the lower clause when *hány lány* is nominative, whether it surfaces in the embedded clause, right next to the verb (as in (3e)), or in the embedded clause (as in (3a)). In

¹ For excellent work on the rich palette of options and their distribution across speakers, see Gervain (2003, 2005, 2009), on whose findings I will rely heavily in the empirical discussion of the Hungarian facts. See also Jánosi (2014).

between (3a) and (3e), we find a three other possibilities, with different case and agreement profiles. One of these corresponds directly to the resumptive prolepsis strategy illustrated for English and Dutch in (1b) and (2b), resp.: the Hungarian sentence in (3b) represents resumptive prolepsis, with *hány lány* in a delative PP belonging to *gondol*, binding a pronoun in the subject position of the embedded clause that optionally shows number concord with its notionally plural antecedent. The *wh*-phrase can also surface with accusative case (marked by the suffix *-t*), as in (3c,d). Plural inflection on the subordinate verb (*jönnék*), as in (3c), is possible here, though this is a matter of major variation: Gervain (2003, 2005) has found that speakers neatly split into two groups on (3c), with one group finding it perfect and the other very marginal. The counterpart to (3c) with singular *jön* in the lower clause, given in (3d), is broadly accepted throughout the community, though those speakers who like (3c) prefer it to (3d).

The palette of possibilities in (3) is much richer and variegated in colour than what we find in English or Dutch. The pattern is not specific to long subject dependencies: object dependencies also paint quite a rich picture in Hungarian (generally subject to less variability in the judgements):

- (4)
- | | | | | | |
|----|----------|-----------------|-----------------------------------|---------------|-------------|
| a. | ki-t | gondol-od, | hogy meghívott | a buliba? | (Hungarian) |
| | who-ACC | think-2SG.DEF | that PV.invited.3SG.INDEF | the party.to | |
| b. | ki-ről | gondol-od, | hogy meghívta | a buliba? | |
| | who-DEL | think-2SG.DEF | that PV.invited.3SG.DEF | the party.to | |
| c. | ki-t | gondol-sz, | hogy meghívta | a buliba? | |
| | who-ACC | think-2SG.INDEF | that PV.invited.3SG.DEF | the party.to | |
| d. | ki-t | gondol-sz, | hogy meghívott | a buliba? | |
| | who-ACC | think-2SG.INDEF | that PV.invited.3SG.INDEF | the party.to | |
| e. | mi-t | gondol-sz, | hogy ki-t hívott | meg a buliba? | |
| | what-ACC | think-2SG.INDEF | that who-ACC invited.3SG.INDEF PV | the party.to | |
- ‘who do you think that he invited to the party?’

The points of variation in (4) are located in (a) the placement of the *wh*-argument (upstairs, as in (4a–d), or downstairs, as in the scope marking construction in (4e)), (b) the filler’s case form (accusative or delative), (c) the the upstairs verb’s definiteness inflection, and, this time around, (d) the downstairs verb’s definiteness inflection.² Again we find that (4a)³ and (4e) behave alike when it comes to the form of the verb in the lower clause: it can only bear indefinite inflection. For (4b), the resumptive prolepsis construction, downstairs definite inflection is the only option — not surprisingly, given that the resumptive pronoun linked to the proleptic matrix object is definite (cf. English *of whom do you think that he invited her to the party?*). But in (4c,d), we see variation with respect to the inflectional form of the verb in the subordinate clause. The patterns for long object dependencies are thus very similar to those for long subject dependencies in Hungarian: in both cases, we get five possible surface outputs; and in both cases, there are four loci of variation.

2 Since number does not matter for object agreement in Hungarian, I have used simple *ki* ‘who’ in (4) rather than *hány lány* ‘how many girls’, in order not to make the examples needlessly complex.

3 The acceptability of (4a) seems generally to be rated higher than that of (3a), which is in line with the picture emerging for other languages: long \bar{A} -dependencies involving the subject give rise to complications (such as that *that*-trace effect in standard English, or the ‘allergic reaction’ to things like (2a) in Dutch).

A *wh*-question pattern that adult English and standard Hungarian lack but which is frequently found in other languages, including child English (Thornton 1990, 1995) and possibly also non-standard varieties of Hungarian,⁴ is so-called ‘*wh*-copying’, illustrated for German in (5).

- (5) a. wer glaubst du wer kommen wird? (German)
 who think you who come will
 ‘who do you think will come?’
- b. wen denkst du wen von den Studenten man einladen sollte?
 who think you who of the students one invite should
 ‘which of the students do you think one should invite?’
- c. *wessen Studenten denkst du wessen Studenten man einladen sollte?
 whose students think you whose students one invite should

Sentences of this type are usually analysed in the literature in a way that is entirely true to the label they are standardly given, in terms of the spell-out of multiple copies of the same *wh*-constituent. If this analysis were correct, *wh*-copying would supply a powerful argument in favour of successive-cyclic movement via SpecCP (assuming that the second instance of *wer* in (5) occupies the SpecCP position of the embedded clause). But the term ‘*wh*-copying’ is really a misnomer, in a way that dampens the prospects of a multiple copy spell-out approach. In constructions referred to by this name, it is possible for the two *wh*-expressions to be non-identical, with one being larger than the other; and when the *wh*-expression is complex, it is never possible for it to show up *in toto* in both clauses. The German examples in (5b) and (5c) (taken from Fanselow & Ćavar 2001, cited in Bruening 2006) illustrate this.

In addition to the patterns for long-distance \bar{A} -dependencies that we encountered in the preceding paragraphs, there may be reason to include yet another type, revealed by an original observation due to Stowell (1981). As we saw above, one way in which long-distance \bar{A} -dependencies involving the subject dependency stand out from other such dependencies is in their unusual sensitivity to the presence of a complementiser (recall English (1a) and Dutch speakers’ reactions to (2a)). Another way in which \bar{A} subject dependencies transgressing a finite clause boundary are special is in their sensitivity to the matrix environment. Stowell (1981:410–13) points out that with verbs of the *convince* class, long-distance *wh*-dependencies involving the object are unproblematic, but long subject dependencies fail categorically, regardless of whether *that* is pronounced or left out:

- (6) a. what did they convince/persuade/remind him that he should do *ec*?
 b. *who did they convince/persuade/remind him (that) *ec* should do it?

4 Balázs Surányi (p.c.) has pointed out to me that there are varieties of Hungarian where something resembling ‘*wh*-copying’ seems possible (see (ia,b)). There is reason to doubt, however, that these Hungarian cases involve German-style ‘*wh*-copying’. Surányi finds such cases marginally possible with *complex wh*’es in both clauses. I will return to the proper treatment of (i) in the discussion of *wh*-copying later in this chapter.

- (i) %kit mondtál hogy kit hívtál meg? (Hungarian)
 who.ACC said.2SG.INDEF that who.ACC invited.2SG.INDEF PV
 b. %kit mondtál hogy ki jött be?
 who.ACC said.2SG.INDEF that who(NOM) came in

I will discuss this pattern in detail in chapter 5, which is dedicated in its entirety to the special behaviour of subjects. My discussion of *convince*-class constructions in section 5.3.14 leads to the conclusion that, in this particular niche of the grammar, the subject ‘sets the standard’: the ungrammaticality of (6b) follows from an analysis according to which a direct filler–gap dependency across the CP under *convince*-type verbs is systematically impossible. The fact that (6a) is acceptable then indicates that, with objects but not with subjects, an alternative to a direct \bar{A} -dependency exists — one that cannot be collapsed with any of the types of \bar{A} -dependencies catalogued previously. This alternative strategy is what I will call ‘*wh*-control’ — involving, as its name suggests, a control relation between the *wh*-dependency in the matrix clause and the one in the subordinate clause.

All things considered, Universal Grammar recognises a rich typology of \bar{A} -dependencies, including the following members:

- (7) *a syntactic typology of \bar{A} -dependencies*
- a. direct filler–gap dependencies
 - b. scope marking
 - c. concordial scope marking (a.k.a. ‘*wh*-copying’)
 - d. prolepsis *cum* resumption
 - e. prolepsis *cum* topic drop
 - f. control

This chapter will be dedicated to the development of analyses for each of the members of this typology of long-distance \bar{A} -dependencies. Because of its richness, the pattern shown by the Hungarian paradigms in (3) and (4) offers a particularly useful starting point for an investigation of the syntax of long-distance \bar{A} -dependencies. Section 4.2, which constitutes the bulk of the chapter, will be focused on the Hungarian facts, and a discussion of its repercussions for the syntax of \bar{A} -dependencies in general. The so-called ‘*wh*-copying’ construction enters the stage in section 4.3, after which section 4.4 presents a preliminary investigation of *wh*-control. Section 4.5 presents a concise conclusion to the discussion of long-distance \bar{A} -dependencies in this chapter.

Conspicuously absent from the typology in (7) is what is commonly called ‘successive-cyclic *wh*-movement’ — a staple of the generative approach to long-distance \bar{A} -dependencies. Their recent discussion of Dinka, a Nilotic language whose \bar{A} -dependencies present an unusually rich set of morphosyntactic ingredients, leads Van Urk & Richards (2015) to assert that the facts of this language provide ‘particularly compelling evidence for the idea that long-distance dependencies involve intermediate movement steps through the edge of every verb phrase and every clause’ (p. 152). In section 4.6, I will examine the validity of this assertion, and show that it does not go through. Not having discovered any compelling evidence to the contrary (see also the critical survey of extant arguments for successive cyclicity in Den Dikken 2009, to appear), I conclude that Universal Grammar’s inventory of \bar{A} -dependencies does not include successive-cyclic movement.

Caveat lector: a great deal of the discussion in this chapter will concern itself with the minutiae of the various types of long-distance \bar{A} -dependencies, based on in-depth investigations of the morphosyntactic facts presented by individual languages, with particular emphasis on Hungarian. Readers satisfied with the presentation of the overall typology in (7) and not interested the nitty-gritty of these \bar{A} -dependencies may choose to move right along to chapter 5.

4.2 Hungarian in focus

Let us now launch into a detailed study of the picture painted by the Hungarian examples in (3) and (4), whose essence is summed up in the table below:⁵

<i>parameters of variation</i>		SUBJECT	OBJECT
CASE form of the <i>wh</i> -argument	NOM	(3a,e)	
	ACC	(3c,d)	(4a,c–e)
	DEL	(3b)	(4b)
PLACEMENT of the <i>wh</i> -argument	upstairs	(3a–d)	(4a–d)
	downstairs	(3e)	(4e)
VERBAL AGREEMENT upstairs	with CP (DEF)	(3a,b)	(4a,b)
	with <i>wh</i> (INDEF)	(3c–e)	(4c–e)
VERBAL AGREEMENT downstairs	SG	(3a,b,d,e)	
	PL	(3b,c)	
	<u>not</u> with <i>wh</i> (DEF)		(4a,d,e)
	with <i>wh</i> (INDEF)		(4b,c)

TABLE 1

In the following subsections, I will work my way through this table. My starting point will be the distribution of upstairs definite inflection, because this gives us the most natural way into the analysis.

4.2.1 Upstairs agreement with the complement clause (DEF)

4.2.1.1 The *a*-sentences: Direct long-distance filler–gap dependencies

In the *a*-examples in (3) and (4), repeated below, we are dealing with a case of a long-distance dependency between the *wh*-constituent in the matrix clause and a gap in the embedded clause.

- (3a) hány lány gondol-od, hogy jön a buliba?
 how.many girl(NOM) think-2SG.DEF that come.3SG the party.to

5 Note that under ‘verb inflection’, wherever it says ‘INDEF’ this really means ‘with the *wh*-phrase’: in all the examples in (3/4), the *wh*-phrase is indefinite; when it is a definite *which*-phrase all ‘INDEF’ in Table 1 flips to ‘DEF’.

I would like to point out here that, although my illustrations will continue to be confined to *wh*-fronting cases, the patterns laid out in Table 1 can be reproduced for non-*wh* focus fronting, which actually presents one additional point of variation: upstairs *person* agreement (*csak TÉGED akar-lak hogy gyere* ‘only you.ACC want-2(OB).1 SG(SU) that come’).

- (4a) ki-t gondol-od, hogy meghívott a buliba?
 who-ACC think-2SG.DEF that PV.invited.3SG.INDEF the party.to

In the long subject dependency in (3a), the *wh*-phrase is caseless on the surface (nominative case is not overtly marked in Hungarian, which is why ‘NOM’ is consistently put in parentheses), just as it is in (3e), where it actually occurs inside the subordinate clause; and in both (3a) and (3e) it also exhibits the kind of number agreement with the verb of the embedded clause that one expects on the basis of the fact that *hány lány*, despite its notional plurality, is formally singular in Hungarian. (3a) is uniformly considered less than perfect — all Hungarian speakers prefer (3e), the scope-marking construction, to (3a); some find (3a) highly marked, others less so.

This relative preference for the scope-marking construction also manifests itself in (4): speakers prefer (4e) to (4a), involving a long-distance object dependency. In (4a), the *wh*-phrase is morphologically marked for accusative case (with the suffix *-t*), but it does not control definiteness agreement in the matrix clause: the fact that *ki-t* is indefinite is reflected in the inflection on the verb in the embedded clause but not on *gondolod*, which is a form from the definite/objective conjugation. Everything suggests that in the a-examples in (3) and (4), the *wh*-expression itself binds a gap inside the embedded clause: these are genuine long-distance filler-gap dependencies.

That the a-sentences do indeed involve a dependency between the *wh*-expression upstairs and a gap in the downstairs clause is strongly confirmed by the emergence of familiar island patterns. Thus, note that (8a,b) are both ungrammatical: the embedded clauses in these examples are questions — a *wh*-question in the case of (8a), and a yes/no-question (marked in Hungarian by the particle *-e* attached to the finite verb of the clause) in (8b).⁶

- (8) a. *hány lány kérdezted, hogy mikor jött el?
 how.many girl(NOM) ask.PST.2SG.DEF that when came PV
 b. *hány lány kérdezted, hogy eljött-e?
 how.many girl(NOM) ask.PST.2SG.DEF that PV.came-QPRT

The analysis for the a-sentences thus must build a filler-gap dependency across the CP boundary of the subordinate clause. It can do so provided that the embedded CP is transparent for the formation of such a dependency. It is here that the matrix verb’s inflectional form plays a key role.

It has been a hallmark of generative syntactic analysis for decades now that clausal constituents are ordinarily opaque, impenetrable for the establishment of long-distance filler-gap dependencies, and that something special needs to happen in order for them to be ‘opened up’. As we saw in chapters 2 and 3 already, Rackowski & Richards (2005) argue (see also more recent work by Van Urk & Richards 2015) that an embedded CP is made transparent for the formation of filler-gap dependencies across its boundaries when it is in an Agree-relation with the verb of the matrix clause. In (3a) and (4a), that is exactly what we find: the definiteness inflection on the matrix verb, *gondolod*, is a reflex of the fact that it is in an Agree-relation with the subordinate finite CP, which is treated by the grammar of Hungarian as a definite object. We see this in (9).

⁶ I only illustrate this island effect for the case of long subject dependencies, (3a). For long object dependencies of the type in (4a), the facts are the same, *mutatis mutandis*, but for reasons of space I will not present them here.

- (9) gondol-od/*-sz, hogy húsz lány jön a buliba
 think-2SG.DEF/*INDEF that twenty girl come.3SG the party.to

Note that the Agree-relation for definiteness between the matrix verb and the embedded CP is not a sufficient condition for the establishment of a long-distance filler–gap dependency across the CP: in the examples in (8), the matrix verb *kérdezted* also unmistakably bears definite inflection; but this time around, the presence of a *wh*- or yes/no-question operator in the left periphery of the embedded clause makes it impossible for *hány lány* to link up with its trace in the embedded clause. This is the familiar ‘intervention effect’ capitalised on in Rizzi’s (1990) Relativised Minimality (recall chapter 3). So when a CP is not an *absolute* barrier because it is an Agree-goal, a filler–gap dependency can nonetheless be obstructed by an *intervention* effect. But importantly, when the CP is not an Agree-goal, and therefore an absolute barrier, a direct dependency between a *wh*-filler in the matrix clause and a gap in the embedded clause becomes impossible to establish: (3a’) is entirely unacceptable for all speakers of Hungarian.⁷

- (3a’) *hány lány gondol-sz, hogy jön a buliba?
 how.many girl(NOM) think-2SG.INDEF that come.3SG the party.to

We see, therefore, that even though a definiteness agreement relation between the matrix verb and the embedded CP is not a sufficient condition for the creation of a filler–gap dependency between a nominative *wh*-expression outside the CP and a gap inside the CP, such an agreement relation IS most certainly a necessary condition for a successful dependency of this sort. We can understand this if it is precisely the Agree-relation for definiteness between the matrix verb and the subordinate clause that renders the latter transparent in principle, as argued by Rackowski & Richards (2005).

We can now summarise the syntax of the constructions in (3a) and (4a) as in (10), where the ‘smiley face’ attached to the embedded CP indicates that, thanks to the Agree-relation that it is engaged in with the matrix verb, the CP is transparent, making a dependency between the *wh*-expression upstairs and the gap downstairs legitimate (provided that no intervention effect arises).

- (10) *wh* V+DEF [_{CP}☺ ... *t_{wh}* ...] [☺ marks transparency]

There is one thing that I would like to stress about (10): the fact that the subordinate CP contains just one gap associated to the *wh*-filler. In particular, there is no additional gap in the intermediate SpecCP position — because there is no need for one: the Agree-relation between the matrix verb and the embedded CP renders the latter transparent for the formation of a direct filler–gap dependency between the *wh*-filler in the matrix clause and its single gap in the lower clause. An additional gap would not make a difference at all: it would just complicate the syntactic representation unduly. It is relevant here to look back for a moment at (3a’), which differs minimally

⁷ For (4a), just substituting *gondolsz* for *gondolod* does not turn the output ungrammatical: (4d) is well-formed. I will argue below that the syntactic derivations of (4a) and (4d) are very different, with *ki-t* binding a gap in the matrix clause in (4d), on a par with the derivation of (3d). The case difference between (3a) and (3d) is effaced in (4a,d) due to the fact that the case assigned to the *wh*-expression is the same (viz., accusative) regardless of whether it is a matrix or embedded object.

from (3a) in not featuring definiteness agreement between the matrix verb and the finite CP in its complement. No number of intermediate traces would ever succeed in making (3a') grammatical: any intermediate trace in SpecCP would still be dominated by CP, an opaque category in (3a') due to the fact that it is not an Agree-goal. The opacity of CP in (3a') makes it entirely impossible for a filler in the matrix clause to link up to a gap inside this CP. So-called 'successive-cyclic movement' cannot eliminate absolute barriers.

In the theory that I am advocating in this work, the grammar builds structures from the top down — i.e., from the root of the tree to the branches and leaves. When we arrive, at a certain point in the structure-building process, at the point at which the embedded CP is merged, we know right away whether it is an Agree-goal or not: after all, the matrix verb has already been merged by that time, including its morphological feature composition. When the Hungarian matrix verb is adorned with a DEF feature, it can establish an Agree-relation with the finite CP once it is merged, and immediately upon establishing this Agree-relation, the subordinate CP is declared transparent for the creation of a filler–gap dependency across its boundaries. But when the matrix verb is adorned with an INDEF feature, the CP in its complement cannot serve as an Agree-goal for it. Since there is nothing else in the matrix clause with which the newly merged CP could engage in an Agree-relationship, the embedded clause is thus declared an absolute island as soon as its top node ('CP') is included in the structure constructed thus far. The opacity of CP entails that this clause cannot contain a gap that is supposed to be directly derivationally linked to a filler in the matrix clause: such a dependency is doomed as soon as it turns out that the CP is not an Agree-goal. The postulation of an intermediate gap 'on the edge of' the embedded CP is entirely pointless from this point of view: it could not help solve the problem. No amnesty can be given to the edge of a local domain: when a particular domain is an absolute barrier in the sense of (11) (repeated from chapter 3),⁸ no direct filler–gap dependency can be established across its boundaries.

- (11) *absolute opacity*
- a. domains that serve as goals in an Agree relationship with an asymmetrically c-commanding probe are transparent for the establishment of filler–gap dependencies across their boundaries
 - b. domains that do NOT serve as goals in an Agree relationship with an asymmetrically c-commanding probe are absolute islands for filler–gap dependencies

It is important to note that the opacity of the embedded CP in (3a') does not entail that literally no dependencies of any kind can be established across its borders at all. After all, not all dependencies in grammar are filler–gap dependencies; and it is only these dependencies whose distribution is reined in by the opacity of a node containing the gap. The examples in (12) (where (12a) is of type (3a) and (12b) of type (3d)) bring out the difference between filler–gap dependencies and referential dependencies when it comes to their sensitivity to absolute islands:

⁸ For our purposes in this chapter, Agree can be understood pre-theoretically, as 'agreement'. In the broader picture, it is important to understand Agree as 'Downward Agree' — the kind of agreement relation in which the probe c-commands the goal (as opposed to 'Upward Agree', a.k.a. Spec–Head agreement, in which the goal c-commands the probe). This matters in the context of an account of the Subject Condition: see chapters 2 and 5.

- (12) a. ki gondol-ja, hogy tetszik neki?
 who(NOM) think-3SG.DEF that please.3SG him/her.DAT
 b. ki-t gondol- \emptyset , hogy tetszik neki?
 who-ACC think-3SG.INDEF that please.3SG him/her.DAT

Only in the a-example (where *gondolja* bears definite inflection) is the matrix verb in an Agree-relation with the embedded clause. Yet in both examples the dative pronoun in the embedded clause can be referentially dependent on the matrix subject. The establishment of a referential dependency between *neki* and the (pro-dropped) subject of the matrix clause is thus entirely oblivious to the question of whether the embedded CP is an opaque domain for the establishment of a direct filler-gap dependency or not. The example in (13) makes the same point on the basis of bound-variable anaphora: it is perfectly possible for dative *neki* to serve as a variable pronoun bound by the universal quantifier *mindenki* in the matrix clause despite the fact that the embedded clause is not an Agree-goal.

- (13) mindenki ki-t gondol- \emptyset , hogy tetszik neki?
 everyone who-ACC think-3SG.INDEF that please.3SG him/her.DAT

This is as it should be, precisely because, in contradistinction to filler-gap dependencies, neither referential dependency nor bound-variable anaphora is computed as part of the on-line syntactic structure-building process. After all, the dative pronoun *neki* in the embedded clauses in (12) and (13) by no means has to be linked (whether via coreference or through binding) to the matrix-clause subject: whether it is or not depends on the semantic intentions of the discourse participants. Referential and binding dependencies are post-syntactic: these are discourse-sensitive in ways that filler-gap dependencies clearly are not. Filler-gap dependencies are computed in the course of the on-line syntactic structure-building process. Thus, the fact that absolute islands obstruct filler-gap dependencies and not referential dependencies or bound-variable anaphora does not need to be stipulated: it falls out from the nature of the dependencies in question.

4.2.1.2 The b-sentences: Resumptive prolepsis with an oblique matrix wh-expression

With the conclusion reached in the previous subsection in mind, we can now proceed to an examination of the syntax of the other member of the paradigms in (3) and (4) featuring definite inflection upstairs: the b-sentences, with oblique case on the *wh*-expression.

- (3b) hány lány-ról gondol-od, hogy jön(-nek) a buliba?
 how.many girl-DEL think.2SG.DEF that come-3PL the party.to
 (4b) ki-ről gondol-od, hogy meghívta a buliba?
 who-DEL think-2SG.DEF that PV.invited.3SG.DEF the party.to

These sentences are unmistakably resumptive prolepsis constructions, with relative *hány lányról* ‘of/about which girls’ in (3b) and *kiről* in (4b) as proleptic objects of the matrix verb, binding a pro-dropped pronoun in the subordinate clause, just as in English (1b):

- (1b) of how many girls do you think that they are coming to the party?

The hypothesis that the subject of the subordinate clause is a *pronoun* in (3b) helps us explain the otherwise incomprehensible fact that plural inflection on the lower verb is possible here: while *hány lány* is formally singular, the agreement form of the verb in the lower clause is third-person plural.

It is well known that pronouns do not have to match up their number specification with that of their antecedent. In English, we see this, for instance, in dependencies between formally singular quantificational noun phrases such as *everyone* and a bound-variable pronoun further downstream, as in (14a); or in referential dependencies involving a collective noun phrase such as *the management*, which can control singular agreement on the verb but still antecede a plural pronoun, as shown in (14b).

- (14) a. everyone thinks that they are smart
 b. the management is convinced that they deserve a raise

For (14a) it would be entirely impossible to have *everyone* in the matrix clause control plural agreement with the verb (**everyone think* is out of the question). And while in British English varieties it would be possible for *the management* to team up with *are* rather than *is* in (14b), the point that interests us here is that (14b) is perfectly good with *is* upstairs and *they* downstairs.

The lack of number-feature matching seen in (14) manifests itself also in resumptive prolepsis constructions in English: (15) is a case in point.

- (15) of which company's management do you think that they actually deserve a raise?

For Hungarian (3b), an assimilation to the English example in (15) is now straightforward. The analysis is schematised in (16):

- (16) wh_{DEL} V+DEF t_{wh} [$_{\text{CP}}$... *pro* ...]

The matrix relative PP contains a formally singular but notionally plural *wh*-phrase that binds a silent pronoun in the embedded clause. The relative originates in the matrix clause as a dependent of the verb; but since the relative phrase is a PP, not a DP, the matrix verb cannot establish an Agree-relation for definiteness with it: such agreement relations are the prerogative of structurally case-marked dependents of the verb. Since the verb cannot establish an agreement relation with the relative, it is thus free to Agree with the subordinate clause, as reflected in the DEF agreement form of *gondolod* in (3b).⁹

9 The Agree-relation between *gondolod* and its CP complement renders the CP transparent (as marked by the 'smiley face' on the CP-label). For wh_{DEL} , this is uninteresting: it is engaged in a filler-gap dependency with a trace within the matrix clause; the relationship between it and the subject of the lower clause is one of binding, insensitive to the intervention of opaque CP boundaries (recall (12b) and (13)). When we build into (16) a direct filler-gap dependency across the embedded CP, we would, all else equal, expect it to come out good. But (ia), where the additional *wh* is italicised, is entirely ungrammatical. This is because all else is *not* equal: the trace of the matrix relative *wh*-element sets up an intervention effect for the *kit*-dependency across it. We see this also in (ib), where the matrix dative is the intervener.

The resumptive *pro* bound by the relative proleptic object can calibrate its number specification either to the formal number of its antecedent (which delivers *jön* downstairs) or to its notional number (yielding plural *jönnek*).¹⁰ This does not conflict with the observation made by Gervain (2005:12) (see also Farkas 2006) that the pronoun in (17) *must* be plural. What seems to be behind the distribution of notional number concord with pronouns in Hungarian is the degree of distance between the pronoun and its antecedent. In (14) the two are in different sentences, while in (16) *pro* is in the same complex sentence as its antecedent — just as in (18), where notional plural marking on *pro* is likewise only optional (Katalin É. Kiss, p.c.).¹¹

- (17) két lány jött be a szobába; leültettem őket /*őt
 two girl came PV the room.into seated.1SG.DEF them him
 ‘two girls entered the room; I offered them a seat’
- (18) megkértem a hat fiút, hogy siessen/siessenek
 PV.ask.1SG the six boy.ACC that hurry.SUBJUNC.3SG/3PL
 ‘I asked the six boys to hurry’

So far we have looked only at cases of delative prolepsis *cum* resumption in which the silent pronoun serves as the subject of the embedded clause. But from (4b) we know that the delative prolepsis strategy is available also for long \bar{A} -dependencies involving an object gap. By the logic of the account of (3b), we are led to postulate a *pro* in the object position of *meghívta* ‘invited.3SG.DEF’ in the structure of (4b). Since *pro* is invariably definite, the fact that the verb in the embedded clause must bear definite inflection (*meghívta* vs. **meghívott*) now follows immediately. There can be no speaker variation on this point.

Less certain is what is expected with respect to the distribution of pro-drop in the case of delative prolepsis *cum* object resumption. Object pro-drop in Hungarian appears to be possible in the third person only for singulars: third-person plural objects generally cannot be pro-dropped.¹²

- (i) a. *hány lány-ról kit gondol-od t_{DEL} hogy imád-nak?
 how.many girl-DEL who.ACC think.2SG.DEF that adore-3PL.INDEF
- b. *{kinek/Marinak} ki mondtad t_{DAT} hogy a barátod volt?
 who.DAT/Mari.DAT who(NOM) said.2SG.DEF that your girlfriend was

10 Gervain (2005) notes the same variation, for a delative containing a non-*wh* proleptic phrase:

- (i) hat meghívott-ról tudom, hogy *pro* {jön/jönnek}
 six invitee-about I.know that come.3SG/3PL
 ‘about/for six invitees, I know that they are coming’

11 Katalin É. Kiss (p.c.) also notes the optional plural inflection in *vártam a hat fiút, és meg is jött(SG)/jöttek(PL)* ‘I expected the six boys, and indeed they came’, a case of sentence-level coordination. It may be that the close semantic connection between the coordinated sentences is sufficient to facilitate singular agreement here (unlike in (17)). The definiteness of the antecedent (*a hat fiút*) may also be playing a role in making singular agreement possible.

12 I took the example in (19) from Keresztes (2012), who points out that the ban on pro-drop for plurals does not carry over (for all speakers) to first- and second-person objects (*pace* Farkas 1987, who was the first to draw attention to the number restriction on Hungarian object pro-drop). Németh T. (2016) discusses the link between object pro-drop and singularity in Hungarian in detail, and claims that dropped definite objects can be plural in all persons. Unfortunately her discussion does not control for the distinction between pro-drop and argument ellipsis, which muddies the waters.

- (19) (ők) elbújtak elölünk, (mi) mégis megtaláltuk *(őket)
 they PV.hide.PST.3PL.INDEF from.us we still PV.find.PST.1PL.DEF they.ACC
 ‘they hid from us, still we found them’

In light of this, an interesting question that arises for delative resumptive prolepsis constructions is whether substitution of morphologically singular but notionally plural *hány lány-ról* ‘about how many girl(s)’ for *ki-ről* in (4b) continues to allow the embedded object to remain silent (as *pro*), or instead forces the insertion of overt *őket* ‘them’ in the object position of the embedded clause. It appears that speakers disagree on the acceptability of (20) without *őket* included in it.

- (20) hány lány-ról gondol-od, hogy meghívta %(*őket*) a buliba?
 how.many girl-DEL think-2SG.DEF that invited.3SG.DEF them the party.to

Speaker variation on (20) may be rooted in two factors: (a) the degree to which speakers allow a third-person plural object pronoun to undergo pro-drop (recall fn. 12), and, probably more importantly (in light of the robustness of the generalisation that (19) represents), (b) the question of whether speakers permit the resumptive pronoun linked to delative *hány lány-ról* to be singular (i.e., to show formal number concord with its antecedent). A particularly interesting pattern is the one reported by Krisztina Szécsényi (p.c.): for her both versions of (3b) are perfectly acceptable, but in (20) *őket* cannot be omitted (just as it cannot in (19)). This suggests that for speakers of this type, the resumptive pronoun in delative prolepsis constructions can be either singular or plural when it is the subject of the lower clause, but must be plural when it serves as the object of that clause. This is coherent if distance is a key factor in the licensing of formal number concord between a resumptive pronoun and its antecedent: the subject of the lower is closer to the proleptic delative phrase than the object is; if, as is likely, the propensity to use notional number increases with the distance between the two relata, the fact that (3b) shows optional plural marking while (20) forces the object to be plural can be understood. For Katalin É. Kiss (p.c.), however, plural marking in (3b) and pronoun realisation in (20) are both optional, which can be taken to indicate that for her the distance between the object resumptive and the proleptic delative in the latter example is not too large for formal number concord (recall also fn. 11).

Be that as it may, there appears to be widespread agreement on the grammaticality of (20) with *őket* included. The fact itself that an overt pronoun is grammatical in (20) suggests strongly that the resumptive prolepsis analysis proposed for the examples in (3b) and (4b) is on target. On this note, I will close the discussion of the b–sentences, and move on to the c–examples.

4.2.2 Upstairs indefinite agreement with the wh–phrase (INDEF)

The c–sentences, repeated below, show clear parallels with the b–sentences: the plural subject inflection seen in (3c) is also licit in (3b); and (4b,c) both feature definite object agreement downstairs.

- (3c) hány lány-t gondol-sz, hogy jön-nek a buliba?
 how.many girl-ACC think-2SG.INDEF that come-3PL the party.to
- (4c) ki-t gondol-sz, hogy meghívta a buliba?
 who-ACC think-2SG.INDEF that PV.invited.3SG.DEF the party.to

These c-examples are the first of three constructions that each feature indefinite object agreement in the matrix clause. I will analyse their syntax in section 4.2.2.1, after which I will turn to the other two cases involving upstairs indefinite inflection: the d- and e-examples.

4.2.2.1 *The c-sentences: Resumptive prolepsis with an accusative matrix wh-expression*

The example in (3c) differs from (3a) in three respects — they go separate ways for each of the three possible points of variation: the case of the *wh*-filler (NOM in (3a) but ACC in (3c)); the object inflection of the matrix verb (DEF in (3a) but INDEF in (3c)); and the subject inflection of the embedded verb (SG in (3a), PL in (3c)). Downstairs plural inflection is grammatical only for a subset of speakers (Gervain 2003, 2005); but for those for whom it is good, it is felt to be somewhat better than absence of downstairs singular inflection (as in (3d)).

The only way in which we can sensibly analyse (3c) is by treating it as a variant of the example in (3b) — in other words, as a resumptive prolepsis construction. (3c) differs from (3b) in (a) adorning the *wh*-phrase with accusative rather than relative case, and (b) having the matrix verb agree in definiteness with the accusative *wh*-phrase. The fact that the downstairs verb in (3c) bears plural inflection while the *wh*-phrase is morphologically singular is a secure indication that (3c) involves resumptive prolepsis: the plural inflection of *jönnek* must be controlled by a silent plural pronoun in the embedded clause, showing notional number concord with the proleptic accusative *wh*-phrase in the matrix clause.

The fact that the matrix verb does not entertain an Agree-relation with the subordinate CP makes the establishment of a filler-gap dependency between *hány lányt* and a trace in the embedded clause impossible, so it is a good thing that we do not need to establish such a dependency. The fact that *hány lányt* is marked accusative would of course also have made a filler-gap dependency between it and the nominative subject position of the embedded clause hard to uphold. Within a single chain there cannot be mismatches in case between its members: a gap in a nominative case position can only be bound by a nominative filler. ‘Case switches’ are impossible, regardless of one’s perspective on the directionality of structure building and syntactic derivation — but particularly so in a top-down approach: the upstairs filler is whatever it is; it cannot pick up or lose a case feature further down the tree.

So (3c) features a silent pronoun in the lower clause, linked to the *wh*-phrase in the matrix clause, which originates upstairs. The schematic structure for (3c) is thus as in (21).

(21) wh_{ACC} V+INDEF t_{wh} [_{CP} ● ... *pro* ...] [● marks opacity]

A prediction that this analysis of (3c) makes in conjunction with what I said about (3a) is the following. Since in (3c)/(21) the embedded clause is not in an Agree-relation with the matrix verb, we expect it to be impossible to establish a filler-gap dependency across its CP boundary. So we now predict that (22) should crash: *kit* should not be able to bind its trace in the embedded clause because CP is opaque. This is borne out: examples like those in (22a,b) are ungrammatical, regardless of the relative order of the two *wh*-phrases. Note that the ill-formedness of (22a) cannot be blamed on the co-occurrence of two structural accusatives in the same clause: (22b) is equally ungrammatical.

(22) a. *{hány lány-t kit} gondol-sz, hogy imád-nak?
 how.many girl-ACC who.ACC think.2SG.INDEF that adore-3PL.INDEF

- b. *{hány lány-t kinek} akar-sz, hogy küldje-nek
 how.many girl-ACC who.DAT want.2SG.INDEF that send.SUBJUNC-3PL.INDEF
 levelet?
 letter.ACC

The analytical claim that the subject of the lower clause in (3c)/(21) is a pronoun linked to the *wh*-phrase that binds a trace in the matrix clause also makes a further prediction, in the realm of objects. Recall that object pro-drop is impossible in Hungarian for third-person plural objects. Recall also that resumptive *pro* gets its number specification via concord with its antecedent — concord for *notional* number in cases in which the antecedent is semantically plural but morphologically singular, such as *hány lány* ‘how many girl(s)’. For the relative resumptive prolepsis construction, we had already seen in (20) that there are speakers for whom the resumptive must be overt in this case. For (4c), with an accusative *wh*-constituent in the matrix clause, (23) now shows that the same thing is true: the same speakers who need *őket* to be overt in (20) also require it in (23).

- (23) hány lány-t gondol-sz, hogy meghívta %(*őket*) a buliba?
 how.many girl-ACC think-2SG.INDEF that invited.3SG.DEF them the party.to

The facts in (24) and (25) further undergird the conclusion that (4c), just like (3c), involves resumptive prolepsis.

- (24) a. ?kit szeretnél, hogy meghívjam?
 who.ACC would.like.2SG.INDEF that PV.invite.SUBJUNC.1SG.DEF
 b. kit szeretnél, hogy meghívjam őt?
 who.ACC would.like.2SG.INDEF that PV.invite.SUBJUNC.1SG.DEF 3.ACC
- (25) a. *kiket szeretnél, hogy meghívjam?
 who.PL.ACC would.like.2SG.INDEF that PV.invite.SUBJUNC.1SG.DEF
 b. kiket szeretnél, hogy meghívjam őket?
 who.PL.ACC would.like.2SG.INDEF that PV.invite.SUBJUNC.1SG.DEF 3PL.ACC

The fact that the matrix verb *szeretnél* has INDEF inflection is a secure indication, from the perspective of the preceding discussion, that the accusative *wh*-elements in the matrix clause in the examples in (24) and (25) cannot bind a trace in the embedded clause. Now, interestingly, when the object position in the lower clause is occupied by an overtly realised pronoun, as in (25b), this pronoun is welcome to be a plural associated with plural *kiket* upstairs; but when the object pronoun associated with the *wh*-word in the higher clause remains silent, only singular *kit* is grammatical: (21a) is ill-formed. The contrast between (24a) and (25b) strongly confirms the hypothesis that the gap associated with the upstairs *wh*-element in (24) and (25) is pronominal: only then can this contrast be directly related to an independently known fact about Hungarian object pro-drop, viz., the obligatoriness of overt pronoun realisation in the third-person plural (as demonstrated in (19), above).

The conclusion that the embedded object position in (24) and (25) is filled by a pronoun also makes direct sense of the fact that in all grammatical examples, the verb in the subordinate clause bears DEF inflection: the *wh*-words *kit* and *kiket* are themselves indefinite, hence a silent copy of *kit* or *kiket* in the lower clause could never have legitimated the morphological form of the lower verb.

All the evidence indicates, therefore, that (3c) and (4c) feature a pronominal empty category in the subordinate clause, linked to the *wh*-operator originating in the matrix clause. One final confirmation comes from the insensitivity of the examples in (3c) and (4c) to islands.

Since there is no cross-clausal antecedent–trace dependency in the syntax of (3c) and (4c), we expect the c–sentences to be oblivious to *wh*-islands. And indeed, there is a marked contrast between the examples in (8), repeated below as (26), and the ones in (27).

- (26) a. *hány lány kérdezted, hogy mikor jött el? (= (8))
 how.many girl(NOM) ask.PST.2SG.DEF that when came PV
 b. *hány lány kérdezted, hogy eljött-e?
 how.many girl(NOM) ask.PST.2SG.DEF that PV.came-QPRT
- (27) a. hány lány-t kérdeztél, hogy mikor jött-ek el?
 how.many girl-ACC ask.PST.2SG.INDEF that when came-3PL PV
 b. hány lány-t kérdeztél, hogy eljött-ek-e?
 how.many girl-ACC ask.PST.2SG.INDEF that came-3PL-QPRT

The examples in (27) are grammatical to the same degree (and with the same degree of speaker variation) as (3c). Unlike in the case of (26), therefore, there is no penalty for building the ‘girl’-dependency across an intervening *wh*- or yes/no-question operator. This is a strong confirmation of the analysis proposed for (3c), as schematised in (21), according to which the empty category in the subject position of the subordinate clause is a pronoun, not a *wh*-trace.

In this regard, (3c) is just like (3b). But (3c) differs from (3b) in the way in which it represents the proleptic object in the matrix clause: the accusative *wh*-phrase is projected as an object of the verb *gondolsz*. This is something (standard) English cannot do.¹³ I contend that this difference between English and Hungarian has to do with another salient difference between the two languages: the fact that in Hungarian, the clausal complement of a verb like ‘to think’ or ‘to say’ is customarily accompanied by a pronominal proleptic pronoun bearing accusative case. This is shown in (28).

- (28) a. az-t gondolom hogy ...
 it-ACC think.1SG.DEF that
 b. I think (*it) that ...

So the claim that I would like to make is that *hány lányt* ‘how many girl.ACC’ in (3c) originates in the same position that the proleptic object *azt* originates in in the syntax of (28a). In (28a), the verb is responsible for the assignment of accusative case to *azt*, and concomitantly *azt* controls definite agreement with the verb. In (3c), it is again the verb that assigns accusative case to *hány lányt*; and since this time the verb’s accusative dependent is indefinite, the verb bears indefinite inflection.

In chapter 2, I argued that the syntax of the object domain includes at least two positions — one in the complement of V and the other serving as V’s specifier. The structures in (29) reproduce the ones I presented in chapter 2.

13 It is imaginable that *whom shall I say is calling?* (Jespersen 1954, Zwicky 2007) can be analysed along these lines, though the fact that the ‘*that*-trace filter’ seems to be in effect in these constructions (*whom shall I say that is calling?* is probably degraded) suggests that the silent subject of the embedded clause is a trace rather than a pronoun.

- (29) a. $[_{VP} \text{SUBJECT } [_{\nu'} \nu [_{VP} \langle \text{OBJECT} \rangle [_{V'} V \langle \text{OBJECT} \rangle]]]]$
 b. $[_{VP} \text{SUBJECT } [_{\nu'} \nu [_{VP} \text{PROFORM}_i [_{V'} V [_{CP} \dots]_i]]]]]$

In (28a), *azt* is associated with the SpecVP position, with the clause occupying the complement-of-V slot, as depicted in (29b). The specifier position of V can alternatively be exploited for the benefit of a proleptic object that binds a gap inside the subordinate clause. This is how we arrive at the syntax of (3c) and (4c). The proleptic object in SpecVP is the closest possible goal for the case probe ν , which explains the accusative case form of *hány lány-t* in (3c) as well as the fact that this *wh*-phrase controls indefinite agreement on the matrix verb: despite the fact that the *wh*-phrase in this example corresponds to the subject of the subordinate clause, it gets accusative case and agrees in definiteness with the matrix verb because it serves as the proleptic object of the matrix verb. With these things in mind, I now propose (30) (a more precise version of (21), above) as the full structure for (3c). Here the *pro* in the matrix clause represents the addressee, and the *pro* in the subordinate clause engages in a relationship of notional plural concord with *hány lányt* in the matrix clause.

- (30) $[_{FocP} \text{hány lányt } [F [_{TP} \text{pro}_{2SG} [T [_{\nu'} \nu [_{VP} \text{hány lányt } [V [_{CP} \text{hogy pro}_{3PL} \dots]]]]]]]]]]$

At this point, recall from chapter 2 that SpecVP is the position for *specific* or *presuppositional* objects: while non-factive complement clauses occupy the complement-of-V position, factive complement clauses, which by nature contain presupposed content, get mapped into SpecVP or associated to an argumental *azt* in that position. Now note that *hány lányt* in (30) binds a silent copy in this very same position, SpecVP. This makes a prediction regarding the interpretation of *hány lányt* in sentences of the type in (3c): because of its location outside the domain of existential closure, the accusative *wh*-phrase in (30) should only be able to support a *presuppositional* interpretation. This turns out to be true: the existence of a group of girls is presupposed; a pure cardinality reading is not available. That this is the case is shown particularly clearly by the fact that (31), a variant of (3c) which differs from it only in the addition of *összesen* ‘in total’, is not well-formed: even speakers who accept (3c) make a clear contrast between it and (31), which they reject.¹⁴

- (31) (~(3c)) **összesen* hány lány-t gondol-sz, hogy jön-nek a buliba?
 in.total how.many girl-ACC think-2SG.INDEF that come-3PL the party.to

One might respond to this by saying that the fact that my analysis of (3c) in (30) postulates a pronoun in the embedded clause linked to the *wh*-phrase in the matrix clause should in itself already bias (3c) towards a specific interpretation for *hány lányt*. But there are at least two cogent empirical reasons why such a response to the datum in (31) would not be on target.

The first is that in the relative prolepsis construction in (3b), for which I argued in section 4.2.1.2 that the empty category in the lower clause is likewise a pronoun, no specificity effect crops up: (32), the version of (3b) with *összesen* ‘in total’ included, is perfectly well-formed.

- (32) (~(3b)) *összesen* hány lány-ról gondol-od, hogy jön(-nek) a buliba?
 in.total how.many girl-DEL think-2SG.DEF that come-3PL the party.to

14 See Obenauer (1994) for an entirely parallel specificity effect for past-participle agreement constructions in French, analysable along the same lines. For some discussion, I refer the reader to section 4.2.5.1, below.

From the perspective of the above discussion of (31), this is entirely as expected: the relative phrase *hány lány-ról* is not an object of the matrix verb; it is an adjunct, its trace not situated in the SpecVP position for specific/presuppositional internal arguments.

The second indication that the specificity effect seen in (3c) cannot be blamed on the resumptive pronoun is the exact converse of the one emerging from (32): this obligatorily presuppositional interpretation also surfaces in (3d), and as the discussion in the next subsection will show, there is no pronoun in the structure of the downstairs clause in (3d).

4.2.2.2 The *d*-sentences: Wh-topic drop

The construction exemplified in (3d) and (4d) occupies a special place in the paradigm of long \bar{A} -dependencies in Hungarian: it shares properties with both the *a*-sentences and the *c*-constructions yet cannot be assimilated to either. For easy comparison, I am repeating (3a,c,d) and (4a,c,d) below.

- | | | | | | | | | |
|-----|----|----------|-----------|-----------------|------|----------|-----|----------|
| (3) | a. | hány | lány | gondol-od, | hogy | jön | a | buliba? |
| | | how.many | girl(NOM) | think-2SG.DEF | that | come.3SG | the | party.to |
| | c. | hány | lány-t | gondol-sz, | hogy | jön-nek | a | buliba? |
| | | how.many | girl -ACC | think-2SG.INDEF | that | come-3PL | the | party.to |
| | d. | hány | lány-t | gondol-sz, | hogy | jön | a | buliba? |
| | | how.many | girl -ACC | think-2SG.INDEF | that | come.3SG | the | party.to |
-
- | | | | | | | | |
|-----|----|---------|-----------------|------|----------------------|-----|----------|
| (4) | a. | ki-t | gondol-od, | hogy | meghívott | a | buliba? |
| | | who-ACC | think-2SG.DEF | that | PV.invited.3SG.INDEF | the | party.to |
| | c. | ki-t | gondol-sz, | hogy | meghívta | a | buliba? |
| | | who-ACC | think-2SG.INDEF | that | PV.invited.3SG.DEF | the | party.to |
| | d. | ki-t | gondol-sz, | hogy | meghívott | a | buliba? |
| | | who-ACC | think-2SG.INDEF | that | PV.invited.3SG.INDEF | the | party.to |

On the surface, (3d) seems to be a mix of (3a) and (3c), and (4d) looks like a combination of (4a) and (4c). With the *a*-sentences, the *d*-examples share the downstairs portion of the string: singular inflection on the lower verb in (3a,d); indefinite inflection on *meghívott* in (4a,d). On the other hand, the *d*-examples have in common with the *c*-sentences the entire upstairs portion of the string: accusative case on the *wh*-phrase (esp. relevant for (3)), and indefinite inflection on the matrix verb. The ‘Janus-faced’ nature of (3d) and (4d) actually runs deeper than these morphological parallels. There seem to be fundamental ways in which *downstairs* the *d*-sentences are very much like the *a*-examples while *upstairs* they are just like the *c*-sentences. In the discussion to follow, I will concentrate on (3d), whose ‘Janus-faced’ behaviour is more salient than that of (4d).

One salient interpretive feature that (3d) shares with (3c) is the specificity effect I noted at the end of the previous subsection: (33), based on (3d), is equally bad as (31), above.

- (33) (~(3d)) **összesen* hány lány-t gondol-sz, hogy jön a buliba?
 in.total how.many girl-ACC think-2SG.INDEF that come.3SG the party.to

In this respect, (3c,d) as a pair are markedly different from both (3c) (recall (32)) and (3a):

- (34) (~(3a)) *összesen* hány lány gondol-od, hogy jön a buliba?
 in.total how.many girl(NOM) think-2SG.DEF that come.3SG the party.to

This confirms that in the upstairs portion of the clause, where specificity is encoded for the *wh*-constituent, (3d) patterns with (3c) rather than with (3a) — even though with (3a) the construction in (3d) shares the singular inflection of the downstairs verb.

Of course this specificity effect is not *just* a function of the fact that the *wh*-constituent in (3c,d) is accusative and agrees in definiteness with the matrix verb: it is entirely possible for accusative agreeing objects to receive a non-specific interpretation, as the simple transitive in (35) shows.

- (35) *összesen* hány lány-t szeret-sz?
 in.total how.many girl-ACC like-2SG.INDEF
 ‘how many girls in total do you like?’

What is responsible for the specificity effect in (31) and (33) is the fact that the accusative and upstairs-agreeing *wh*-constituent in these examples, unlike in (35), can only be mapped onto SpecVP: the complement-of-V object position is already taken by the complement clause.

It is important to stress that unlike in the case of the c-examples, we cannot arrive at an account of the specificity effect for the d-sentences by capitalising on the presence of a silent pronoun in the lower clause. The fact that the lower verb in (4d) shows up in with indefinite (or subjective) inflection (*meghívott*, as opposed to *meghívta* in (4c)) forces us to conclude that the object gap in the structure of the d-sentences is not pronominal: third-person object pronouns in Hungarian, whether silent or overt, always control definite agreement with the finite verb. So while (4c) definitely contains a silent pronoun in the object position of the lower clause, (4d) does not. We need a non-pronominal empty category in the lower clause in the d-sentences.

This is further confirmed by the fact that the d-examples share with a-sentences the ability to license a parasitic gap in the lower clause. In this respect as well, the d-sentences differ starkly from the c-examples, in which this is impossible. Let me illustrate this for the subject dependencies in (3), with the aid of the triple in (36)–(38) (modelled on original material in Horvath 1992). In Den Dikken (2009b), I point out that while (3c) does not support a parasitic gap in the downstairs clause (see (37)), (3d) does (to the degree that parasitic gaps are generally acceptable in Hungarian), as (38) shows. To this picture I now add that (3a) is just like (3d) in this regard: (36) is grammatical.

- (36) (~(3a)) ^(?) hány lány mondtad, hogy jött a buliba anélkül hogy meghívtál volna?
 ?Q girl(NOM) said.2S.DEF that came.3SG the party-to without that invited.2SG would
 ‘how many girls did you say came to the party without you having invited (them)?’
- (37) (~(3c)) *hány lány-t mondtál, hogy jött-ek a buliba anélkül hogy meghívtál volna?
 ?Q girl-ACC said.2S.INDF that came-3PL the party-to without that invited.2SG would
- (38) (~(3d)) ^(?) hány lány-t mondtál, hogy jött a buliba anélkül hogy meghívtál volna?
 ?Q girl-ACC said.2S.INDF that came.3SG the party-to without that invited.2SG would

The fact that no parasitic gap is licensable in the embedded clause in (37) is of course straightforwardly predicted by the resumptive prolepsis analysis that I proposed for (3c): no \bar{A} -movement obtains within the embedded clause there; but parasitic gaps are dependent for their licensing on the

presence of a variable in the local syntactic environment. From (36) and (38), on the other hand, we are led to conclude that there IS an \bar{A} -bound variable present in the lower clause.

The analysis of the a–sentences presented in section 2.1 delivers this variable: recall (10). For the d–examples, we are now on the look-out for a syntactic analysis that shares with (10) the presence of a variable in the downstairs clause linked to the *wh*-constituent in the upstairs clause. But a direct dependency between *hány lányt* ‘how many girl.ACC’ and a *wh*-trace in the subject position of the embedded clause cannot be established in (3d): the *wh*-expression has the wrong case to be directly associable to the structural subject position of the lower clause. Another cogent reason to believe that building a direct \bar{A} -dependency across the embedded CP, as is done in (10), is impossible in (3d) and (4d) is the fact that the matrix verb does not entertain an Agree relation with the embedded clause in the d–sentences (whose matrix verbs bear indefinite inflection): this lack of definiteness agreement makes CP opaque.

That there is no direct \bar{A} -dependency across the embedded CP in the syntax of the d–sentences is confirmed by another striking bifurcation in the paradigms in (3) and (4), this time with regard to sensitivity to intervening *wh*-operators. For (3a), the facts in (26), repeated below as (39) (with a c–example added), had shown us in section 4.2.2.1 that it evinces a clear *wh*-island effect — unsurprisingly, in light of the analysis in (10), which correctly leads us to expect an intervention effect. For (3c), I had emphasised in that same section that it does not show any sensitivity to intervening *wh*-operators. This was illustrated in (27), repeated here as (40) (again, with a c–example added). In (41), I now complete the *wh*-island picture by adding the new facts for (3d). While (3d) is just like (3a) when it comes to parasitic gap licensing (recall (36) and (38)), it patterns with (3c) in showing no sensitivity to the presence of a *wh*- or yes/no-question operator in the subordinate clause.

- (39) a. *hány lány kérdezted, hogy mikor jött el? (= (26); ~(3a))
 how.many girl(NOM) ask.PST.2SG.DEF that when came PV
 b. *hány lány kérdezted, hogy eljött-e?
 how.many girl(NOM) ask.PST.2SG.DEF that PV.came-QPRT
 c. *hány lány vagy kíváncsi, hogy eljön-e?
 how.many girl(NOM) be.2SG curious that PV.come-QPRT
- (40) a. hány lány-t kérdeztél, hogy mikor jött-ek el? (= (27); ~(3c))
 how.many girl-ACC ask.PST.2SG.INDEF that when came-3PL PV
 b. hány lány-t kérdeztél, hogy eljött-ek-e?
 how.many girl-ACC ask.PST.2SG.INDEF that came-3PL-QPRT
 c. hány lány-ra vagy kíváncsi, hogy eljön-nek-e?
 how.many girl-SUBL be.2SG curious that PV.come-3PL-QPRT
- (41) a. hány lány-t kérdeztél, hogy mikor jött el? (~(3d))
 how.many girl-ACC ask.PST.2SG.INDEF that when came.3SG PV
 b. hány lány-t kérdeztél, hogy eljött-e?
 how.many girl-ACC ask.PST.2SG.INDEF that came.3SG-QPRT
 c. hány lány-ra vagy kíváncsi, hogy eljön-e?
 how.many girl-SUBL be.2SG curious that PV.come.3SG-QPRT

These observations about English, which presumably reveal UG properties of null-operator dependencies (see section 5.3.13 in chapter 5), militate against an identification of the binder of the variable in the embedded clause of (3b) as a null operator, in the technical sense of the term.

An important property of the d-sentences is that the *wh*-operator in the matrix clause truly does seem to ‘belong’ in the lower clause: it is in some sense selected there. In this respect, the d-sentences again behave differently from established null-operator constructions such as *tough*-movement cases. As Wilder (1991:123) has pointed out, selectional restrictions imposed by the predicate of the *to*-infinitive can strikingly fail to be satisfied by the subject of the *tough*-adjective:

- (46) a. [_{CP} for him to be top of the class] is hard to believe
 b. *I can’t believe [_{CP} for him to be top of the class]

As is well known, *believe* does not normally select *for-to* infinitives: (46b) is ill-formed. Yet (46a) is fine. So the subject of the *tough*-adjective does not seem to (have to) ‘belong’ in the infinitival clause. But the *wh*-operator of the d-sentences does behave in almost every way as though it belongs to the subordinate clause — it is only its case and its control of upstairs verb agreement that tell us that the physical *wh*-operator is a constituent of the matrix clause alone.

Hungarian-internal considerations also plead strongly against (43). From parasitic gap constructions in Hungarian, we know that whenever the parasitic gap corresponds to the object of the inflected verb, the verb always bears *indefinite* inflection. This is true entirely regardless of the (in)definiteness of the binder of the ‘real’ gap. Thus, consider the pair of examples in (47). In (47a) the *wh*-phrase *hány lányt* is indefinite, as shown by the indefinite inflection on the matrix verb; in (47b), *melyik lányt* is definite, controlling the definite conjugation of the verb in its clause. But in both examples, the parasitic gap occurs with a verb exhibiting indefinite agreement. Regardless of the (in)definiteness of the binder of the ‘real’ gap, the inflection of the verb in the clause containing the parasitic gap is always indefinite.

- (47) a. *hány lányt* *vártál* *anélkül hogy meghívtál* *volna?*
 how.many girl.ACC waited.2SG.INDEF without that invited.2SG.INDEF would
 b. *melyik lányt* *vártad* *anélkül hogy meghívtál* *volna?*
 which girl.ACC waited.2SG.DEF without that invited.2SG.INDEF would
 ‘how many girls/which girl did you wait for without having invited?’
 ‘which girl did you say came to the party without you having invited (her)?’

Succinctly put, null operators are indefinite. With this in mind, consider (48b) (from Den Dikken 2009b:fn. 17) and compare it with (47) and (48a). If the operator in the lower clause in the d-sentences were a null operator, it ought to consistently behave like an indefinite, and hence, as an object, systematically trigger indefinite inflection on the finite verb. But the fact of the matter is that when the *wh*-constituent in the matrix clause corresponds to the object of the embedded verb and is definite, we never find indefinite object agreement downstairs. Sentences such as (48b) are grammatical only with definite inflection (*meghívta*) in the subordinate clause.¹⁶

16 This stands in the way of an application to Hungarian of analyses such as Schneider-Zioga’s (2009) interesting account of Kinande long *wh*-dependencies in terms of partial null operator movement.

- (48) a. *hány lányt gondolsz, hogy {meghívott/*meghívta}?*
 how.many girl-ACC think-2SG.INDEF that PV.invited.INDEF/*DEF
 ‘how many girls do you think he invited?’
- b. *melyik lányt gondolod, hogy {meghívta/*meghívott}?*
 which girl-ACC think-2SG.DEF that PV.invited.DEF/*INDEF
 ‘which girl do you think he invited?’

With the null-operator analysis hereby discarded, let me put forward the analysis which I would like to advocate for the d–sentences: an analysis that could be captured under the rubric of ‘*wh*-topic drop’. The essence of this account goes back to a point I made in section 3.4.1.3, in the discussion of *wh*-island effects.

There I mentioned that the copy of a matrix *wh*-constituent that is re-uploaded onto the pushdown stack of an intervening *wh*-element in a subordinate clause is anaphoric to the *wh*-expression in the matrix clause, hence serves as a topic in the lower clause. Though in d–type long \bar{A} -dependencies in Hungarian there is no re-uploading of the matrix *wh*-expression going on, what these constructions share with argumental *wh*-island cases is that they involve a silent *wh*-operator in the embedded clause that is anaphoric to the overt *wh*-operator in the matrix clause, and therefore functions as a topic within the lower clause. That the silent *wh*-operator in the subordinate clause is indeed a topic is clear from the absence in the lower clause of d–type sentences of the ‘preverb–verb inversion’ triggered by focus fronting in Hungarian. We see this in (48): the preverb *meg* precedes rather than follows the finite verb, and must so do; (49a,b), with inversion, are ungrammatical:

- (49) a. **hány lányt gondolsz, hogy hívott meg?*
 how.many girl-ACC think-2SG.INDEF that invited.INDEF PV
- b. **melyik lányt gondolod, hogy hívta meg?*
 which girl-ACC think-2SG.DEF that invited.DEF PV

So what we are looking for in the embedded clause is a silent element that is anaphoric to the *wh*-chain in the matrix clause, but not a pronoun. My proposal, following in the footsteps of the discussion of argumental *wh*-island cases in section 3.4.1.3, is that there is a silent copy of the matrix *wh*-phrase in the topic field of the embedded clause — a surface anaphor, lacking a PF matrix:

- (50) $[_{CP1} wh_i \dots [_{CP2} \dots [_{TopP} wh_i [_{TopP} Top [_{COMMENT} \dots]]]]]$

Unlike in the case of *wh*-island constructions, the silent *wh*-topic in the embedded is not a literal copy of the matrix *wh*-constituent: it is not the case that, in the course of the derivation of d–type constructions, the a copy of the matrix *wh*-phrase is re-uploaded onto the pushdown stack of a *wh*-operator on the edge of the embedded clause; there is no such intervening *wh*-operator in the structure of d–type sentences. As a matter of fact, when there *is* an intervening *wh* in d–type sentences, no ‘*wh*-island effect’ manifests itself: recall (41). That the silent *wh*-topic in (50) is not a literal copy of the *wh*-phrase in the matrix clause is also clear from the fact that the two *wh*-operators can have non-identical sets of morphological features: in (3d), the overt *wh*-phrase in the matrix clause bears an accusative case feature (spelled out as the *-t* suffix on *lány-t*) while the *wh*-element in the embedded finite clause serves as its subject, and hence must be nominative.

The silent *wh*-topic in the embedded clause in (50) is thus created independently of the *wh*-operator in the matrix clause, and forms a chain of its own. No chain is formed between *hány lányt* in the upstairs clause and the *wh*-dependency in the embedded clause: the chain that *hány lányt* is in terminates in the VP-specifier position for the matrix object; a separate chain is formed between a silent *wh*-operator and a variable in the subject position — a *topic* chain, unhindered by any *wh*- or question operator in the embedded clause (whence the grammaticality of (41)).¹⁷

The ‘*wh*-topic’ in (50) is necessarily silent: while (51a) is a grammatical example of the d-type, (51b) is ill-formed.

- (51) a. ki-t gondol-sz, hogy meghívott? (cf. (4d))
 who-ACC think-2SG.INDEF that invited.3SG.INDEF
 b. *ki-t gondol-sz, hogy ki-t meghívott?
 who-ACC think-2SG.INDEF that who-ACC invited.3SG.INDEF

The ungrammaticality of (51b) is not difficult to understand: the *wh*-topic in the embedded clause is formally identical with the matrix *wh*-operator, and is c-commanded by it; its entire content is recoverable from the fact that it is anaphoric to the matrix *kit*, which is fully licensed as a matrix-clause constituent independently of the embedded *wh*-topic (as we know from the c-sentences). The recoverability of the *wh*-topic in the embedded clause in conjunction with the fact that its antecedent is autonomously licensed leads to the obligatory silence of the second *kit* in (51b).

Though there are non-trivial differences, there is a significant parallel between Hungarian d-type sentences and English cleft-like constructions of the type in (52) (analysed by Reeve 2012).

- (52) who did you see *Op* that you liked?

Here, too, the *wh*-operator in the embedded clause is anaphoric to the *wh*-operator in the matrix clause, and formally fully identical with it and recoverable from it. Moreover, the matrix *wh* c-commands the operator-variable dependency in the embedded clause (though it is not entirely clear whether the foot of the chain of *who* c-commands the *that*-clause; much depends here on the details of the analysis of sentences of the type in (52), which I will not delve into). I find it significant that in (52) the operator in the embedded clause is silent, and does not alternate at all naturally with *who* (^{??}*who did you see who you liked?*, which contrasts with the ordinary cleft *it was John {that/who} I liked best*, where an overt *wh*-operator is perfectly fine).

Analysed as in (50), the d-examples are a case of obligatory topic drop affecting an anaphoric *wh*-operator. That topic drop exists as a feature of Universal Grammar is clear from the grammaticality of Dutch sentences of the type in (52b). Dutch also shows that *wh*-operators can be dropped: ‘*wh*-drop’ is illustrated in (53b).

17 That topicalisation and focalisation do not interfere with one another is particularly clear from the fact that the two happily co-occur in a single sentence, and give rise to no superiority effect (recall the discussion in chapter 2).

- (i) a. [_{Topic} Mari] [_{Focus} hány lányt] mutatott be Jánosnak?
 Mari how.many girl.ACC introduced.3SG.INDEF János.DAT
 b. [_{Topic} Jánosnak] [_{Focus} hány lányt] mutatott be Mari?
 János.DAT how.many girl.ACC introduced.3SG.INDEF Mari
 both: ‘how many girls did Mari introduce to János?’

- (52) a. dat boek heb ik al gelezen (Dutch)
 that book have I already read
 ‘that book I’ve read already’
 b. _____ heb ik al gelezen
 have I already read
- (53) a. wat heb je nou gedaan? (Dutch)
 what have you now done
 ‘now what have you done?’
 b. _____ heb je nou gedaan?
 have you now done

The embedded clause of (50) is a combination of the properties of (52b) and (53b). The silent *wh*-operator in the embedded clause is anaphoric to the *wh*-constituent upstairs, hence topical (unlike the silent *wh*-operator in Dutch (53b)) and subject to topic drop (as in Dutch (52b)).

Because of the fact that the operator in the embedded clause in (50) is identical (*modulo* case) to the *wh*-operator in the matrix clause, the representation of the d–sentences looks very much like that of the result of a traditional, bottom-up successive-cyclic movement derivation. But on a successive-cyclic movement derivation, the complex tangle of facts presented by the d–sentences would be very difficult to straighten out. The ‘case switch’ seen in (3d) might be reconciled with bottom-up successive cyclicity — indeed, some might even herald it as a bonus of that analysis, with accusative case assignment to an intermediate copy of the chain (see Kayne 1984 and work inspired by it). But (a) the absence of intervention island effects (recall (41)) and (b) the information-structural difference between the matrix dependency and the one in the embedded clause both bring sharply into focus the fact that it is emphatically not the case in the derivation of the d–sentences that a single chain is formed that spans two clauses. The derivation of the a–examples, by contrast, does involve a single chain that spans two clauses — but here, no intermediate copy on the edge of CP is needed because Agree with V makes CP transparent.

4.2.2.3 The *e*–sentences: Wh-scope marking

After this fairly lengthy discussion of the d–sentences, I would like to complete the survey of the empirical landscape of Hungarian long-distance \bar{A} -dependencies by saying a few words about the *e*–examples in (3) and (4), the last members of the Hungarian typology of long \bar{A} -dependencies with upstairs agreement with the *wh*-operator (here *mit* ‘what.ACC’).

- (3e) mi-t gondol-sz, hogy hány lány jön(*-nek) a buliba?
 what-ACC think-2SG.INDEF that how.many girl(NOM) come-*3PL the party.to
- (4e) mi-t gondol-sz, hogy ki-t hívott meg a buliba?
 what-ACC think-2SG.INDEF that who-ACC invited.3SG.INDEF PV the party.to

In the scope-marking constructions in (3e) and (4e), the embedded clause is a *wh*-question, with *hány lány/kit* in the focus position of the embedded clause (as witness the ‘preverb–verb inversion’ in (4e)); its scope is indicated by the *wh*-scope marker (*mit* ‘what-ACC’ in the cases under

discussion) in the matrix clause. At no point does the *wh*-constituent in the embedded clause move out of that clause to ‘replace’ the *wh*-scope marker — there is only an indirect dependency between the *wh*-scope marker and the *wh*-constituent in the embedded clause (see Dayal 1994, Horvath 1997, Bruening 2006): the radical impossibility of accusative case checking performed by the ‘real’ *wh*-constituent (**mit gondolsz hogy hány lány-t jön a buliba*) testifies to this. The *wh*-subject of the embedded clause in (3e) checks ϕ -features and nominative case downstairs, and has no feature-checking relationship with the matrix v -V, from which it is separated by an opaque CP boundary.

The literature is brimming with accounts of the scope marking construction. I will not be able to do it justice here (see Dayal 1994, Horvath 1997, Lipták 2001, Lutz *et al.* 2000, Fanselow 2006, and references cited there for detailed discussion). Instead, I will present one particular outlook on the gross structure of the scope marking construction, the one proposed by Felser (2001).

The central insight of Felser’s (2001) approach is that the embedded clauses in *wh*-scope marking constructions is predicated of the scope marker (*mit* in the Hungarian examples). Felser’s particular way of giving this idea structural shape is to say that the scope marker originates in the specifier position of a Larsonian VP whose head takes the CP that contains the ‘real’ *wh*-constituent as its complement, as depicted in (54) (with the lexical material provided by the Hungarian example in (3e)).

(54) v [_{VP} SM=*mit* [_V V [_{CP} *hogy* [_{FocP} DP=*hány lány* [*jön a buliba*]]]]]

The structure in (54) provides a way of giving formal substance to the idea that the Hungarian scope marker *mit* originates in the same position as the ‘clausal expletive’ *azt* in (55). Recall from chapter 2 that in the structure of *azt*+CP constructions with bridge verbs, *azt* originates in SpecVP and the CP sits in the complement-of-V position.

(55) *azt* gondolom, [_{CP} *hogy* *húsz* *lány* *jön*]
 it/that-ACC think.1SG.DEF that twenty girl come-3SG
 ‘I think that twenty girls will be coming’

Once the ‘clausal expletive’ and the scope marker are assimilated in this way, it follows that scope marking is unavailable in constructions featuring the ‘clausal expletive’ *azt* in Hungarian.¹⁸

As expected, the scope-marking construction in (3e) behaves like (3a) in allowing *hány lány* to license a parasitic gap in the embedded clause: compare (36), repeated below, and (56). An \bar{A} -dependency is established between *hány lány* in the focus position in (55) and a variable below the adjunct containing the parasitic gap.

18 By the same token, it is impossible in German to use scope marking in the presence of the ‘clausal expletive’ *es* (see (i)).

(i) a. womit scheint *es* (dir), dass man ihm helfen kann? (German)
 where-with seems it you that one him help can
 ‘with what does it seem (to you) that one can help him?’
 b. **was* scheint *es* (dir), womit man ihm helfen kann?
 what seems it you where-with one him help can

(36) (~(3a)) ^(?) hány lány mondtad, hogy jött a buliba anélkül hogy meghívtál volna?
 ?Q girl(NOM) said.2S.DEF that came.3SG the party-to without that invited.2SG would

(56) (~(3e)) ^(?) mi-t mondtál, hogy hány lány jött a buliba anélkül hogy meghívtál volna?
 what-ACC said.2S.IND that ?Q girl came.3SG the party-to w/o that invited.2SG would
 ‘how many girls did you say came to the party without you having invited (them)?’

Felser (2001) points out that, thanks to serving as a predicate of the scope-marker pronoun in the matrix clause, the embedded CP in *wh*-scope marking is similar to a relative clause (which likewise serves as a predicate). But importantly, the subject of the *wh*-CP in *wh*-scope marking must be an interrogative *wh*-pronoun. Felser takes this to be an automatic reflex of an ‘interrogative concord’ relationship between the CP and the scope marker. She likens this to the case concord relationship in effect between the subject and a predicate nominal. I will come back to the role of concord in *wh*-scope marking constructions in my analysis of so-called ‘*wh*-copying’ in section 4.3. But before turning there, let me first summarise the results of the discussion of the Hungarian facts.

4.2.3 The landscape of Hungarian long \bar{A} -dependencies in a nutshell

At this point, we have analyses on the table for all the ways in which Hungarian can form a biclausal *wh*-question in which the identity of an argument of the embedded verb is enquired about. Let me summarise the picture that has emerged by going through the various constructions again briefly.

In the a-sentences, a long-distance filler-gap dependency is formed between the indefinite *wh*-operator in the matrix clause and a gap in the embedded clause. This dependency is facilitated by the agreement relation (in definiteness) between the matrix verb and the subordinate clause, which renders the latter transparent. No intermediate trace is present on the edge of the embedded clause because none is necessary: the dependency between the *wh*-operator upstairs and the gap in the subordinate clause can be established directly. The structure in (57) (cf. (10), above) sums this up.

(57) $[wh_{IND} V+DEF \quad [_{CP\oplus} \dots t_{wh} \dots]]$ [⊙ marks transparency]

In the b-sentences, we are dealing with one particular instantiation of resumptive prolepsis. The proleptic object binds a pronominal gap in the matrix clause. This is summarised in (58) (cf. (16)). Because it has oblique (delative) case, the matrix verb cannot agree with it; the verb instead agrees with the embedded clause, which is therefore transparent in principle for the establishment of filler-dependencies across it. But the dependency between the oblique proleptic object and the gap in the embedded clause is a binding relation, not a filler-gap dependency; so the transparency of the embedded clause is irrelevant for it. The silent resumptive is definite, and therefore controls definite agreement with the embedded verb when it plays the role of object in the embedded clause. When it is the subject, its number feature is determined under notional concord with its antecedent; so whenever the antecedent is morphologically singular but denotes a multitude of individuals, *pro* will be plural and control plural agreement with the lower verb.

(58) $[wh_{DEL} V+DEF \quad t_{wh} \quad [_{CP\oplus} \dots pro \dots]]$

The c-examples are also resumptive prolepsis constructions, and hence surface variants of the examples in (3b) and (4b). The only difference between them is the case assigned to the indefinite proleptic object, and, concomitantly, the agreement form of the matrix verb: in the c-sentences, the verb in the matrix clause assigns case to and has its definiteness agreement controlled by the proleptic object. Because the verb does not agree with the embedded clause, it is opaque. The proleptic object is linked, via binding, to a resumptive pronoun in the downstairs clause. The syntax of the c-sentences is summarised in (59) (cf. (21)).

(59) $[wh_{ACC} V^{+INDEF} t_{wh} [_{CP} \bullet \dots pro \dots]]$ [● marks opacity]

The member of the paradigms in (3) and (4) that turned out to be toughest nut to crack is the pattern shown by the d-sentences. For these, I first rejected an analysis in terms of a null operator dependency *à la* parasitic gap or *tough*-movement constructions: when serving as an object, the null operator in such constructions would lead to systematic indefinite agreement in the downstairs clause (because the null operator, i.e., PRO, is formally indefinite in Hungarian); but what we find instead is that the definiteness inflection on the embedded verb is a function of the definiteness properties of the filler in the matrix clause. This might suggest a long-distance movement derivation. But the fact that no intervention island effects manifest themselves in the d-sentences, unlike in the a-examples, strongly obstructs such a treatment of the d-sentences from an empirical point of view. Theoretically (esp. urgently in a top-down model), the fact that the filler in the matrix clause can have a case form that is different from the one expected in the embedded clause (cf. (3d)) also militates against a direct filler-gap dependency between the matrix operator and a gap downstairs. What I ended up defending as the analysis for the d-sentences is a syntax that assimilates them in certain ways to topic-drop, *wh*-drop, and argumental *wh*-island traversing constructions. There are two operators in the structure, one in the focus position in the matrix clause and the other in a topic position in the embedded clause: see (60) (cf. (50)).

(60) $[wh_{ACC} V^{+INDEF} t_{wh} [_{CP} \bullet \cancel{wh}_{Topic} \dots t_{wh} \dots]]$

The two operators have the same referent, and they are formally identical (*modulo* case); moreover, they are in an asymmetric c-command relation. As a result of this combination of identity and asymmetric c-command, the operator in the embedded clause obligatorily remains unexpressed. This is marked by strike-through in (60). It is important to bear in mind, however, that ‘~~*wh*~~’ in (60) does not hint at the application of successive-cyclic movement: it is not the case that a single *wh*-chain is formed in the derivation of (60); rather, two separate chains (a focus chain upstairs, and an anaphoric topic chain downstairs) combine in this special type of prolepsis construction, with the *wh*-operator in the matrix clause uniquely binding a gap in the position of proleptic objects.

In the e-examples, too, this position is utilised — this time by the *wh*-scope marker *mit*. Just as in the d-sentences, two \bar{A} -chains are formed in the syntax of the scope marking constructions in (3e) and (4e). But unlike in the d-examples, the \bar{A} -chain in the lower clause is not anaphoric to the one in the matrix clause: the head of the \bar{A} -chain in the subordinate clause is in fact the ‘real’ *wh*-phrase; the element heading the matrix \bar{A} -chain merely ‘foreshadows’ it, and ensures that it has matrix scope. While the *wh*-element in the lower clause is an anaphoric topic in (60), it is the focus in (61) (cf. (54)).

(61) $[wh_{SM} \ V+INDEF \ t_{wh} \ [_{CP} \bullet \ wh_{Focus} \ \dots \ t_{wh} \ \dots]]$

Now that we have the landscape laid out, we can ask an important question — one which you are probably dying to get answered by now. What makes Hungarian so unusually rich in its options for constructing long \bar{A} -dependencies? A number of things team up to make this possible.

A very important one is the fact that Hungarian, unlike, say, English, allows accusative case to be assigned to a proleptic object in the SpecVP position of a clause whose verbal head selects a clausal complement. (English does have *I hate it that S*, but here the CP is factive, hence not V's complement.) We see this in (62), but we also see it in the c–, d– and e–examples, each of which exploits the SpecVP position to accommodate the trace for the *wh*-operator in the matrix clause.

(62) azt gondolom hogy szép idő van
 PRON.ACC think.1SG.DEF that nice weather is
 ‘I think it is nice weather’

A second ingredient that gives Hungarian so many ways of building long \bar{A} -dependencies that look very much like straight filler–gap dependencies is the fact that Hungarian is quite liberal in the realm of pro-drop. As we have seen, the fact that no pro-drop is allowed in the language for third-person objects that are plural actually helps us confirm that *pro* plays a key role in the syntax of some of the constructions in (4) (in particular, in the b– and c–sentences).

The third thing that contributes to giving Hungarian such a rich palette of options is the fact that it very actively exploits the topic and focus functions of its sentence constituents and encodes these functions in the left periphery. These three things taken together give Hungarian an exceptionally wide array of strategies for forming long \bar{A} -dependencies.

Knowing all this, one may find it surprising that one type of \bar{A} -dependency is conspicuously absent from the tools for forming long *wh*-dependencies in Hungarian: as I pointed out in the discussion of the d–pattern, null operator movement of the type familiar from *tough*-movement and parasitic gap constructions does not participate in the spectrum of long \bar{A} -dependencies in (3) and (4). This is not a radical impossibility: it will turn out in chapter 5 that null operator dependencies do sometimes form an integral part of the syntax of long *wh*-fronting constructions; they just cannot be employed in the relatively simple constructions under discussion here (i.e., the ones in (3) and (4)). In the grand scheme of things, the use of the null operator (i.e., PRO) is a last resort, an option chosen only when no other strategy could deliver a grammatical result. For the \bar{A} -dependencies in (3) and (4), there is an abundance of strategies available that do not involve the postulation of a null operator (PRO). So null operator movement never comes into the picture.

4.2.4 Appendix: Further issues in the syntax of Hungarian long \bar{A} -dependencies

Now that we have a complete picture of the spectrum of Hungarian long \bar{A} -dependencies presented by the examples in (3) and (4), this appendix will turn to some further details concerning accusative prolepsis constructions, to round things out. This appendix, as its name suggests, provides additional information that is not strictly needed for a proper understanding of the general topic addressed in this chapter. It can readily be skipped, therefore. Readers who have had their fill of Hungarian long \bar{A} -dependencies by now should feel free to move on to section 4.3.

4.2.4.1 A specificity effect with a parallel in French

Recall from the discussion of (3c) and (3d) above that the *wh*-operator *hány lányt* ‘how many girl.ACC’ binds a gap in the matrix clause, in the VP-specifier position for proleptic objects. This position, outside the complement domain the verb, gives the *wh*-operator a specific interpretation. We saw evidence for this specificity effect in the ban on the use of *összesen* ‘in total’:

- (63) **összesen* *hány lány-t gondol-sz,* *hogy jön(nek) a buliba?*
 in.total how.many girl-ACC think-2SG.INDEF that come.3(PL) the party.to

Let us study this specificity effect in a little more detail now.

In the structure of (3c) and (3d), the *wh*-operator binds a trace in the specifier position of VP (as the complement-of-V position is occupied by the subordinate clause). This is depicted in (64).

- (64) $[_{\text{FOCP}} \textit{hány lány}_i \textit{ [F [}_{\text{TP}} \dots [_{\text{VP}} t_i \textit{ [}_{\text{V}'} \textit{ V [}_{\text{CP}} \bullet \dots]]]]]]$

In the Romance syntax literature, this position has been reserved, at least since Kayne (1989), for objects agreeing with a past-participle. In light of this, it is interesting to consider the French examples in (65) in tandem with the Hungarian c- and d-constructions.

- (65) a. *combien de fautes a-t-il fait?* (French)
 how.many of mistakes has-he made(F.PL)
 b. *combien de fautes a-t-il fait-es?*
 how.many of mistakes has-he made-F.PL
 ‘how many mistakes has he made?’

Obenauer (1994) and several scholars since have studied the semantic effects associated with the apparently optional past-participle agreement in (65) in detail. Obenauer points out that while optional on the surface, the presence of past-participle agreement in (65b) has a clear semantic effect: *combien de fautes* is necessarily interpreted specifically when there is past-participle agreement (*faites*). This follows if, in order to control past-participle agreement, the *wh*-phrase must bind a trace in SpecVP, which is the position for specific/presuppositional objects. The structure in (66) is directly parallel to that of the c- and d-examples from Hungarian, given in (64). The specificity parallel between these various sentences presents a reassuring convergence of facts and analyses.

- (66) $[_{\text{CP}} \textit{combien de fautes}_i \textit{ [C [}_{\text{TP}} \dots [_{\text{VP}} t_i \textit{ [}_{\text{V}'} \textit{ V } \dots]]]]]]$

4.2.4.2 Inner islands

The *wh*-quantifier *combien* can famously participate in ‘quantification-at-a-distance’, creating overtly split quantification constructions of the type in (67a). When a *wh*-operator is overtly severed from its restriction, we find robust sensitivity to the presence or absence of sentential negation between the quantifier and the restriction: (67b) contrasts sharply with (67a).

- (67) a. combien a-t-il fait de fautes? (French)
 how.many has-he made of mistakes
 b. *combien n'a-t-il pas fait de fautes?
 how.many NEG.has-he not made of mistakes

We see this also in the Dutch/German *wat voor/was für*-split construction, as illustrated in (68):

- (68) a. wat heeft hij voor fouten gemaakt? (Dutch)
 what has he for mistakes made
 ‘what kinds of mistakes has he made?’
 b. *wat heeft hij niet voor fouten gemaakt?
 what has he not for mistakes made

Split quantification generally resists the intervention of a negation, for reasons discussed in detail in Honcoop (1998) (see also chapter 4). The significance of this fact in connection with the Hungarian c- and d- constructions is that upstairs-agreeing, ‘case-switching’ \bar{A} -dependencies in this languages turn out to exhibit an ‘inner island’ effect as well.¹⁹

Consider the pair of examples in (69a) and (69b). Of these, the former features upstairs definite agreement, as in (3a), while the latter has upstairs *indefinite* agreement and ‘case switch’, like (3c,d). Though the judgements are complex and call for a more systematic investigation, there are speakers who accept (69b) without matrix negation but reject it outright (except on a strongly echoic reading) with negation included; and there are also speakers who would normally prefer upstairs indefinite agreement *cum* ‘case switch’ to the a-pattern, but who reverse their preference in the presence of a sentential negation in the matrix clause — that is, for these speakers, (69b) with negation included is worse than (69a), even though without negation (69b) would be better for them than (69a).

- (69) a. ki (nem) gondolod, hogy el fog menni?
 who(NOM) not think.2SG.DEF that PV will go
 b. kit (?*nem) gondolsz, hogy el fog menni?
 who.ACC not think.2SG.INDEF that PV will go
 ‘who do(n’t) you think will go away?’

Speakers for whom (69b) is degraded when negation is included in it bring evidence for an inner island effect in (3c,d). This follows straightforwardly if in the LF representation of c- and d-type constructions in Hungarian, the restriction of the *wh*-operator must be interpreted at SpecVP — in line with the specificity/presuppositionality of the *wh*-phrase in c- and d-sentences. The quantifier is in the left periphery but its restriction is interpreted in the specifier position for objects:

- (70) $[_{FocP} \text{hány}_i [F [_{TP} \dots [_{VP} [t_i \text{lányt}] [_{V'} V [_{CP} \bullet \dots]]]]]]$

19 The discussion in the following paragraph is based on Den Dikken (2009b), where the original empirical observation is made.

So at LF, c/d-type constructions are very much like overtly split quantification constructions such as French (67b) or Dutch (68b). The fact that the *wh*-phrase of c/d-sentences behaves like split quantifiers with regard to sensitivity to sentential negation suggests that, at the relevant level of representation (i.e., at LF), it in fact IS a split quantifier.²⁰

4.2.4.3 A word-order restriction in multiple *wh*-fronting

With this in mind, let us now examine an interesting word-order restriction on long subject dependencies of types (3c) and (3d) in the realm of multiple *wh*-fronting.

The relative order of two fronted *wh*-constituents in Hungarian is generally free: (71a) and (71b) are both fine (though they have different interpretations, as É. Kiss 1993 shows in detail).

- (71) a. ki hány lányt hívott meg?
 who how.many girl.ACC invited.INDEF PV
 b. hány lányt ki hívott meg?
 how.many girl.ACC who invited.INDEF PV
 both: ‘who invited how many girls?’

But in (72) we see that long subject dependencies with ‘case switch’ (i.e., an accusative in the matrix clause corresponding to the subject of the lower clause, as in (3c,d)) present us with an ordering restriction on the two fronted *wh*’s: ‘case-switched’ *hány lányt* must be placed to the immediate left of the finite verb.²¹

- (72) a. ki hány lányt gondol, hogy jön(%-nek)?
 who how.many girl.ACC think.INDEF that come-3PL
 b. ^{??}hány lányt ki gondol, hogy jön(-nek)?
 how.many girl.ACC who think.INDEF that come-3PL

The ‘??’ diacritic is used here in exactly the way that Beck (1996) used it in her discussion of intervention effects: as Beck (1996:3, fn. 3) puts it, ‘[t]he ‘??’ means that the data are incomprehensible (uninterpretable) rather than simply ungrammatical’. The fact that (72b) may sound at first like a fine sentence but upon inspection turns out to be uninterpretable strongly suggests that the problem with (72b) manifests itself in the LF representation.

In the following paragraphs I will try to relate the pattern in (72) to the fact that in *wh*-scope marking construction, we also find a word-order restriction in multiple *wh*-fronting contexts: the scope-marker *mit* insists on being immediately next to the finite verb, as shown in (73).

20 The present analysis makes significant strides in relating the inner island effect in (69b) to similar effects found in split quantification constructions. But it does not carry over immediately to the complex conditions under which ‘plain’ scope-marking constructions exhibit sensitivity to sentential negation (as discussed in Horvath 1997). The empirical material here is complex. I cannot do it justice here.

21 As always, there is speaker variation regarding the acceptability of notional plural agreement on the embedded verb. But as far as I have been able to determine, there is no variation with respect to the contrast between the a- and b-examples in (72). Thanks to Márton Kucsera and Krisztina Szécsényi (p.c.) for their help with these sentences.

- (73) a. ki mit gondol, hogy hány lány jön?
 who what.ACC think.INDEF that how.many girl(NOM) come
 b. *mit ki gondol, hogy hány lány jön?
 what.ACC who think.INDEF that how.many girl(NOM) come

For the scope-marker, it is unsurprising that it wants to be in the focus position (i.e., the position immediately adjacent to the finite verb). In a Hungarian multiple *wh*-fronting construction, the *wh*-expression in the position to the immediate left of the finite verb is focused while all *wh*'s preceding it are in topic-like positions suitable only for universal quantifiers. The fact that the scope-marker *mit* completely lack of descriptive content makes it unsuitable for the topic-like left periphery position occupied by non-verb-adjacent *wh*-elements in Hungarian. But for *hány lányt* in (72) it is not immediately obvious why it should resist placement in a non-focus position. What is wrong with (72b)?

Recall, once again, that in (3c) and (3d) the *wh*-operator *hány lányt* ‘how many girl.ACC’ binds a gap in SpecVP, the position for specific objects. At LF, the syntax of these constructions features just the *wh*-element in the operator position; the restriction is mapped onto the trace position. Now consider what this entails for (72), featuring an additional *wh*-element in the matrix left periphery. If we place the proleptic *wh* in the focus position and the additional *wh* in a pre-focus position, similar to that of non-*wh* topics (and annotated as ‘TopP’, accordingly), the LF representation, given in (74a), is fully interpretable. But if instead we place the additional *wh* in the focus slot, that forces us to put the proleptic *wh* in the topic field. The result of this derivation crashes at LF: (74b) cannot interpret the ‘bare’ *wh*-operator (*hány*) in the topic position; placement in the pre-focus topic position requires the presence of presuppositional material in that position, and the *wh*-operator *hány* ‘how many’ is obviously not presuppositional; its restriction, *lány* ‘girl’, most certainly does have descriptive content and would serve fine as a topic, but it must be linked to the SpecVP position in the LF representation in order that the specificity of the *wh*-expression can be marked appropriately, on the restriction.

- (74) a. $[_{\text{TopP}} ki [_{\text{Top}} [_{\text{FocP}} hány_i [_{\text{Foc}} [_{\text{TP}} \dots [_{\text{VP}} [t_i lányt] [V [_{\text{CP}} \bullet \dots]]]]]]]]]$
 b. $*[_{\text{TopP}} hány_i [_{\text{Top}} [_{\text{FocP}} ki [_{\text{Foc}} [_{\text{TP}} \dots [_{\text{VP}} [t_i lányt] [V [_{\text{CP}} \bullet \dots]]]]]]]]]$

The LF representation of the matrix portion of (3c) and (3d) wants the *wh*-phrase to be ‘split’ into a ‘bare’ operator upstairs and its restriction in SpecVP. But it expressly does not want the former to be in a position in which it cannot be interpreted. Because of its bareness, the *wh*-operator by itself is only welcome in the focus position. If the matrix *wh*-phrase is in the pre-focus position in overt syntax, this is where the ‘bare’ *wh*-operator will be at LF. As this is not a position in which it can be interpreted, the LF of (72b), given in (74b), is rejected, in exactly the same way as that of (73b) is.

4.3 *Wh*-copying as concordial scope marking

In the *wh*-scope marking construction, illustrated for Hungarian in (3e) and (4e) (repeated here as (75a,b)), the ‘real’ *wh*-phrase shows up in the embedded clause, which is linked to a bare *wh*-operator in the matrix clause.

- (75) a. mi-t gondol-sz, hogy hány lány jön(*-nek) a buliba?
 what-ACC think-2SG.INDEF that how.many girl(NOM) come-*3PL the party.to
 b. mi-t gondol-sz, hogy ki-t hívott meg a buliba?
 what-ACC think-2SG.INDEF that who-ACC invited.3SG.INDEF PV the party.to

Though impossible in standard Hungarian (see below for discussion of examples found in dialects of Hungarian), there are languages in which the *wh*-operator in the matrix clause can take on a shape that makes it look at least in some respects like the *wh*-phrase in the subordinate clause. Thus, consider the following German examples side by side:

- (76) a. was denkst du wen von den Studenten man einladen sollte?
 what think you who of the students one invite should
 b. wen denkst du wen von den Studenten man einladen sollte?
 who think you who of the students one invite should
 ‘which of the students do you think one should invite?’ (German)

The example in (76b) repeats (5b), from the introduction to this chapter. It represents what is often called the ‘*wh*-copying construction’, which occurs, in many dialects of Dutch and German, alongside the ‘plain’ *wh*-scope marking construction in (76a). The question under discussion in the present section is how the ‘*wh*-copying’ construction in (76b) should be analysed.

4.3.1 Not literal copying

Simple examples of ‘*wh*-copying’ such as (77), from German, seem to lend themselves easily to a bottom-up derivational account according to which *wer* starts out in the subject position of the lower clause, then raises to the SpecCP position of the embedded clause, and subsequently gets launched into the SpecCP slot of the matrix clause. On the assumption that a copy of the moved *wh*-operator is left behind at the edge of the embedded clause, and that this copy can, in German, be given a phonological matrix along with the *wh*-copy in the matrix SpecCP, the surface output in (77) arises.

- (77) wer glaubst du wer kommen wird? (= (5a)) (German)
 who think you who come will
 ‘who do you think will come?’

But examples of the type in (76b) and (78) (both taken from Fanselow & Ćavar 2001, cited in Bruening 2006) should give us pause.

- (78) *wessen Studenten denkst du wessen Studenten man einladen sollte?
 whose students think you whose students one invite should

If one could literally spell the intermediate copy (or copies) in the chain of successive-cyclic *wh*-movement out at PF, there is no obvious reason why that copy can be partially the same as the upstairs *wh*-phrase, as in (76b), and may never be a full replica of a complex *wh*-constituent upstairs, as in (78).

Note also that the multiple copy spell-out analysis of *wh*-copy constructions has never been straightforward. One of the questions it raises is how we can allow multiple members of the same chain to be spelled out simultaneously, in a syntactic configuration in which the higher copy asymmetrically c-command the lower one. Barbiers *et al.* (2009) suggest a solution to this puzzle that capitalises on the idea that the copy of the *wh*-constituent in the lower SpecCP undergoes *morphological merger* with C, and is thereby rendered ‘invisible’ to the linearisation algorithm (Kayne’s 1994 Linear Correspondence Axiom). But plainly, the lower *wh*-element IS linearised, and so are subparts of words. Kayne (1994) himself clearly deemed his LCA operative below the level of the word (see his discussion of clitic clusters, for instance). If the LCA ‘cannot look into the structure of words’ (Barbiers *et al.* 2009:15), a separate linearisation algorithm is needed to order subparts of words, which will undermine the LCA, potentially to the point of becoming redundant. Absent an explanatory perspective on the way a *wh* that undergoes morphological merger with C is rendered ‘invisible for the LCA’ but does get pronounced and linearised, it seems to me that a multiple copy spell-out analysis of *wh*-copying cannot be pursued.

So for both descriptive and theoretical reasons, the term ‘*wh*-copying’ is really a misnomer. I will continue to use it because of its instant recognisability, but put it in inverted commas throughout.

4.3.2 *Concordial scope marking*

What I would like to pursue here as the analysis of ‘*wh*-copying’ is a proposal that takes quite literally what I said about (76) above: the *wh*-operator in the matrix clause takes on a shape that makes it look in some respects like the *wh*-phrase in the subordinate clause. More precisely, what I would like to suggest (following Den Dikken 2009b) is that ‘*wh*-copying’ is a variant of the *wh*-scope marking construction, with the scope marker entertaining a *feature concord* relationship with the ‘real’ *wh*-phrase.²²

Thus, a sentence like (76b) involves concord between the *wh*-constituent downstairs (*wen von den Studenten*) and the *wh*-scope marker upstairs for just the φ - and case-features of the former — features for which concord is known to exist in natural language (cf. the concord relation between adjectives and nouns for number and gender, and the case concord between subjects and predicate nominals). This is depicted in (79), where ‘SM+FF _{φ /case}’ represents a scope marker that entertains a concord relationship with the *wh*-DP downstairs for the latter’s φ - and case-features.

22 In Den Dikken (2009b), I presented an analysis of d-sentences in (3) and (4) that took the silent element \bar{A} -binding the variable in the embedded clause to be a ‘full-concordial scope marker’: the abstract scope-marker in the matrix clause acquires *all* of the features of the ‘real’ *wh*-constituent in the lower clause, via concord, which has two effects: (a) the scope-marker comes to look very much like the ‘real’ *wh*-constituent; and (b) because the ‘concordial scope marker’ is a featural superset of the ‘real’ *wh*-constituent in the embedded clause, the latter must remain silent. The nature of the concord process that is supposed to be responsible for the fact that the scope marker comes to walk and talk almost exactly like the ‘real’ *wh*-constituent remains rather unclear in Den Dikken’s (2009b) analysis, however. Abels (2012:56) rightly chastises the ‘full-concordial scope marking’ proposal for this lack of clarity or similarity with known cases of concord. The φ -feature concord that the main-text analysis of ‘*wh*-copying’ resorts to is much less mysterious: in fact, it is well known that two terms can engage in a concordial relationship confined to φ -features.

One might ask how ‘ φ -concordial scope marking’ fits into a top-down approach to the building of syntactic structure and the filler–gap dependencies established in them. There really is no conflict here: all we need to do is endow the scope marker with a set of φ -features at the outset, and make sure that these φ -features are matched, under concord, with those of the ‘real’ *wh*-phrase downstairs.

(79) [SM+FF_{φ/case}=*wen* ... [CP [DP *wen von den Studenten*] ...]]

The scope marker ‘inherits’ from the *wh*-constituent the fact that it is singular, human and accusative, which (in conjunction with the fact that it is a *wh*-element) leads the scope marker to be spelled out as *wen* — and crucially not as *wen von den Studenten* ‘which of the students’.

In a situation in which the upstairs scope marker and the *wh*-constituent in the embedded clause are engaged in a concord relation for the φ- and case-features of the ‘real’ *wh*-constituent, we expect that (a) the upstairs *wh* is always a bare *wh*-word (because it inherits nothing but the φ- and case-features of the ‘real’ *wh*), and (b) the concordial scope marker and the downstairs *wh*-constituent are both allowed to be pronounced, as the two *wh*-elements are featurally non-identical, and their identical φ- and case-features are not in a c-command relation. This description characterises the ‘*wh*-copying’ construction very well.

We expect to find cases in which the two *wh*-elements entertaining a concord relation are both bare (and then identical), as in (77), and cases in which the *wh*-element in the higher clause is bare but shares its φ- and case-features with a complex *wh*-expression in the lower clause, as in (76b). What we do not expect to find is ‘*wh*-copying’ constructions in which the *wh*-expressions in both clauses are complex: and indeed (78) is not attested. And we also do not expect to come across sentences in which there is a *wh*-element in each clause, one of them bare, but with the bare *wh* located in the embedded clause rather than in the matrix.

Observationally speaking, this last expectation is not borne out. This leads me to a brief interlude on a different type of ‘*wh*-copying’ pattern — one that cannot be treated in terms of concordial scope marking, and is more likely to involve a reduced cleft.

4.3.3 A different kind of ‘*wh*-copying’

‘*Wh*-copying’ constructions in which the larger *wh*-expression appears in the embedded clause and the simple *wh* is in the matrix clause (as in (76b) and (80a)) are attested in colloquial varieties of adult Dutch (see also Koster 2009),²³ and one also finds them in learners of English as a second language (interestingly, even in speakers whose native languages do not have anything like *wh*-copying; see Gutiérrez 2005). For these we now have an account in terms of concordial scope marking. But ‘*wh*-copying’ constructions of the inverse type, with the larger *wh* upstairs and the simple one downstairs, as in (80b), have also been reported. Thus, Van Kampen (1997, 2009) points out that they appear frequently in the speech of children acquiring Dutch as a first language; and similar such utterances have been recorded for child English (Thornton 1990, 1995) and colloquial varieties of adult Dutch (see Barbiers *et al.* 2009).

23 Here are some attested examples, culled from the internet using Google:

- | | | | | | | | | | | |
|-----|----|-----|-------|-----|-------|-----------------|-------|----------|------------|--------------------|
| (i) | a. | wie | denk | je | welk | type | je | bent? | | (colloquial Dutch) |
| | | who | think | you | which | type | you | are | | |
| | b. | wie | denk | je | welk | team | er | kampioen | wordt? | |
| | | who | think | you | which | team | there | champion | becomes | |
| | c. | wie | denk | je | welke | topsportvrouw | dit | is? | | |
| | | who | think | you | which | top-sport-woman | this | is | | |
| | d. | wie | denk | je | welke | stellen | nog | bij | elkaar | zijn? |
| | | who | think | you | which | couples | still | with | each-other | are |

- (80) a. wie denk je welke jongen ik gezien heb? (colloquial Dutch)
 who think you which boy I seen have
 b. welke jongen denk je wie ik gezien heb? (colloquial and child Dutch)
 which boy think you who I seen have

I agree with Barbiers *et al.* (2009) that (80b) must be given an analysis that is fundamentally different from the one assigned to (76b) and (80a). A reduced cleft analysis, perhaps along the lines suggested recently by Koster (2009), seems plausible, with either (81a) (construing *denk je* as a parenthetical) or, more likely, (81b) as the underlier.²⁴

- (81) a. welke jongen ~~is het~~ denk je wie ik gezien heb? (Dutch)
 which boy is it think you I seen have
 b. welke jongen denk je ~~dat het is~~ wie ik gezien heb?
 which boy think you that it is who I seen have

If the reduced cleft analysis can be upheld, (80b) does not present a new type of *wh*-dependency: it reduces to the syntax of clefts with a *wh*-phrase as the focus. There is a large literature on cleft constructions (see Reeve 2012 for recent discussion, and references to the earlier literature). I refer the reader to that literature for details on the syntax of clefts, which space does not allow me to go into here (see Den Dikken 2013 for my perspective on clefts).

4.3.4 On the cross-linguistic distribution of concordial scope marking

Returning now to ‘*wh*-copying’ constructions of the type in (76b), (77) and (80a), we expect — in light of the concordial scope-marking hypothesis — that such constructions should be grammatical only in languages that also have ‘plain’ *wh*-scope marking. This seems to be the case. But it does not follow that all languages that have *wh*-scope marking should also be able to accommodate ‘*wh*-copying’.

There are two factors which play a role in making ‘*wh*-copying’ (i.e., ϕ -concordial scope marking) more restricted in its distribution than ‘plain’ *wh*-scope marking. The first is an obvious one: some languages may not have recourse to concord relations at all. The other factor is structural in nature, and concerns the structural environment in which concord between a *wh*-scope marker in the matrix clause and a *wh*-phrase at the left edge of the embedded clause is possible.

Although concord relations are much less well understood than agreement relations, and though it probably cannot be maintained that all concord relations are regulated by Agree (though see Baker 2008 for a proposal that treats both agreement and concord in terms of Agree), it seems entirely reasonable to hypothesise, by way of a baseline, that concord relations are, like agreement relations, subject to a locality condition. For agreement, which is governed by Agree, we know that it cannot be established across the boundaries of an opaque domain, as defined in (82) (see section 3.3 for discussion):

24 The non-elliptical versions of (81a,b) are acceptable in colloquial speech. In the standard language, the clefts in (81) would involve the *d*-word *die* instead of *wie* in the sentence-final relative clause. It is interesting to note that the variant of (80b) featuring *die* instead of *wie* is also attested, possibly lending support for the reduced cleft analysis.

- (82) *opaque domain*
 in $[\alpha \dots \pi \dots [\Delta \dots \beta \dots]]$, Δ is an *opaque domain* for a relation between α and β iff:
 (a) Δ dominates β , and
 (b) $\Delta \neq$ a goal γ in an Agree-relation with an asymmetrically c-commanding probe π

Let us assume, then, that concord relations are likewise confined to members of the same domain.

Hungarian is an interesting language to look at in this connection. We have seen in our discussion in section 4.2 that in Hungarian *wh*-scope marking constructions, the matrix verb engages in an Agree relation with the scope marker: in (75), repeated here, we see indefinite agreement on the matrix verb — not the definite inflection that we would have got if the verb entertained an Agree relation with its finite clausal complement.

- (75) a. mi-t gondol-sz, hogy hány lány jön(*-nek) a buliba?
 what-ACC think-2SG.INDEF that how.many girl(NOM) come-*3PL the party.to
 b. mi-t gondol-sz, hogy ki-t hívott meg a buliba?
 what-ACC think-2SG.INDEF that who-ACC invited.3SG.INDEF PV the party.to

Because of the fact that the embedded clause in (75) is not an Agree-goal to the matrix verb, we expect it to be an opaque domain. It should logically be impossible, therefore, for Hungarian to have ‘*wh*-copying’ — or, in current terms, concordial scope marking. And indeed, while extraordinarily well-stocked in types of \bar{A} -dependencies, standard Hungarian lacks precisely the ‘*wh*-copying’ construction. This is a straightforward and accurate prediction which I would like to present as an argument in favour of the approach taken here.

As I mentioned on a few occasions in the foregoing discussion, however, there appear to be varieties of Hungarian where something resembling ‘*wh*-copying’ is possible. The examples in (83) were provided to me by Balázs Surányi (p.c.):

- (83) a. %kit mondtál, hogy ki jött be? (Hungarian)
 who.ACC said.2SG.INDEF that who(NOM) came in
 b. %kit mondtál, hogy kit hívtál meg?
 who.ACC said.2SG.INDEF that who.ACC invited.2SG.INDEF PV

The verb in the matrix clause is clearly marked with indefinite inflection, which irrevocably seals the embedded CP as an opaque domain. These could not be cases of concordial scope marking, therefore, if ϕ -feature concord observes absolute locality, as I presume it does. What seems to confirm that the Hungarian examples do not involve German-style ‘*wh*-copying’ is that, according to Surányi, the Hungarian examples are marginally possible with complex *wh*’s in both clauses — very much unlike what we see in (78) — and it is even possible to use a focused non-*wh* pronoun as the associate of the matrix *wh*:²⁵

25 The matrix verb in (84) bears the definite conjugation because *melyik diákot* ‘which student.ACC’ is definite. Even in sentences of the type in (84), therefore, the verb is not in an Agree relation with the embedded clause.

- (84) ^{?(?)}melyik diákot szeretném, hogy ŐT rúgjuk ki? (Hungarian)
 which student.ACC would.like.2SG.DEF that HIM throw.1PL.DEF out

Though I do not have a full-fledged analysis of (83) and (84) to offer at this time, what I suspect we are dealing with here is a variant of contrastive left dislocation, with the *wh*-constituent in the matrix clause as a contrastive topic and the *wh*-element or pronoun in the embedded clause as a focus. Hungarian has contrastive left dislocation, both clause-internally and across a CP boundary. The pronominal associate of the left-dislocated constituent is usually a form of the demonstrative *az*; but it can also be a personal pronoun. Example, modelled on (84), are given in (85a,b).

- (85) a. azt a diákot szeretném, hogy azt kirúgjuk (Hungarian)
 that student.ACC would.like.1SG.DEF that DEM out.throw.1PL.DEF
 b. azt a diákot szeretném, hogy őt kirúgjuk
 that student.ACC would.like.1SG.DEF that him out.throw.1PL.DEF
 ‘that student, I wish they would throw him out’

From (85b) it is but a small step to (84): the contrastive topic is turned into a *wh*-phrase, and its associate in the embedded clause is focused (which brings about ‘preverb–verb inversion’). If the associate pronoun is itself turned into a *wh*-form, what we get is constructions of the type in (83).

I present these thoughts here as a direction for more detailed research on the syntax of apparent ‘*wh*-copying’ constructions in varieties of Hungarian. What matters for me here is no so much the precise ingredients of this analytical sketch, but rather the conclusion that it is likely that examples of the type in (83) and (84) are much more directly aligned with contrastive left dislocation constructions than with what the literature commonly refers to as ‘*wh*-copying’. The phenomenon of ‘*wh*-copying’ proper, analysed here in terms of concordial scope marking, cannot — and arguably does not — occur in Hungarian because in Hungarian scope marking constructions the embedded clause is systematically opaque, which precludes the structural relationship between the scope marker in the matrix clause and the *wh*-element in the embedded clause that would facilitate ϕ -feature concord between them.

4.4 *Wh*-control²⁶

In the introduction to this chapter, I pointed out that there are environments in English in which long-distance *wh*-dependencies involving the subject of a finite clause do not ‘just’ show a sensitivity to the presence of a lexical complementiser (as in the familiar ‘*that*-trace effect’) but fail categorically, even with the complementiser omitted. Stowell (1981:410–13) was the first, to my knowledge, to observe that with verbs of the *convince* class, long-distance *wh*-dependencies involving the object are well-formed but long subject dependencies fail regardless of whether *that* is pronounced or not:

- (86) a. what did they convince/persuade/remind him that he should do *ec*? (= (6))
 b. *who did they convince/persuade/remind him (that) *ec* should do it?

26 I use this term to make reference to ‘control by a *wh*-element’, on the analogy of ‘subject control’ and ‘object control’, which stand for ‘control by a subject’ and ‘control by an object’, respectively. Confusingly, some of the literature has used the term ‘*wh*-control’ to refer to ‘control into a *wh*-clause’ (as in *I don’t know what to do*).

This pattern is not shared by verbs of the *show*-type, which like *convince*-type verbs feature both a noun phrase and a clause in their complement domain: as Stowell notes, (87b) is grammatical provided that *that* is omitted, exhibiting the familiar ‘*that*-trace effect’.

- (87) a. what did they show/show him that he should do *e*?
 b. who did they show/tell him (*that) *e* should do it?

We are not dealing, therefore, with some sort of ‘defective intervention’ effect: it is not the presence of a nominal object in the matrix clause *per se* that renders the long subject dependency in (86b) ungrammatical.²⁷

The unusual behaviour of *convince*-class verbs in the realm of long subject dependencies is not specific to English. In Dutch as well, *convince*-type verbs allow object extraction but resist long subject dependencies:

- (88) a. wat hebben ze hem overtuigd dat hij moet doen? (Dutch)
 what have they him convinced that he must do
 ‘what did they convince him that he should do?’
 b. *wie hebben ze hem overtuigd dat het moet doen?
 who have they him convinced that it must do

In Hungarian *convince*-class constructions (with *meggyőz* ‘convince’ as our exemplar), we also find that long subject dependencies are ungrammatical: (89b). But this time around, even the accusative direct object resists *wh*-fronting into the matrix clause: (89a) is impossible, too.

- (89) a. *kit győztek meg (téged), hogy János szeret? (Hungarian)
 who.ACC convinced.3PL.INDEF PV you.ACC that János(NOM) love.3SG.INDEF
 ‘who did they convince you that János loves?’
 b. *ki győztek meg (téged), hogy szereti Marit?
 who(NOM) convinced.3PL.INDEF PV you.ACC that love.3SG.DEF Mari.ACC
 *‘who did they convince you (that) loves Mari?’

As the reader will recall from the discussion of Hungarian long-distance \bar{A} -dependencies in section 4.2, speakers of Hungarian generally disprefer long *wh*/focus fronting to alternative strategies involving *wh*-scope marking or prolepsis. But no speaker of Hungarian would reject examples such as (3a) and (4a) outright: the sentences in (90) fall squarely within the realm of the grammatical.

- (90) a. ki-t gondol-od, hogy meghívott a buliba? (Hungarian)
 who-ACC think-2SG.DEF that PV.invited.3SG.INDEF the party.to
 a. ki gondol-od, hogy jön a buliba?
 who(NOM) think-2SG.DEF that come.3SG.INDEF the party.to

27 I should note that long subject dependencies of the type represented by (86b) are not unattested: as long as the complementiser *that* is omitted, sentences of this type occur in reasonable numbers on the internet. But the relative contrast between (86a) and (86b) and that between (86b) and (87b) both seem real. Bošković & Lasnik (2003) also endorse the judgements that Stowell (1981) reports, both for *tell*-type verbs and for *convince*-type ones.

The fact that both (89a) and (89b) are ungrammatical is thus significant.

What is particularly revealing about the Hungarian examples in (89) and (90), when considered side by side, is that in the latter the matrix verb engages in a definiteness agreement relation with the subordinate clause while in the former it does not. Recall again that in chapter 3 I argued for the definition of ‘opaque domain’ in (82), in which Agree plays the key role.

- (82) *opaque domain*
 in [$\alpha \dots \pi \dots [\Delta \dots \beta \dots]$], Δ is an *opaque domain* for a relation between α and β iff:
 (a) Δ dominates β , and
 (b) $\Delta \neq$ a goal γ in an Agree-relation with an asymmetrically c-commanding probe π

The fact that the embedded CP in (89) is not an Agree-goal leads us to conclude that it is an opaque domain — an absolute island.

For Hungarian, this closes the door on all long-distance filler–gap dependencies crossing the boundaries of the subordinate CP in *convince*-class constructions. But as I will argue at more length in chapter 5, there is an alternative strategy available for object dependencies in English and Dutch *convince*-class constructions — one that represents the *wh*-object as a proleptic object of the matrix clause, controlling a null-operator dependency in the embedded clause:

- (91) [_{CP} *wh*_i *did they convince him* *t*_i [_{CP} *Op*_i ... *ec*_i ...]]

The null-operator dependency in (91) is established entirely within the confines of the subordinate clause, hence unhindered by the opacity of the lower CP. The null operator is an instance of PRO (see chapter 5 for arguments to this effect); and it is in a control relation with the proleptic *wh*-dependency in the matrix clause. This is what I call ‘*wh*-control’ — the \bar{A} -counterpart to the familiar cases of control relations between an argument of a matrix clause and a PRO in a subordinate domain, which, in cases of control into adjuncts, is itself opaque.

- (92) a. they_i tried [PRO_i to convince him of it]
 b. he_i left the room [without/before/after PRO_i being convinced]
 c. he_i left the room [PRO_i convinced]
 d. they_i convinced him_k [PRO_{i/k} drunk]
 e. they_i talked sense into him_k [PRO_{i/*k} drunk]

The ‘*wh*-control’ strategy is not available at all in the case of Hungarian because of the fact that the proleptic object in *convince*-class constructions is an oblique, PP-contained noun phrase — and as we know independently from the fact that, while (92d) is ambiguous between a subject-depictive and an object-depictive reading, (92e) only supports the former interpretation, it is impossible for a PP-contained noun phrase to control PRO. This is why Hungarian (93) is ungrammatical with indefinite inflection on the embedded verb: indefinite agreement signals the presence of a null-operator dependency in the embedded clause (recall parasitic gap examples of the type in (36), above), but since the null operator (i.e., PRO) cannot be successfully controlled by the matrix proleptic *wh*-object, (93) fails without *-i*. When *-i* is present, by contrast, (93) is fine: now the matrix proleptic object is linked to a silent object pronoun (*pro*) in the embedded clause, triggering definite inflection; and as is well known, pronoun–antecedent relations do not require c-command.

- (93) kiről győztek meg (téged), hogy János szeret-*(i)? (Hu.)
 who.DEL convinced.3PL.INDEF PV you.ACC that János(NOM) love.3SG-DEF
 ‘of whom did they convince you that János loves her?’

The ‘*wh*-control’ option is universally unavailable in the case of subjects because of restrictions on null-operator dependencies, which will be discussed, illustrated for a variety of cases, and derived from the general theory in chapter 5. This will provide us with an explanation for the ill-formedness of the subject extraction examples in (86b) and (88b).

I will develop the syntax of *wh*-control in much more detail in section 5.3.14, which will also lead us back in a surprising way to the discussion of Pesetskian ‘cascade structures’ in chapter 2 — for it will turn out that *convince*-class constructions actually have a ‘cascade structure’ that serves both c-command and constituency. For the purposes of the discussion in the present section, however, what is important to take away from this brief preview of the syntax of *convince*-class constructions is that Universal Grammar exploits the control relation in its treatment of long-distance \bar{A} -dependencies (just as it does in its treatment of long A-dependencies). This is just as it should be: if the control relation exists as an ingredient of the universal syntactic toolkit, as I believe it must, we expect it to be exploited wherever possible. In an approach in which control is not reduced to movement or binding but exists as a syntactic relation in its own right, *wh*-control (or \bar{A} -control, more generally) ought to exist alongside the familiar cases of A-control.

4.5 Conclusion

With *wh*-control now in place, we have completed the typology of long-distance \bar{A} -dependencies (concentrating primarily on *wh*-dependencies; but the pattern should carry over to focus and topic dependencies), featuring the following ingredients:

- (94) a. direct filler–gap dependency (= (7))
 b. scope marking
 c. concordial scope marking (a.k.a. ‘*wh*-copying’)
 d. prolepsis *cum* resumption
 e. prolepsis *cum* topic drop
 f. control

All of the members of this typology fit neatly into the top-down approach to syntactic structure building pursued in this book. They are all empirically supported, and occur in the environments in which we would expect them to be able to occur. The availability of each type of long \bar{A} -dependency is constrained by syntactic properties of individual constructions (e.g., by the question of whether a subordinate clause is transparent or opaque), and by parametric properties of individual languages or varieties (thus, English does not have (94e) because it does not have topic drop; and it cannot have accusative prolepsis in general because it does not have the Hungarian-style pronominal ‘herald’ of a bridge-verb complement clause). The distribution of the various types of long-distance \bar{A} -dependencies thus does not need to be stipulated. It is as rich as it needs to be, given the facts of natural language; and it is as rich as it can be, given the way that Universal Grammar is equipped.

4.6 The case of Dinka long-distance \bar{A} -fronting²⁸

None of the members of the typology of \bar{A} -dependencies in (94) involves ‘successive-cyclic movement’ — i.e., the generation of a chain between a filler and a gap featuring one or more subsidiary gaps in intermediate positions along the way. The top-down approach to the building of syntactic structures and the filler–gap dependencies inside them does, however, deliver a close relative to successive-cyclic movement. In the syntax of the d–sentences in (3) and (4), summarised given in (60) (repeated below), there is a silent *wh*-operator, identical in form and reference to the matrix *wh* (minus its structural case feature), in the left periphery of the embedded clause, and it is silent *wh* that binds the variable in the argument position in the embedded clause; the overt *wh*-operator upstairs is a proleptic object.

$$(60) \quad [wh_{ACC} \text{ V+INDEF } t_{wh} \text{ [CP} \bullet \text{ } wh_{Topic} \dots t_{wh} \dots]]$$

Representationally, (60) looks a lot like the output of a bottom-up successive-cyclic movement derivation. But it comes about very differently — and because of this, it behaves very differently: there are no *wh*-island effects, and the two chains have different information-structural profiles. In constructions of the type in (3a) and (4a), where the *wh*-filler in the matrix clause *itself* binds the variable in the embedded clause, we have seen that there is no need for an intermediate copy in the left periphery of the embedded clause: the subordinate clause is transparent for the formation of a direct filler–gap dependency across its borders thanks to the fact that the matrix verb takes the CP as its Agree-goal (Rackowski & Richards 2005). When the CP is not an Agree-goal, as in (60) (as marked by ‘●’), it is opaque: no filler–gap dependency can be built straddling its boundaries.

Van Urk & Richards (2015) have recently made a case for a combination of Rackowski/Richards-style CP-agreement and densely successive-cyclic movement. Their argument is based on the facts of Dinka, a Nilotic language spoken by about 3 million people in South Sudan. They argue that Dinka provides not only support for the idea that long \bar{A} -dependencies across CP require CP to be an Agree-goal, but also ‘particularly compelling evidence for the idea that long-distance dependencies involve intermediate movement steps through the edge of every verb phrase and every clause’ (p. 152). On one’s first pass through their paper, the Dinka facts do indeed seem to present a striking case for densely successive-cyclic movement, via the edge of every ν P and CP along the movement path. In this brief section, I will show that upon inspection, it turns out that the Dinka data are at best merely *compatible* with a successive-cyclic movement analysis of long-distance \bar{A} -dependencies: nothing about these data suggests that they could *only* be successfully treated if we adopted such an analysis. To the contrary, there are ingredients of the Dinka data that suggest that an analysis not involving successive-cyclic movement is more likely to be correct.²⁹

Dinka is a Verb Second language, with V2 generalised beyond the pattern familiar from the Germanic languages, in two respects:

28 This section can safely be skipped by readers not interested in the question of whether there is morphosyntactic support for so-called ‘successive-cyclic movement’.

29 In these remarks, I obviously cannot do full justice to Van Urk & Richards’ (2015) detailed study. I refer the interested reader to the original paper for a wealth of material.

- (95) *Dinka generalised V2*
- a. not only does the root-CP generally require there to be something in its specifier position immediately preceding the verb in C, but *non-root* CPs show this requirement as well
 - b. not only CPs evince the V2 restriction, *vPs* do as well

There are two systematic exceptions to this generalised V2 pattern, and in both of these we find a Verb First (V1) pattern instead:

- (96) *Dinka V1*
- a. in yes/no-questions³⁰
 - b. in constructions in which something is extracted from *vP* and CP

Van Urk & Richards capitalise on (96b), and argue persuasively that in these extraction cases, the V2 restriction IS in fact satisfied in the course of the derivation: the extracted element passes through the left edge of every *vP* and CP along the movement path, and checks the EPP (the filled-specifier requirement) at every step along the way.

The proof of the pudding for this analysis, according to Van Urk & Richards, is in the distribution of the plural clitic *ke*. Van Urk & Richards (2015:sect. 3.2) show that plural extractees (of all stripes) force the presence of the plural clitic *ke* in each *vP* along the path of movement. They suggest that this involves ‘stranding’, similar to quantifier float.³¹ The paradigm in (97) shows this.³²

- (97) a. yeyinga ye ke taak, cii Bol ke ting? (Dinka)
 who.PL HAB.2SG PL think PRF.NS Bol PL see
 ‘who all do you think Bol saw?’
- b. *yeyinga ye taak, cii Bol ke ting?
 who.PL HAB.2SG think PRF.NS Bol PL see
- c. *yeyinga ye ke taak, cii Bol ting?
 who.PL HAB.2SG PL think PRF.NS Bol see
- d. *yeyinga ye taak, cii Bol ting?
 who.PL HAB.2SG think PRF.NS Bol see

If indeed the distribution of *ke* (which is not analysed in much detail by Van Urk & Richards) is that of a floating quantifier, and if floating quantifiers articulate the path of movement (as is customarily assumed in the literature), then (97) tells us that extraction of the plural object proceeds via the edge of every *vP* along the movement path.

30 Just as in Germanic. The parallel with the Indo-European V2 languages is hard to miss here.

31 The facts of West Ulster English discussed by McCloskey (2000) seem at first to be a particularly close match.

32 This paradigm incorporates Van Urk & Richards’ (2015:128) examples (25b) and (26). Note that I am leaving off all of the diacritics that adorn Van Urk & Richards’ examples. Whoever would like to quote these examples should base themselves on the originals, not on the simplified orthography presented here.

What we have here is most certainly not ‘particularly compelling evidence for the idea that long-distance dependencies involve intermediate movement steps through the edge of every verb phrase and every clause’ (Van Urk & Richards 2015:152). The distribution of *ke*, which is supposed to nail the argument, cannot give rise to this strong conclusion.

There are other reasons, too, for being deeply skeptical about Van Urk & Richards’ treatment of the Dinka facts. What should make one particularly suspicious about their account is the fact that the occupant of SpecCP ‘occurs in the unmarked case form’. We see this in (99) (Van Urk & Richards’ 2015:120 exx. (12a,c)): in (99a), the sentence-initial *Ayéñ* appears without the (tonal) genitive marking that it bears in clause-internal position in (99b); and the initial constituent in (99b) appears without the locative marking that it bears in (99a).

- (99) a. Ayéñ abe Bol gooc aleth rook? (Dinka)
 Ayen 3SG.FUT Bol buy clothes town.LOC
 b. rok abii Áyèñ aleth gooc Bol
 town 3SG.FUT.NS Ayen.GEN clothes buy Bol
 both: ‘Ayen will buy Bol clothes at the town’

Put differently, there is no case connectivity between the element in SpecCP and the clause-internal position in which it is assumed to bind a trace. Moreover, the occupant of SpecCP can bind into the subject: see (100b) (reproduced, along with (100a), from Van Urk & Richards’ 2015:121 examples (13a,b)).

- (100) a. thokde ace dhuk eben kaac? (Dinka)
 goat.SG.his 3SG.PRF boy every bite
 ‘his_{j/*i} goat bit every boy_i’
 b. dhuk eben acii thokde kaac
 boy every 3SG.PRF.NS goat.SG.his bite
 ‘every boy_i, his_i goat bit’

In both of these respects, the fronted constituent behaves as though it was never any lower in the tree: it was base-generated on its high perch.³⁶

initially watch me eat the first apple, and then watch me eat the second, etc. etc. Of course, ‘thinking’ and ‘seeing’, the verbs used in (97), do not denote much of an activity, so it may be harder to make intuitive sense of pluractionality here. But whether ‘pluractional’ is quite the right term here or not, the idea that *ke* is a modifier of verbal phrases and attributes something to their interpretation will make immediate sense of the fact that *ke* does not show up on the left edge of CPs: the absence of *ke* from the position immediately preceding *cii* in (97a) is no longer a mystery (as it will almost inevitably be if *ke* is taken to be a floating quantifier articulating the successive-cyclic extraction path).

There may be other approaches to *ke* as well. The floating quantifier approach to *ke*, which Van Urk & Richards (2015) adopt without argument, is only one of the logical possibilities. Further investigation is needed here before anything can be said with any certainty about *ke* and its diagnostic power for successive-cyclic movement.

36 That (i) (= Van Urk & Richards’ (13c)) is also grammatical is potentially problematic for high base-generation. But much depends here on the exact height of the position in the tree, and the details of the dislocation phenomenon at hand. Note that Dutch *d*-word contrastive left dislocation, which like Dinka \bar{A} -fronting can show a lack of case connectivity, likewise allows ‘reconstruction’ of a bound variable contained inside the sentence-initial constituent below a

For a movement-based derivation of any construction in any language, the lack of case connectivity is always a major problem: the movement metaphor was invented precisely because of the fact that a displaced constituent seems to have close morphosyntactic ties with a position lower in the structure. At a minimum, the fact that the fronted constituent in what Van Urk & Richards take to be SpecCP is always in the unmarked case form and fails to show case connectivity should lead us to reconsider the proposition that movement (i.e., copying) is involved in the derivation of these Dinka sentences.

If the initial constituent is base-generated in the left periphery, the entire analysis of the facts that Van Urk & Richards (2015) present needs to be re-examined — perhaps along the lines of a resumptive prolepsis analysis, which would be readily compatible with the case-invariant nature of the filler in SpecCP. The unmarked case form of the fronted constituent, which Van Urk & Richards say nothing about in their account, should be the starting point of the analysis, not an afterthought. Once this is properly understood, we can try to find out which of the strategies for the formation of long \bar{A} -dependencies that I have reviewed for Hungarian, or perhaps one that Hungarian is not privy to, would be at work in Dinka. At any rate, the prospects for Dinka as an argument for densely successive-cyclic movement in a bottom-up model seem dim.

quantifier. I show this in (iia), for Dutch. Even the English hanging-topic left dislocation construction in (iib) does not seem to resist a bound-variable reading for *his*. These kinds of examples introduce complications that go well beyond the grammar of a single language.

- | | | |
|------|---|---------|
| (i) | kiteepke aanhier dhuk eben ke | (Dinka) |
| | book.PL.3SG PL.love.NS boy every PL | |
| | ‘his _i books, every boy _i loves’ | |
| (ii) | a. zijn _i boeken, daar is iedere jongen _i gek op | (Dutch) |
| | his books there is every boy crazy on | |
| | b. his _i books, every boy _i is crazy about them | |