# Dependency and Directionality

## Marcel den Dikken

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

# Chapter 5

# The trouble with subjects

# 5 The trouble with subjects

## 5.1 Introduction

Subjects are essential: no clause can do without one. But the structural subject of the clause has brought generative syntax almost nothing but trouble. *(i)* While in non-finite clauses, the structural subject position can be invisible and, as I will show, remain unprojected altogether, in finite clauses of non-pro-drop languages this position must apparently always be filled by something, even if that means having to recruit an apparently meaningless piece of material specifically for the purpose. The 'Extended Projection Principle' (EPP) says as much, without, however, shedding any explanatory light on the question (several decades of valiant attempts at deriving it notwithstanding). *(ii)* When the structural subject position of a clause contains a gap bound by an operator, the C-head of the clause must usually be silent (while in the case of non-subject filler–gap dependencies there is never a ban on filling the C-position) — this manifests itself in the form of the 'complementiser-trace effect' (*who do you think (\*that) is the winner?*, *who do you prefer (\*for) to be the winner?*) and the '*do*-trace effect' (*who (\*did) put the race off?*). Both of the effects in *(ii)* have been attributed to the 'Empty Category Principle' (ECP). The ECP may make also sense of the fact that *(iii)* in a multiple *wh*-question, a subject-*wh* cannot be *in situ* (*who won what/when/where?* versus *\*what/when/where did who win?*), and that *(iv)* the subject is the only argumental expression that roundly resists extraction from a *wh*-island (*\*who do you wonder when won?* versus *?what do you wonder when he won?*; see also *(vi)*, below). But making these all effects fall out from a licensing requirement imposed on the trace of the subject has never been straightforward, not least because of the fact that *(v)* in the presence of a high adverbial, overt subject extraction becomes perfectly legitimate across *that* (*who do you think that for all intents and purposes is the winner?*) but not across a *wh*-expression (*\*who do you wonder when for all intents and purposes won?*). No simple version of the ECP could account for the fact that *(vi)* long-distance dependencies in which a subject gap is found in a clause that is embedded in an island are disallowed altogether, regardless of the presence or absence of a complementiser (*\*who are you looking for someone* [*who says (that) could win the race*]*?*). That absence of the complementiser is not a sufficient condition for the establishment of long subject dependencies is also apparent from the fact that *(vii)* the null operator in *tough*-movement and parasitic gap constructions completely resists binding a trace in the structural subject position (*\*John is tough to think (that) could be the winner*), and that *(viii)* with a specific class of matrix verbs ('*convince*-type verbs'), the absence of *that* has no beneficial effect on extraction of an overt *wh*-subject (*\*who did they convince him (that) must be the winner?*). *(ix)* The 'anti-*that*-trace effect' seen in highest-subject relatives even turns any statement to the effect that overt complementisers bar a trace from the structural subject position right below them completely on its head: in *the man \*(that) is the winner is John*, the complementiser *that* is not only allowed but obligatory. *(x)* Highest-subject relatives also stand out within the family of clauses introduced by an operator in being relatively permeable to the establishment of argumental *wh*-dependencies across them (*this is a race which we need to find someone who can win* versus *\*this is a race which we need to find someone who we can win with*), a fact that poses a threat both to minimality-based accounts of 'intervention islands' and to the theory of absolute locality. The trouble with subjects in the standard theory is so pervasive that it cries out for a fresh look at the problem, from the perspective of the top-down approach to syntactic structure building developed in this book.

In the first part of this chapter, I will introduce the central players in the realm of the structural subject and its vagaries, starting with early principles-and-parameters theory (Chomsky 1981), and examining ways to unify and derive the two key principles (the EPP and the ECP) in the minimalist programme (esp. Chomsky 2013, 2015), against the background of the main conceptual shifts involved in the transition from GB to minimalism. In the second part, I will subsequently present my own outlook on an integrated account of the manifold restrictions imposed on subject dependencies, against the background of the theory of syntactic structure building and the creation of filler–gap dependencies presented in this work.

## 5.2    A concise historical perspective

### 5.2.1    *The subject function: Some preliminaries*

From the very beginnings of the Western linguistic-philosophical tradition, the subject has been problematic. Aristotle, in his *Perì Hermēneías*/*De Interpretatione*, made a useful terminological distinction for the *rhêma* or 'predicate' between the *kategoroúmenon* (the *grammatical* predicate) and the *symbebekós* (the *logical* predicate), but made no parallel distinction for the subject, using only the term *hypokeímenon* (corresponding directly to Latin *subiectum*) on that side of the equasion. The terminological confusion that has surrounded the notion of 'subject' finds its roots right there, and has pervaded the history of linguistic thought in the Western world.

It will be important for us to distinguish between the *notional* subject and the *structural* subject. The notional subject is very difficult to define, with thematic, semantic and pragmatic notions vying for attention. Here the thematic notion of subject (equivalent to 'external argument') will largely play no role: in the discussion to follow, both 'deep' subjects (i.e., subjects that are external to the predicate at every point in the derivation) and 'derived' subjects (i.e., subjects that start out within the predicate and are 'externalised', via passivisation or raising) by and large behave exactly the same way. By 'notional subject', I mean 'what the predicate is predicated of' — where the predication relation in question is not necessarily a relation of external thematic role assignment. Thus, I will call both the clause-initial constituent of (1a) and that of (1b) a notional subject or subject of predication at the level of the clause, generalising in this way over unergative/transitive and unaccusative constructions.

(1)    a.    John runs (a mile) every morning
        b.    John falls down the stairs a lot

By 'clause', I do not mean the semantic unit of the 'proposition' (which is commonly understood to contain a specification for tense and illocutionary force) but rather some syntactic constituent, the structural unit of predication. In my work (see Den Dikken 2006a), the structural unit for predication relations is the RELATOR phrase, depicted in (2), where one of the two terms of the structure is the predicate and the other the subject of predication.

(2)            *the structural configuration of predication*
            [$_{RP}$ XP [$_{R'}$ RELATOR [YP]]]

A variety of functional categories can serve the RELATOR function. The head of the clause, T, is one of them. The occupant of SpecTP can be the notional subject of the clause (the subject of predication), but it can also be its predicate, in so-called predicate inversion constructions such as (3b), grammatical alongside the 'canonical' copular sentence in (3a).

(3)     a.      John is the president
        b.      the president is John

SpecTP thus cannot be equated with the 'notional subject' function. But it does have one defining subject characteristic: it is the *structural* subject position. The occupant of the SpecTP position of an English finite clause controls agreement with the finite verb whenever it has a set of phi-features (person and number), and it inverts with the finite verb in Subject–Aux Inversion constructions.

(4)     a.      these facts {are/*is} the problem, aren't they?
        b.      the problem {is/*are} these facts, isn't it?

Another important property of the structural subject is that it is constrained by a variety of restrictions on its silence and its participation in filler–gap dependencies. These will be at the fore-front of the discussion throughout this chapter.

### 5.2.2    The filler requirement for the structural subject position: The 'EPP'

Predicates take subjects, by definition. For every predicate, therefore, there will always be a subject represented in the syntactic structure of the predication relation that it participates in. In a finite clause, that subject is very often found in the structural subject position, SpecTP. Thus, in (3a), *John* is the subject of predication, and the simplest imaginable representation of the structure of this copular sentence (hence the one that is to be preferred on grounds of economy of representation and derivation; see Chomsky 1995:chapter 2 for these notions) has *John* occupying SpecTP 'in the base', with T serving as the RELATOR of the predication relation between *the president* and *John*, as depicted in (5).

(5)             $[_{TP} [_{DP=Subject} John] [_{T'} T=RELATOR [_{DP=Predicate} the president]]]$

The subject of predication is not always in SpecTP. In the predicate inversion construction in (3b), what occupies the structural subject position is the predicate nominal, *the president*, with *John* residing in the specifier position of a RELATOR phrase in the complement of T. In (6), the predicate has *raised* into the structural subject position: the pronounced predicate nominal is a filler that binds a silent copy of itself in the predicate position of the small clause (RP) in the complement of T. Silence of the copy is indicated by strike-through.

(6)             $[_{TP} [_{DP=Pred} the president] [_{T'} T [_{RP} [_{DP=Subject} John] [RELATOR [_{DP=Pred} \text{the president}]]]]]$

The silent copy of *the president* in (6) is generally assumed to have to be the one in the structurally lower position, within the small clause: there is a consensus in the literature that a structure like (6) cannot give rise to the same surface output as does (5), with *John* preceding *the*

*president* in the surface string. It is certainly not unthinkable that (6) could give rise to two different PF outputs, one corresponding to (3a) and the other to (3b). It could even be interesting to say that (6) can accommodate both word orders: we could then use (6) as the syntactic representation of the version of (3a) that is information-structurally on a par with (3b), with *John* as the narrow focus of the utterance (bearing a distinctive pitch accent) and *the president* as its topic (with low and flat prosody). To be sure, the neutral information structure of (3a) is not that: rather, when (3a) is pronounced with its neutral prosody of (3a) (with the integrative sentence accent on *president*), we either get an information structure with *John* as the topic and *the president* as the information focus (delivering a categorical judgement in the sense of Brentano), suitable to serve as a reply to a question like *what's John's role in the organisation?*, or one in which the entire sentence is focused (yielding a thetic judgement), in reply to a question such as *what's new?*. It would certainly be worth exploring the possibility that the non-neutral information structure of (3a), in which *John* is focused and *the president* is topical, is syntactically represented as in (6), in exactly the same way as (3b) (which always has this particular information structure, and hence cannot serve as a reply to a question like *what's new?*). The prosody of this information-structural pattern would then be derivable from a single syntactic representation, with variation at PF being confined to the position in which the topic is pronounced: in the structural subject position (SpecTP) or inside the small clause. But the merits of such an approach are unknown at this time because, to my knowledge, it has not been pursued in the literature.

The reason why it has not been pursued is that it is customary to assume, for a language such as English, that the structural subject position of its finite clauses cannot remain silent. Something must always occupy this position, it seems — barring cases in which a filler–gap dependency is constructed that goes beyond the confines of TP. I will turn to such dependencies shortly; but for now, let us consider only constructions in which nothing is extracted from TP. It is entirely standard to assume for all such constructions, in a language like English, that the SpecTP position of finite clauses must be filled. This is why, whenever nothing semantically meaningful is placed in the structural subject position of such a clause, SpecTP must be 'plugged up' with a meaningless piece of phonologically overt material: a so-called expletive (also called 'pleonastic' or 'dummy').

(7)     a.      someone is standing at the door
        b.      there is someone standing at the door
        c.      *is someone standing at the door

In Chomsky (1981), this requirement that the structural subject position must be filled is smuggled into the definition of the Projection Principle, which regulates the relationship between the lexicon and the syntax in demanding that all the lexical properties of a predicate head be syntactically represented (i.e., 'projected' into the syntax). Chomsky adds to the need to satisfy the lexical-semantic properties of predicate heads in (8a) (which is the original Projection Principle) the need to fill SpecTP, as in (8b).

(8)             *Extended Projection Principle (EPP)*
        a.      the lexical properties of predicate heads must be syntactically represented
        b.      the structural subject position (SpecTP) must be filled

The second clause of (8), which is what makes the original Projection Principle (i.e., clause (8a)) different from the Extended Projection Principle, is usually referred to by itself as 'the EPP'. It is one of the most mysterious ingredients of generative-syntactic theory, and it is still very much at the forefront of syntactic research today: sometimes it seems as if it defines the vast majority of research efforts in the minimalist programme (Chomsky 1995 *et passim*).

     Not only is it thoroughly unclear what might be responsible for the EPP (i.e., (8b)), it is not even clear where in the grammar this constraint should be thought to hold. Standard Chomskian theorising has taken it to be a condition of 'narrow syntax'. Plainly, it is not an LF constraint: some of the stuff recruited specifically to satisfy the EPP is semantically rather useless, such as the expletive *there* in (7b). But there are indications that it might be a PF restriction. As Van Craenenbroeck & Den Dikken (2006) show, when T is marked for ellipsis (hence not present at PF), the requirement that something be placed in SpecTP is suspended, apparently as a function of the fact that T itself does not get any PF exponence.

     We see this, for instance, in (9b). For both examples in (9), Den Dikken, Meinunger & Wilder (2000) argue at length that the NPI-foci of these specificational pseudoclefts can only be licensed on an analysis in which they are part of an elliptical TP that contains their licensing negation. For (9a), the account of NPI-licensing is then entirely straightforward: analysed as in (10a), this pseudocleft gets *any wine* licensed by the negation in the elliptical portion of the postcopular clause. For (9b), such an analysis delivers the desired outcome as well — but crucially, only if the subject of the elliptical clause remains within the c-command domain of the negation at LF, i.e., does not raise to SpecTP, as depicted in (10b) versus (10b′).

(9)     a.      what they didn't bring was any wine
       b.      what didn't work was any of the printing equipment
(10)    a.      what they did*n't* bring was [$_{TP}$ ~~they did*n't* bring~~ *any* wine]
       b.      what did*n't* work was [$_{TP}$ — ~~did*n't*~~ *any* of the printing equipment ~~work~~]
       b′.    *what did*n't* work was [$_{TP}$ *any* of the printing equipment ~~did*n't* work~~]

Importantly, NPI-licensing succeeds in the structure in (10b) thanks to the fact that *any of the printing equipment*, the subject of the elliptical clause, does not raise to SpecTP. The fact that T itself is part of the ellipsis allows the subject to stay put in its VP-internal position, c-commanded by the negation. On the basis of this and other evidence, Van Craenenbroeck & Den Dikken (20xx) argue that the EPP, i.e., the requirement that the structural subject position (SpecTP) be filled, is a PF condition.

     Though it seems likely that PF considerations play a role in the context of the requirement that the structural subject position be filled, the EPP should still arguably be thought of as primarily a *syntactic* condition. I will return to a way in which it may be derived later in this chapter.

## 5.2.3   *Infinitives: Raising, control, and PRO*

The requirement that the structural subject position be filled seems to be blatantly violated in infinitival clauses. This problem manifests itself in three different contexts: *(a)* raising infinitives, *(b)* infinitival Obligatory Control complements, and *(c)* Non-Obligatory Control infinitives, esp. those in subject positions. The examples below provide illustrations for each type.

(11)          John seems/used [$_{TP}$ __ to be a good person]
(12)          John tries/wants [$_{TP}$ __ to be a good person]
(13)          [$_{TP}$ __ to be a good person] is not always easy

For raising infinitives, the problem for the EPP is only apparent, on standard derivational accounts: the subject of the raising verb, *seems* or *used*, was at some point within the bracketed portion of the structure, and satisfied the EPP for the embedded TP at that point. Rephrased representationally, what we would say is that there is a silent copy of *John* in the SpecTP position of the infinitival clause, and that this silent copy satisfies the EPP for that clause. Such an account of EPP satisfaction in raising infinitives would be hard to reconcile with a 'strictly PF' approach to the EPP. There is nothing in the embedded SpecTP position that has a phonological signature: not only is the SpecTP position not physically filled with anything, there is reason to believe that whatever occupies this position in syntax has no role to play in the phonological component at all. We see this particularly clearly in the case of *used to* constructions: the raising modal and the infinitival marker heading its complement can be contracted to form *usta*; such contraction is ostensibly not obstructed by anything in the specifier of the embedded TP.

For (12) we see this, too: here *want* and *to* can famously contract to form *wanna*. And for *want*, the facts are particularly interesting because it has been known for a long time that *wanna*-contraction IS in fact obstructed by otherwise phonologically invisible material: a variable of the subject in between *want* and *to* makes *wanna*-contraction impossible: (14) is interpretable only in such a way that *who* is interpreted as the *object* of *leave* (cf. *I want to leave her*), not with *who* read as the *subject* of *leave* (cf. *I want her to leave*).

(14)   a.      I wanna leave
       b.      who do you wanna leave
               = who do you want to leave ~~who~~
               ≠ who do you want ~~who~~ to leave

The proper structural interpretation of the contraction facts remains a topic of lively debate, which I will not engage in here. In the present context, all that matters is that while there are indications that certain silent elements rear their heads in the PF component by wielding veto power over PF operations (variables obstruct *wanna*-contraction), the underscores in (11) and (12) seem not to be occupied by such silent elements.

Principles-and-parameters theory has traditionally treated the underscores in (11) and (12) discretely: the standard approach to raising constructions of the type in (11) takes the specifier position of the embedded TP to be filled by a trace (or silent copy) of the subject; the 1980s consensus on (12) was that its embedded SpecTP is occupied by a different kind of silent element, called PRO, which is not a trace of movement. A central tenet of principles-and-parameters theory in fact compelled us not to treat the silent subject of the infinitival clause in (12) as a trace: if *John* were to raise from the lower clause into the matrix clause in (12), its chain would be associated with two thematic roles (the role assigned by the predicate *a good person*, and the role assigned by *try/want*), which the Theta Criterion proscribes.

In Chomsky (1995), the Theta Criterion was abolished, however — primarily on the basis of an Occam's Razor-style lack-of-need argument. The condition had been invented to prevent

sentences like (15a,b), involving, respectively, raising of the object into the thematic subject position (i.e., (15a) cannot mean that John killed *himself*) and raising of a direct object into the indirect object position (i.e., (15b) cannot mean that Mary showed John *himself* in the mirror).

(15)    a.    *John$_i$ killed $t_i$
        b.    *Mary showed John$_i$ $t_i$ in the mirror

But what early principles-and-parameters theory ruled out with the aid of the Theta Criterion can in fact be ruled out entirely independently in the minimalist programme, with an appeal to the general requirement that movement be *triggered*, and that triggers are uniformly *functional* categories: the assigners of θ-roles are *lexical* categories, hence θ-positions reside in the domain of lexical categories, not in the domain of functional probes; so no movement operation could ever be triggered to target a θ-position. The ungrammaticality of (15) thus falls out entirely independently of the Theta Criterion: the movements depicted are untriggered, hence illegal.

        With the Theta Criterion abolished, the possibility of legal cases of movement 'into a θ-position' then opens up in principle. Work by Hornstein and colleagues has argued that such movement does indeed exist, and that what are standardly called 'control constructions' instantiate it. Thus, in (12), for Hornstein, the underscore is occupied by the same kind of element that the underscore in (11) represents: a silent copy of the matrix subject. The 'movement theory of control' raises the question, however, of why movement operations can equip chains with two θ-roles in 'control' environments but not in (15). It faces a number of more specific obstacles as well. There has been an agitated debate in the literature of the past ten years about the movement theory of control (see Hornstein 1999 and later work). But no matter what the best way of analysing (12) may turn out to be (control, movement, or some other way), there remains an irreducible residue of infinitival clauses whose structural subject position is unlikely to be occupied by a trace or copy of something that has moved out of the infinitival clause: Non-Obligatory Control cases like (13) are in this set.

        For infinitival clauses such as the bracketed portion of (13), therefore, the question of what satisfies the EPP here arises perhaps particularly urgently. There is no obvious (or even non-obvious) way of treating such cases in terms of raising: there is no overt antecedent; and even if we were to capitalise on the interpretive presence of an antecedent (a silent experiencer belonging to *easy*, in particular), we would at best shift the weight of the problem from one silent element (the subject of the infinitival clause) to another (the silent experiencer of *easy*). Let us focus, therefore, on sentences of the type in (13) and ask how the EPP might be satisfied here.

        Baltin (1995) argues that the EPP is in fact *not* satisfied here — that is, nothing occupies SpecTP; the specifier position of T is not projected in control infinitives (OC and NOC cases alike). He does so on the basis of the distribution of floating quantifiers, pointing out that while (16) is grammatical with *all* to the left of *to*, (17) and (18) are not. On the assumption that the distribution of floating quantifiers is restricted in such a way that they are only allowed to occur in positions immediately local to a member of the chain of some noun phrase that the quantifier can be taken to quantify over, the contrast between (16), on the one hand, and (17) and (18) tells us that there is no member of the chain of the subject of the infinitive in the infinitival SpecTP.

(16)        they seem <all> to <all> be good people
(17)        they try <*all> to <all> be good people

(18)               <*all> to <all> be good people is not easy

     If Baltin's conclusion is on target, the underscores in (12) and (13) are not placeholders for some empty category: the SpecTP positions of these infinitival clauses is unoccupied. But T is clearly present: it is occupied by *to*. The vacancy of SpecTP in these cases hence cannot be linked to ellipsis of T (as in the PF approach to the EPP briefly reviewed in the previous subsection). The Q-float facts also threaten a movement approach to Obligatory Control constructions: the fact that (12) behaves so markedly differently from (11) in the context of Q-float (as seen in (16) *vs* (17)) discourages a structural assimilation of the two sentence types (though I hasten to add that these facts do not, *per se*, refute the movement theory of control, of course: facts by themselves can never refute any theory; it is the way the facts are analysed that could).

     Not only do these examples raise the question of how the EPP is satisfied, and what the EPP might be, they also pose the problem of how the silence of subjects can be regulated. It is certainly not the case that subjects have a blanket licence to remain silent if they so desire. In finite clauses, in languages such as English (so-called non-pro-drop languages), the structural subject position can never remain entirely unoccupied. And even in non-finite clauses, the silence of subjects is governed by severe restrictions. Thus, while a verb like *want* is peculiar in that it allows the subject of its infinitival complement clause to be either silent or overt (and interestingly, when it is overt it can even be coreferential with the matrix subject: see *himself* in (19a)), epistemic verbs that take a *to*-infinitival complement require their subjects to be overt.

(19)  a.        John wants {Bill/himself} to be president  
     b.        John wants to be president

(20)  a.        John considers {Bill/himself} to be president  
     b.        *John considers to be president

     The distribution of the silent subject of infinitival clauses was regulated in the principles-and-parameters era with the aid of the Binding Theory — more specifically, by a corollary of the binding-theoretic status of PRO called the PRO Theorem.

(21)               *PRO Theorem*  
                 PRO is ungoverned

The PRO Theorem is a theorem of the Binding Theory on the understanding that *(a)* PRO is featurally specified as [+anaphoric, +pronominal], and *(b)* the distributions of [+anaphoric] and [+pronominal] proforms are governed by the conditions in (22) and (23), which are part of the Binding Theory of Chomsky (1981).

(22)               *Condition A*  
                 an anaphor must be bound in its governing category

(23)               *Condition B*  
                 a pronoun must be free in its governing category

The precise definition of 'governing category' is immaterial for our purposes here; what is relevant is just that, as its name suggests, the governing category of α contains the governor of α. With this said, we realise quickly that PRO's only chance not to violate Conditions A and B of the Binding Theory is for it to have no governing category (in which case it satisfies both conditions vacuously) — hence for it not to have a governor. Thus, (21) follows.

Returning now to (19) and (20), the way the principles-and-parameters theory accounted for the contrast between *want*-type and *consider*-type verbs with respect to the legitimacy of a silent subject of their infinitival complement was to say that while *want* takes a CP complement, the infinitival complement of *consider* is no larger than TP. The C-head of the complement of *want* then 'shields' PRO from being governed by the matrix verb in (24). Conversely, the absence of such a protective shield in (25) causes the PRO Theorem inevitably to be violated here.

(24)     $[_{VP}$ want $[_{CP}$ C$_⌀$ $[_{TP}$ PRO to be president]]]
(25)     *$[_{VP}$ consider $[_{TP}$ PRO to be president]]

The success of the account of the distribution of silent subjects of non-raising infinitives has always been muted by the fact that it has not turned out to be easy to provide clear independent support for the hypothesis that the PRO-infinitival complement of *want* is a CP with a silent head, and that *consider* cannot take such a complement. On the one hand, the silent C-head in (24) has never been easy to bring to light: the grammaticality of *wanna*-contraction in (14a) would, if anything, seem to suggest that there is very little structure in between *want* and *to*. On the other hand, the fact that *consider*-type verbs can clearly take *finite* CP complements (*John considers that Bill should be president*) would make it difficult to intelligently formulate a selectional restriction preventing such verbs from taking infinitival CP complements.

The entire enterprise of accounting for PRO's distribution has largely been left in limbo with the abolishment, in Chomsky (1995), of the structural relation of government and the binding-theoretic conditions in (22) and (23). It is customary for syntacticians to still refer to the ghost of the PRO Theorem and to assume without comment that it can continue to rein in the distribution of PRO. But of course, with government and Conditions A and B off the table, the account of the distribution of PRO should be sent right back to the drawing board.

Later in this chapter, I will assume the existence of PRO, and present a new perspective on its distribution that makes no appeal to government or the Binding Theory.

### 5.2.4    Subject extraction and the 'ECP'

#### 5.2.4.1 The empirical core

A very well-known fact — and one that basically every syntactic theory worth its salt has tried its hand at — is that while the complementiser *that* is optional in (26a), it is obligatorily omitted in (27a), involving long-distance fronting of the subject of the embedded finite clause. This is the '*that*-trace effect'. Also well known (since Bresnan 1977:194, fn. 6 and, more recently, Culicover 1993) but less frequently discussed is the fact that the presence of certain material in the left periphery of the clause whose subject is fronted into the matrix clause lifts the ban on an overt complementiser — as a matter of fact, the complementiser in (27c) *must* be overt, but this is independent of subject

extraction: in (26c) as well, *that* is inomissible.[1] Seldom noticed in the literature,[2] however, is the fact that the '*that*-trace effect' is not specific to long-distance fronting of the *notional* subject. Instead of the subject of predication, the predicate nominal can be placed in the structural subject position ('predicate inversion'; see Den Dikken 2006a and references cited there). I did this in (26b). As (27b) shows, in the absence of a high adverbial, long extraction of the raised predicate (in the form of *which*) requires *that* to be silent in order for a grammatical result to emerge, just as long extraction of the notional subject does in (27a). Predicate inversion constructions with long-distance fronting of the raised predicate nominal in fact mimic the behaviour of familiar long subject extraction cases perfectly: the Bresnan/Culicover effect manifests itself in (27d) in the same way as it does in (27c).

(26)   a.      everybody thinks (that) John Jones is the mayor of this town
       b.      everybody thinks (that) the mayor of this town is John Jones
       c.      everybody thinks *(that) for all intents and purposes John Jones is the mayor of this town
       d.      everybody thinks *(that) for all intents and purposes the mayor of this town is John Jones

---

1        Kandybowicz (2006) discusses a very similar high-adverbial effect on long subject extraction in Nupe. More recently, Erlewine (2016) shows that Kaqchikel exhibits this effect as well. Though Dutch is not usually considered to be a '*that*-trace language' (though see Den Dikken 2007b and references cited there), it gives rise to a high-adverbial effect in constructions in which in English long subject extraction is impossible regardless of the presence or absence of the complementiser, except when a high adverbial is present (see (i), based on Stowell 1981:410; and see section 5.3.14 below for discussion). The Dutch equivalent of (ia) is ungrammatical as well; and just as in English (ib), the result improves significantly when a high adverbial is inserted to the immediate right of the complementiser, as in (iib). (Object extraction is unproblematic in both English and Dutch in these constructions.)

(i)     a.      *which terrorist organisation did they convince you (that) was preparing an attack?
        b.       which terrorist organisation did they convince you that *in all likelihood* was preparing an attack?
(ii)    a.      *welke terroristische organisatie hebben ze je overtuigd dat een aanval aan het voorbereiden is?
                which terrorist organisation have they you convinced that an attack at the prepare is
        b.      ?welke terroristische organisatie hebben ze je overtuigd dat *naar alle waarschijnlijkheid*
                which terrorist organisation have they you convinced that to all probability
                een aanval aan het voorbereiden is?
                an attack at the prepare is

2        The reason why this is so often overlooked is that Ā-fronting of the inverted predicate of a predicate inversion construction is very often ungrammatical. But as Den Dikken (2006a:125) points out, as long as the predicate of a predicate inversion construction is not moved into a focus position, Ā-movement of the raised predicate is informationally coherent. As we see in (27b), it succeeds (as in the case of Ā-movement of the notional subject in (27a)) just in case the complementiser is silent.
        Bresnan (1977:186) points out that locative inversion (a predicate inversion operation involving PP-predicates) also gives rise to *that*-trace effects under onward Ā-extraction: (i) is an example due to Bresnan.

(i)                     in which villages do you think (*that) are found the best examples of this cuisine?

The analysis of locative inversion is sufficiently controversial (see Den Dikken 2006a for discussion) to make it difficult to present (i) as a straightforward case of a *that*-trace effect. For the analysis of copular inversion constructions of the type *the mayor of this town is John Jones* there is a broad consensus in the literature today that its syntax involves raising of the predicate nominal into the structural subject position. See Heycock (2012) for recent discussion and references.

(27)  a.      John Jones, who everybody thinks (*that) is the mayor of this town, must be a hero
      b.      the mayor of this town, which everybody thinks (*that) is John Jones, must be a hero
      c.      John Jones, who everybody thinks *(that) for all intents and purposes is the mayor
              of this town, must be a hero
      d.      the mayor of this town, which everybody thinks *(that) for all intents and purposes
              is John Jones, must be a hero

Though this is ignored surprisingly commonly in more specialised studies, introductory textbooks and general surveys of the territory of complementiser-trace effects (such as Pesetsky 2015) will typically impress upon the reader the fact that the '*that*-trace effect' does not stand on its own but finds a close parallel in '*for*-trace effects' seen in infinitival clauses, and '*wh*-trace effects' emerging from extraction of the subject of a subordinate question.

In (29a) we see that long-distance fronting of the subject of a *for-to* infinitive in English requires the complementiser (here *for*) to be omitted. Once again, the effect is not specific to fronting of the *notional* subject: the predicate inversion example in (28b) allows the predicate nominal to be relativised, in the form of *which*; but when this happens, *for* must once again be absent, as we see in (29b). The Bresnan/Culicover effect cannot be demonstrated for *for-to* infinitives. Quite independently (for case reasons, as I will argue below), it is impossible to insert adverbial material between the complementiser *for* and the subject of the infinitival clause: the baseline examples in (28c,d) are already ungrammatical regardless of the presence or absence of *for*, and they do not improve under extraction of the structural subject of the infinitive.

(28)  a.      everybody prefers/wants (for) John Jones to be the mayor of this town
      b.      everybody prefers/wants (for) the mayor of this town to be John Jones
      c.      *everybody prefers/wants (for) for all intents and purposes John Jones to be the mayor
              of this town
      d.      *everybody prefers/wants (for) for all intents and purposes the mayor of this town to
              be John Jones
(29)  a.      John Jones, who everybody prefers/wants (*for) to be the mayor of this town, must
              be a hero
      b.      the mayor of this town, which everbody prefers/wants (*for) to be John Jones, must
              be a hero

In (31a,b) (derived from (30a,b)), we see that the structural subject (again, regardless of whether it is the notional subject or a raised predicate nominal) radically resists extraction from an embedded question. This time around, nothing can save the result — not even the insertion of a high adverbial of the Bresnan/Culicover type: (31c,d) are no better (in fact, because of their additional complexity, even worse) than the a– and b–examples.

(30)  a.      everybody wonders why John Jones is the mayor of this town
      b.      everybody wonders why the mayor of this town is John Jones
      c.      everybody wonders why for all intents and purposes John Jones is the mayor of this
              town
      d.      everybody wonders why for all intents and purposes the mayor of this town is John
              Jones

(31)  a.      *John Jones, who everybody wonders why is the mayor of this town, must be a hero
      b.      *the mayor of this town, which everybody wonders why is John Jones, must be a hero
      c.      *John Jones, who everybody wonders why for all intents and purposes is the mayor of this town, must be a hero
      d.      *the mayor of this town, which everybody wonders why for all intents and purposes is John Jones, must be a hero

These facts form the empirical core of the problem of long subject extraction. Most of the ingredients of it have been on the menu for a very long time. Perlmutter (1968, 1971) discovered the '*that*-trace effect'. Ross (1967:445ff.) is credited with the observation that the subject of a *for-to* infinitive resists extraction. Bresnan (1977) was the first unify the '*that*-trace effect' and the '*for*-trace effect',[3] added the ameliorating influence of a high adverbial to the empirical puzzle, and emphasised that the effects in question are not confined to the *notional* subject: she herself mentions the fact that onward extraction of a PP having undergone locative inversion in the subordinate clause gives rise to the *that*-trace effect (recall fn. 2, above); the copular inversion cases presented above make the point more generally and less controversially. Well before the advent of the principles-and-parameters framework (Chomsky 1981), therefore, the nature and the extent of the problem with subject extraction were largely known. Nonetheless, as Pesetsky (2015:9) puts it in a recent survey of the literature on *that*-trace effects and related phenomena, '[e]ven at a relatively broad level of generality, it has proven frustratingly hard to determine just what kind of phenomenon the complementizer-trace effect is'.

In the following subsections, I will begin by sketching the standard principles-and-parameters approach to complementiser-trace effects, couched in terms of the Empty Category Principle (ECP), and will then proceed to a brief discussion of 'minimalist' perspectives on the problem.[4] The historical overview will then be drawn to a close, and I will subsequently embark on the presentation of a new and fully integrated account of the trouble with subjects.

*5.2.4.2 The ECP*

The central ingredient of Chomsky's (1981) account of the distribution of traces left behind by movement is the Empty Category Principle, given in (32).

---

3      Chomsky & Lasnik (1977:455ff.) reject a unification on account of the fact that 'Ozark English' allows *for*-trace strings but not *that*-trace ones. But Pesetsky (2015) points out that Chomsky & Lasnik's argument was never backed up by a systematic empirical investigation. As a matter of fact, Sobin (1987) presents the results of a judgement study conducted at the University of Iowa from which it emerged that many university students there detected no contrast of the type characterised by the familiar *that*-trace effect. Pesetsky (2015:28) writes that 'the lack of a *that*-trace contrast is often believed to be localized to the American Midwest', but that it is unclear whether this is truly the case or whether dialectal or idiolectal differences are involved. Indeed, over the years, I have discovered that many of my native-English-speaking graduate students, coming from various parts of North America, found the deviance of *that*-trace sequences much less severe than they have standardly been claimed to be in the generative literature.

4      My discussion here will not aim at comprehensiveness or exhaustivity. Pesetsky (2015) presents an eminently detailed and informed discussion of the many different kinds of approaches to complementiser-trace effects. I refer the reader to this readily available reference work for details that the present discussion has found no space to include.

(32)             *Empty Category Principle (ECP)*
                 a non-pronominal empty category must be properly governed

The ECP bears directly on the distribution of traces because traces are precisely the only phrasal empty categories that are non-pronominal: in the typology of silent noun-phrase types, traces fit into the [+anaphoric, –pronominal] and [–anaphoric, –pronominal] cells (NP-traces and variables, resp.); the [+anaphoric, +pronominal], as we have seen already, is filled by PRO, and its little brother *pro* is the representative of the [–anaphoric, +pronominal] category. Though (32) is not strictly equivalent to the statement that a *trace* must be properly governed,[5] for our purposes here such a simplification is innocuous.

In order to determine how (32) reins in the distribution of traces, we need to know how an empty category can be properly governed. For most of its existence, the principles-and-parameters framework has worked with a *disjunctive* formulation of the ECP, assuming two ways in which proper government can come about: *(a)* θ-government, and *(b)* antecedent-government. For an element to be θ-governed, it must be governed by the head that assigns a θ-role to it. Structural subjects (i.e., the occupants of SpecTP) are never governed by the head that assigns them their θ-role. So for traces in SpecTP we have no business with *(a)* at all. Antecedent-government *(b)* is a very local relation between a trace and its immediate antecedent. It is this relation that Chomsky (1981) exploits in his quest to account for the complementiser-trace effect and the '*wh*-trace effect'.

Let us begin with the '*wh*-trace effect' in (31): the fact that nothing can bind a trace in the structural subject position across an intervening *wh*-constituent. This follows entirely and straightforwardly from the ECP — more specifically, from Minimality:

(33)             *Minimality*
      a.         no constituent α can have more than one governor
      b.         whenever α has more than one potential governor, only the governor closest to α
                 governs it

In the configuration in (34), the trace in SpecTP is separated from its antecedent by a *wh*-element that is in a position (the SpecCP of the subordinate clause) from which Ā-bound traces can in principle be antecedent-governed — i.e., the *wh* in SpecCP is a *potential* antecedent-governor for the subject trace, but it does not actually serve as an antecedent for the subject trace; since it is closer to the subject's trace than the subject-*wh*, it prevents the latter from antecedent-governing its trace. The fact that insertion of a high adverbial in between the *wh* in SpecCP and the subject trace does not change this at all is something we will revisit later in this section, in the context of a discussion of the Bresnan/Culicover effect in *that*/*for*-trace cases.

(34)             *[$_{CP1}$ *wh*$_{SU}$ ... [$_{CP2}$ *wh* [C$_{\varnothing}$ (*for all intents and purposes*) [$_{TP}$ *t*$_{SU}$ ...]]]]

---

5         Stowell (1981) in fact made interesting use of the original 'non-pronominal' formulation to reduce the distribution of complementiser omission to the ECP: a silent C$_{\varnothing}$ is a non-pronominal empty category, hence subject to the ECP in (32).

The intervening *wh* in (34) does not just prevent the licensing of a *subject* trace in the subordinate clause, CP2: it is also expected to make it impossible for the trace of a non-argument to be properly governed. Just like subjects, non-arguments can satisfy the ECP only via antecedent-government. Since the *wh*-element in the specifier position of CP2 interferes with the proper antecedent-government of a trace inside CP2 by an element outside CP2, it should be just as bad to extract a non-argument from CP2. And indeed it is: (35) is ungrammatical as well, on the readings indicated there (in which the sentence-initial *wh*-element is understood to modify the lower clause).

(35)   a.   *how does everyone wonder [whether John Jones got elected mayor of this town *t*]?
       b.   *when does everyone wonder [whether John Jones got elected mayor of this town *t*]?

For the '*that*-trace effect' and the '*for*-trace effect', Chomsky (1981, 1986) presents an ECP account that capitalises directly on the role of the complementiser in the licensing of the subject trace. The active ingredient in the analysis is once again Minimality. To see how Minimality accounts for the basic '*that/for*-trace effects', consider the structure in (36), a structure in which the subject of the clause has raised to the local SpecCP:

(36)              *$[_{CP}$ *wh*$_{SU}$ [C=*that/for* [$_{TP}$ *t*$_{SU}$ ...]]]

In this structure, the complementiser *that/for* governs the trace in the structural subject position. Because C is closer to the trace than its local antecedent in SpecCP, the complementiser thus precludes the establishment of an antecedent-government relation between the trace and its immediate antecedent. No closer antecedent for the subject trace can be provided (see further below for discussion), so the derivation crashes.

Note that in the cases of long subject extraction under discussion (36) reflects an *intermediate* state in the derivation: one in which the subject-*wh* has raised to the subordinate SpecCP on its way up, eventually, into the matrix SpecCP. The hypothesis that movement proceeds successive-cyclically is a pillar of all work in the principles-and-parameters framework. But of course (36) should be equally ungrammatical if it were the final stage in the derivation. Thus, the ECP also rules out 'doubly-filled Comp effects' with subjects — accurately, as it turns out, but rather unhelpfully because 'doubly-filled Comp effects' are not just confined to subjects: in the varieties of English that show 'doubly-filled Comp effects', they are manifest for any and all fillers of SpecCP.

(37)   a.   everyone wonders who (*that) is the mayor of this town
       b.   everyone wonders which town (*that) he is the mayor of
       c.   everyone wonders when (*that) he became the mayor of this town

But (36) does helpfully apply to the final derivational stage in root clauses, with a form of the dummy *do* replacing the complementisers *that* and *for*. That (38) is ungrammatical is shown by (39a); this effect is specific to subjects, as shown by (39b,c).

(38)              *$[_{CP}$ *wh*$_{SU}$ [C=*do* [$_{TP}$ *t*$_{SU}$ ...]]]
(39)   a.   what people (*do) like quinoa?
       b.   what *(do) these people like?
       c.   why *(do) these people like quinoa?

(Of course, (32) can only account for the *ban* on *do*-support in (39a); it has nothing to say about the obligatoriness of *do*-support in (39b,c), which does not correspond to anything we find in subordinate clauses.)

While this suggests that the Minimality-based analysis can generalise over more than what it was specifically designed for, it is not immediately clear how it can allow for non-arguments to extract from subordinate clauses with an overt complementiser. One might expect an overt complementiser to systematically interfere with proper government of any and all traces that depend for their licensing on antecedent-government. But this turns out not to be the case: it is only subject extraction that shows a sensitivity to the presence or absence of a lexical complementiser or the dummy *do*. For *do*-support, (39c) already illustrates that non-arguments do not resist the filling of the C-head position. And for *that* and *for*, we see the lack of a '*that/for*-trace effect' under non-argument extraction in (40).

(40)    a.    how does everyone think [(that) John Jones got elected mayor of this town *t*]?
        b.    how does everyone prefer [(for) John Jones to govern as mayor of this town *t*]?

For the Minimality-based account of the '*that/for*-trace effect', the grammaticality of (40) poses a serious problem. Lasnik & Saito (1984) designed a technical solution for this problem (later revisited and streamlined in Chomsky & Lasnik 1993, reprinted as chapter 1 of Chomsky 1995). The solution has three active ingredients: *(a)* an approach to the ECP that splits it into two components, a 'gamma-marking' operation (assigning $[+\gamma]$ to traces that have a proper governor and $[-\gamma]$ to those that do not) and a verification at LF for all traces that are present in the LF representation (such that any $[-\gamma]$-marked trace in an LF structure will rule the corresponding sentence ungrammatical); *(b)* the hypothesis that traces left behind by argument extraction receive a $\gamma$-mark as soon as they emerge whereas traces of non-argument extraction are $\gamma$-marked only at LF; and *(c)* maximal deletion of semantically meaningless material at LF. For subject extraction, this theory delivers the '*that/for*-trace effect' as before: Minimality causes the trace of the subject in SpecTP to be marked $[-\gamma]$ at S-structure; though the complementiser is meaningless and hence an excellent candidate for deletion, complementiser deletion at LF will not save the subject-extraction case because the offending trace in SpecTP received a $[-\gamma]$-mark already at S-structure, and this stigma will stick to the trace throughout the subsequent derivation. But for non-argument extraction, $\gamma$-marks are only handed out at LF, by which time the complementiser that caused the Minimality violation with subject extraction has been deleted; with C empty at LF, nothing prevents the trace of the non-argument from being antecedent-governed, hence $[+\gamma]$-marked, by the intermediate trace in SpecCP.

This solution to the contrast between subject extraction and non-argument extraction in the realm of complementiser-trace effects is technically sound. But it faces two major drawbacks. One is that it splits the ECP (itself already a disjunction of $\theta$-government and antecedent-government) into two mechanisms: a $\gamma$-marking operation (performed by what we might call the 'trace police') and an LF check-up (done by a 'judge' presiding at LF and sentencing to ungrammatically all LF-representations that have one or more $[-\gamma]$-marked traces in them). The other main disadvantage of the Lasnik & Saito (1984) proposal is that it encodes the difference between arguments and non-arguments twice: positionally (A *vs* Ā) as well as in terms of the point at which $\gamma$-marking obtains. Primarily for these reasons, this proposal is not particularly attractive from the perspective of theoretical parsimony and explanatory adequacy.

Can the Minimality-based approach to the '*that/for*-trace effect' make sense of the Bresnan/ Culicover facts? Within the framework of Chomsky (1986), it can indeed (though Chomsky himself does not discuss this). For Chomsky (1986), the trace of the subject in (36) is prevented from being properly governed by the trace in SpecCP, and moreover, as I said before, no closer antecedent for the subject trace can be postulated. This is because, on Chomsky's (1986) assumptions, intermediate adjunction to TP is illegitimate.[6] But suppose now that we insert an additional maximal projection between C and TP, with a silent head $F_{\varnothing}$, and fill its specifier position with a high adverbial. We then procure an intermediate stopover site for the subject on its way to SpecCP: it can adjoin to FP on its way up. That gets us (41), which is well-formed as far as the trace in SpecTP is concerned: it is properly antecedent-governed by the trace adjoined to FP. Of course this intermediate trace adjoined to FP now faces the threat of violating the ECP. But here Lasnik & Saito's (1984) theory comes to the rescue: the meaningless intermediate trace is free to be deleted at LF; since the ECP is verified at LF and since the intermediate trace does not make it there, no ECP violation arises in (41).

(41)         $[_{CP}$ *wh*$_{SU}$ $[C=$*that/for* $[_{FP}$ $t_{SU}$ $[_{FP}$ *for all intents and purposes* $[F_{\varnothing}$ $[_{TP}$ $t_{SU}$ ...$]]]]]]$

Unfortunately, if this is to be the account of the Bresnan/Culicover data, the fact that insertion of a high adverbial has no beneficial effect on long subject extraction across a *wh*-element in SpecCP (recall (31c,d)) comes back to haunt us. We had already presented an ECP-account of the '*wh*-trace effect' in (34). When we now update this analysis and include the 'FP' of (41), with an intermediate trace of the subject adjoined to it, we get (42):

(42)         $[_{CP1}$ *wh*$_{SU}$ ... $[_{CP2}$ *wh* $[C_{\varnothing}$ $[_{FP}$ $t_{SU}$ $[_{FP}$ *for all intents and purposes* $[F_{\varnothing}$ $[_{TP}$ $t_{SU}$ ...$]]]]]]]$

A derivation along these lines is expected to deliver a Subjacency violation (the movement from the FP-adjoined position straight into the matrix clause across SpecCP2 covers too great a distance), but it should satisfy the ECP in the same way that (41) does: the trace of the subject in SpecTP is properly antecedent-governed by the FP-adjoined trace; this trace itself deletes at LF and hence cannot be responsible for any ECP effects. This is not the desired result: (31c,d) are much worse than one's typical Subjacency effect with long object extraction, such as (43).

(43)         [?]?which town does everyone wonder why (for all intents and purposes) John Jones wants to be mayor of?

Placing the high adverbial in the specifier position of a functional projection between C and T, as does Culicover (1993), thus ultimately does not fit in very well with the theory of Chomsky (1986), with its Rigid Minimality.

The FP-based account does not help out for Rizzi's (1990) Relativised Minimality either. If Rizzi were to place the high adverbial in an Ā-specifier position between the trace and its antecedent, the adverbial in SpecFP would qualify as an intervener: it is a phrase in an Ā-position that is closer to the subject's trace than SpecCP is, and should therefore block the relationship between the filler

---

6       Chomsky bars this by blunt stipulation. For subject extraction, at least, this stipulation can be eliminated in favour of an appeal to vacuous movement (Chomsky 1986) or anti-locality (Grohmann 2003): the step from SpecTP to the adjunction position on the edge of TP would be too short.

of SpecCP and that of SpecTP. I will not discuss Rizzi's (1990) analysis of subject extraction in any detail here; I refer the reader to the original work for details, and add merely that this analysis cannot as it stands account for the spectrum of data reviewed above.

We have not talked yet about the question of why the ill-formed examples reviewed above all become grammatical once the filler of the C-slot is dropped. Here, the Minimality-based approach only has a stipulation to offer: a silent C is not a governor, hence a silent C does not interfere with proper antecedent-government of the trace in subject position. So at the end of the day, the Minimality account of '*that/for*-trace effects' is only a partial account: it does reasonably well (once amplified with the theory of Lasnik & Saito 1984) on accounting for the badness of subject extraction across a local complementiser and even extends to *do*-support in root *wh*-questions, but it does not have a deep analysis of what happens when there is no complementiser or dummy *do* in C.

One last note is in order before we leave the discussion of the ECP. The entire account of the distribution of gaps bound by subjects in this subsection was based on the premise that the gaps in question are non-pronominal empty categories — *traces*, in other words. Cinque (1990) has argued, however, that especially specific or D-linked fillers have the ability to bind gaps that are pronominal in nature — *pro*'s. Binding-based Ā-dependencies differ crucially from movement-based ones in that the gap comes under the purview of the ECP only in the latter. So if the ECP is to be held responsible for the (il)legitimacy of subject extraction, and if the ECP is formulated precisely with reference to *non*-pronominal empty categories, we are led to expect that we should freely be able to ignore the ECP when a specific or D-linked *wh*-subject is involved. The facts show clearly that this is not the case: 'complementiser-trace effects' manifest themselves just as robustly for extraction of *which*-phrases (which are specific/D-linked) as they do under extraction of 'bare' *who* or *what*:

(44)  a.  which man does everyone think (*that) is the mayor of this town?
      b.  which man does everyone prefer (*for) to be the mayor of this town?
      c.  *which man does everyone wonder whether is the mayor of this town?

This suggests that an account of 'complementiser-trace effects' that is sensitive quite specifically to the particular nature of the gap in the structural subject position is unlikely to get the facts right.

I will now leave the classic principles-and-parameters era behind, and move on to briefly consider two approaches to 'complementiser-trace effects' in the minimalist programme.

### 5.2.4.3 A positional account and its limitations

In a syntactic account of the facts of subject extraction, it would seem *prima facie* attractive to make a distinction between SpecTP and some position in T's complement, and to argue that extraction of the subject is successful only from the latter, not from SpecTP. Tied to a hypothesis regarding a special feature-sharing relation between C and T, and a late insertion approach to the distribution of the complementiser, this positional approach could account for several of the core facts.

An approach along these lines could could run as follows. Imagine first that the complementisers *that* and *for* are lexicalisations of a C–T feature-sharing complex: for Pesetsky & Torrego (2001), *that* is the reflex of T-to-C movement, while in Chomsky's (2008) 'feature inheritance' (FI)

approach, it could be the spell-out of the finite C–T FI complex in subordinate clauses.[7] Assume furthermore that when C and T form this special unit, the SpecTP position must be projected (i.e., the 'EPP' is in effect) but cannot be occupied by a trace.[8] These assumptions combined ensure that when C and T form a special unit spelled out at C as *that* or *for*, something must raise to SpecTP ('EPP'), and whatever raises to SpecTP cannot be extracted out of the clause. This delivers the basic '*that*/*for*-trace effect' in the a–examples in (27) and (29), and, given one additional assumption, also the Bresnan/Culicover effect in the c–examples. When *that* or *for* is present and nothing else occupies SpecTP, the subject is raised to the subordinate SpecTP and 'frozen' there. When a high adverbial such as *for all intents and purposes* is present, we could assume it to be placed in SpecTP.[9] Assuming so, we can then extract a subject or raised predicate out of the subordinate clause from a position in the complement of T. And when *that* or *for* is absent, the subject or raised predicate is also free to extract from a position lower than SpecTP (cf. Chomsky 2008, Legate 2011), so no 'freezing effect' will arise.

      This gives us a purchase on many of the facts (though not in any obvious way on '*wh*-trace effects'). But unfortunately, the account is only as good as one of its central assumptions: that when C and T form a special unit, the SpecTP position must be projected (i.e., the 'EPP' is in effect) but cannot be occupied by a trace. Let us see if we can do better than this, unifying the 'ECP' and 'EPP', and deriving this assumption.

---

7      Extending this approach to *for* is not entirely straightforward: it is not self-evident what feature(s) *for* and *to* might share. An integrated account covering both the *that*–$T_{Fin}$ and *for*–$T_{Infin}$ pairs might capitalise on (non)nominality as the shared feature. The finite complementiser *that* is evidently nominal in origin (it is historically a demonstrative), and for finite tense inflection one could also make a case that it is nominal (thus, note that Dutch past-tense morphology, *-de* (with a phonologically conditioned allomorph *-te* after stems ending in a voiceless obstruent), is formally identical with the common-gender definite article, *de*; see also Kayne 2015 for the idea that past tense is nominal). So *that* and $T_{Fin}$ could plausibly be thought to share the feature [+D]. It is important to bear in mind that the feature-sharing relation between finite C and T cannot reasonably be considered to involve either [±PAST] or φ: *that* does not inflect for tense (i.e., it has the same form in present-tense and past-tense clauses), nor does it share with $T_{Fin}$ the φ-features of the subject (on the latter, see Haegeman & Van Koppen's 2012 astute demonstration based on complementiser inflection in Dutch dialects: the inflected complementiser and $T_{Fin}$ can have *different* sets of φ-features). The only feature that C and $T_{Fin}$ can truly be assumed to have in common is the feature [+D]. This approach has the benefit of extending to the relation between the infinitival complementiser *for* and $T_{Infin}$=*to*: these two elements are both prepositional, hence [–D].

8      In the terminology of Rizzi's (2006, 2007) 'criterial freezing', we could call SpecTP a 'criterial' position. In this work, I will not have recourse to 'criterial freezing': the concept of 'freezing', taken literally, makes sense in a bottom-up derivational system, but not in the top-down approach taken here.

9      The idea that adverbials can be the occupant of SpecTP is not standard in the literature, which usually reserves SpecTP for argumental material, and occasionally allows predicates to be in this position, as in (26b), but generally does not allow other non-argumental constituents to occupy it. For this idea to be operationalised, we would need to make an ancillary assumption, viz., that finite auxiliaries/copulas do not necessarily surface in T (for if they did, (26c) would not be derivable by placing *for all intents and purposes* in SpecTP: we would then get this adverbial followed by the finite copula, which is impossible: *\*everybody thinks that for all intents and purposes is John Jones the mayor of this town*). The idea that finite auxiliaries/copulas can surface further to the right than T, while unusual, is certainly not indefensible: on the assumption that adverbials can only be attached to maximal projections (not to intermediate projection levels), the grammaticality of *John probably is/will/won't (be) leaving town tomorrow night* may show that *is* and *will/won't* can surface lower than T; if so, with *for all intents and purposes* in SpecTP, we can get the right word order in (26c) by placing *John Jones* in the specifier position of the phrase in the complement of T in whose head *is* sits.

## 5.2.4.4 A labelling account and its limitations

Chomsky (2015) argues that a unification of 'ECP' and 'EPP' is possible in the realm of canonical complementiser-trace effects. He concentrates on *that*-trace effects (nothing is said about *for*-trace cases), and argues that *(a)* like bare roots, T is too 'weak' to provide a label for its own projection, but *(b)* raising the subject to SpecTP provides a labelling opportunity for the [DP TP] structure created, thanks to phi-features shared by DP and T(P). This recasts the 'EPP' in terms of labelling. To now unify this with the 'ECP' (or at least, the subject/object asymmetry that the classic ECP derived), Chomsky proposes that *(c)* moving the subject from SpecTP to SpecCP leaves a silent copy of the subject behind in SpecTP that, at the point the next phase is reached, becomes invisible to the labelling algorithm. Let us examine how this accounts for the English *that*-trace pattern.

The subject, in English, must always raise to SpecTP to help provide a label for the [DP TP] structure. The placement of the subject to the left of modals, finite auxiliaries, and the infinitival marker *to*, all of which are standardly assumed to occupy T, indicates that it is systematically in SpecTP, regardless of whether CP is introduced by an overt complementiser or not:

(45)  a.   I think (that) he *is*/*may* be the winner
      b.   I want/prefer (for) him *to* be the winner

When *that* is present, CP is a phase. Movement of the subject of the finite clause to SpecCP is illegitimate because, with the silent copy of the subject in SpecTP becoming invisible to the labelling algorithm at the point at which CP is reached, TP is left unlabelled. It follows that the subject of a finite clause cannot be extracted when *that* is present.[10]

When there is no overt complementiser present in C, Chomsky assumes that C is deleted in syntax.[11] According to Chomsky (2015:11), when that happens,

---

10      It actually does not matter whether movement proceeds successive-cyclically, via the specifier of the embedded CP, or in one fell swoop: the silent copy that the subject leaves behind is unable to label TP, and the derivation crashes regardless of how it proceeds. And when the complementiser is silent (see the next paragraph for details), there is no reason for the subject to extract via an intermediate stop-over in SpecCP either: when the complementiser is silent, Chomsky assumes that it hands over its phasehood to TP and is itself 'de-phased'. So a null-headed CP is not a phase. The C-deletion *cum* inheritance approach to complementiser omission thus derives *lack* of successive cyclicity for cases of long subject extraction in the absence of an overt complementiser. As I already pointed out in chapter 2, above, the presumed successive cyclicity of extraction from CP is *not* derived in the labelling system of Chomsky (2015). Everything that Chomsky (2015) says about long subject extraction and *that*-trace effects, and the unification of 'EPP' and 'ECP' under labelling, can be preserved (to the extent that it is worth preserving) without the assumption of successive cyclicity — and, for that matter, without the assumption of a bottom-up derivational structure-building model.

11      Note that Chomsky (2015) assumes that C is deleted *in syntax*. He does not assume that when the complementiser is absent, C is present but not assigned a phonological feature matrix, nor does he assume that the deletion process affects only the *content* of C: the C-*head* disappears. The result is a rather roundabout process whereby C is merged with TP first, then hands over all of its properties (including its phasehood) to T, and deletes. If CP stays behind after C is deleted, it creates a major labelling problem further up the tree: it is unclear how a CP with a head that is deleted in syntax can be selected, or even survive without a head at all. Thus, the result derived for 'C deletion' cases (i.e., *that*-less finite complement clauses) is entirely similar to what would come out if, in the absence of a complementiser, the syntax decided to forgo Merge of a CP altogether, and made TP serve as the verb's complement directly (see Bošković & Lasnik 2003, Grimshaw 1997 etc. for proposals to the effect that in the absence of *that*, the matrix verb takes a TP complement).

> phasehood is inherited by T ... along with all other inflectional/functional properties of C
> (φ-features, tense, Q), and is activated on T when C is deleted

Activation of T as a phase-head, via inheritance from C, for some reason does not endow T with the ability to label its own phrase, however: recall from (45) that the subject must still help T out for labelling purposes when C is silent. Why, then, can the subject be extracted when C is deleted in syntax, whereas it cannot when C is spelled out by a complementiser? Key here is that after T has inherited phasehood from a deleted C, the [DP TP] structure is a single phase, already labelled by phi-feature agreement, and that 'memory is phase-level' (Chomsky 2015:11): since [DP TP] has been labelled *and* [DP TP] is a phase when C is deleted in syntax, the phi-features of the subject that contribute to the labelling of [DP TP] remain available even after extraction of the subject in a long-distance *wh*-fronting construction. So when C is deleted and T inherits phasehood, the labelling of [DP TP] is unaffected by extraction of DP — at least, within the confines of the subordinate clause.

  But what happens by the time the *wh*-subject reaches the matrix SpecCP? Setting aside the question (irrelevant here) of whether movement proceeds via the edge of the matrix $v$P, what we get is the structure below (46a), which should be compared to the one underneath (46b).

(46) a.  who do you think won the race?
     $[_{CP=\Phi}$ *wh* ... $[_{VP}$ V $[_{CP\neq\Phi}$ C$_\varnothing$ $[_{TP=\Phi}$ ~~*wh*~~ ...]]]]

  b.  *which race do you wonder John won?
     $[_{CP=\Phi}$ *wh* ... $[_{VP}$ V $[_{CP=\Phi}$ ~~*wh*~~ $[$ C$_Q$ $[_{TP}$ ...]]]]]

The example in (46b), the ungrammatical result of an attempt at extracting the specifier of a *wh*-question in the complement of a verb like *wonder*, illustrates what Rizzi (2013) has called 'the halting problem': once *which race* has arrived in the specifier position of the question selected by *wonder* (where it is grammatically spelled out in *I wonder which race John won*), it cannot proceed further into the matrix clause. Chomsky (2015) argues that by the time the matrix CP is reached, the silent copy of *wh* in the lower SpecCP in (46b) is invisible for the labelling algorithm. That causes the subordinate CP to be labelled by the feature content of C alone (i.e., by 'Q'), which identifies CP as a yes/no-question. The result is gibberish. Note that the embedded CP in (46b) is itself a phase (by hypothesis), and that ~~*wh*~~ is in the specifier position of that phase (by hypothesis), agreeing (again, by hypothesis) with C for the feature 'Q'. (This Q-agreement is what makes the embedded CP labellable in *I wonder which race John won*.) Apparently the fact that the silent copy of *wh* in the lower SpecCP in (46b) is present right at the phase level does not contribute to the preservation of its visibility: even though 'memory is phase-level', the silent *wh*'s features cannot be preserved for memory. When we now look at (46a), where we do want the phi-features of the silent copy of the *wh*-subject to be preserved for labelling purposes, we see that in all relevant respects the situation is the same as in (46b): the silent copy is at phase level in both cases; the next phase that is reached is in the matrix clause in both cases; yet (46a) is supposed to converge while (46b) does not.

  Note, however, that there are important reasons to believe that when the complementiser is omitted, a CP is still projected in the syntax. Not only would deletion of CP be a violation of structure preservation, it also would create trouble for the account of *want*-type infinitival complements with a lexical subject. For sentences such as *she wants/prefers (for) him to win*, Kayne (1984) argues cogently that the source of accusative case for the subject of the infinitive is C — whether it be overt (*for*) or silent. So in *who does she want/prefer (*for) to win?*, an infinitival 'complementiser-trace effect', we need CP to be projected even when its head is silent or, in Chomsky's terms, has been deleted in syntax.

This apparent contradiction regarding the visibility of silent copies of *wh*-extraction for labelling is reminiscent of the many, typically *ad hoc* and selective, appeals made throughout the history of principles-and-parameters syntax (incl. its minimalist phase) to the idea that traces lack features. At the end of the previous subsection, I pointed out that a positional approach to complementiser-trace effects must assume a condition to the effect that traces cannot satisfy the 'EPP'. The presumed selective inability of the *wh*-subject's trace to help label the [DP TP] structure downstairs is a particularly unpalatable one. There can be no reasonable doubt that *wh*-traces can have phi-features: they can control subject–finite verb agreement, participle agreement, and concord with secondary predicates, no matter how far away their antecedent might be, even in the presence of an overt complementiser in the lower clause (*which of these people do you think that for all intents and purposes <u>are</u> the strongest candidates?*).[12]

Be that as it may, the basic assumption that it is the shared phi-features of DP and TP that label [DP TP] structures, which is the cornerstone of Chomsky's (2015) proposal, is itself untenable. If filling SpecTP is to solve the labelling problem that is assumed to arise in [DP TP] structures, then T had better match the phi-features of the occupant of SpecTP. But examples such as those in (47) and (48) show that it is not obviously the case that the occupant of SpecTP needs to share phi-features with T.

(47)          I think that there are too many *ad hoc* assumptions in this analysis
(48)          I think that what hobbles this analysis are the many *ad hoc* assumptions

In (47) T's phi-features do not match those of its specifier (*there*, which is not itself endowed with a number feature) but instead those of the expletive's associate, further downstream. For this *there*-expletive construction, one might want to assume as a way out that *there* somehow matches the phi-features of its associate — notwithstanding the fact that *there* is plainly invariant on the surface: there is no discernible evidence for such feature matching, therefore, which raises the familiar acquisition and falsifiability issues. But the problem arises outside the realm of expletive constructions as well: for the pseudocleft in (48), the conclusion that the occupant of SpecTP does not need to have the same phi-features as those represented on T seems entirely inevitable (cf. *what you are eating is/\*are disgusting*: a free relative with *what* cannot by itself control plural agreement with T).

Apart from these issues, Chomsky's attempt to unify the 'EPP' and the 'ECP' also shares with earlier principles-and-parameters attempts the stipulation T has some sort of 'weakness'. This stipulation was particularly prominent in Chomsky (1986), where T has a whole compendium of handicaps: its projection is never an inherent barrier; its projection is never a minimality barrier; its projection does not allow intermediate adjunction. It resurfaces in Chomsky (2015) — though this time around, T is made out to be a member of a family of 'weaklings', including all acategorial roots.

---

12      The assumption that traces are, under certain circumstances, invisible to the labelling algorithm also fits in awkwardly with the standard-minimalist bottom-up derivational theory that treats movement as copying. Before the *wh*-subject is moved to its Ā-position in the left periphery of a higher clause, it occupies the SpecTP position of the embedded clause. It should *at that point* be able to help label the [DP TP] structure in the same way that non-*wh* subjects can. It should continue to be able to so do even after having moved away: the silent copy left behind in SpecTP should be different from its antecedent only in its lack of phonological features; it is not the phonological features that label the [DP TP] structure. The EPP may well be sensitive to PF properties (see the brief discussion in section 5.2.2). But it is hardly likely that it is the *phonological* features of the subject that label TP — esp. if phonological features are inserted only at PF.

Before closing this subsection, let us ask whether Chomsky's (2015) labelling approach could have anything to say about the Bresnan/Culicover facts. It can probably account for the obligatoriness of *that* in the case of subject extraction across a high adverbial, if adverbials like *for all intents and purposes* are placed in SpecTP, with subject extraction in the presence of such high adverbials taking place from the complement of T. According to Chomsky (2015), when C is deleted in syntax, CP is 'de-phased' and TP becomes the phase, with its complement transferred. So the subject cannot extract from the complement of T when TP becomes a phase due to C-deletion. In order to make subject extraction possible when *for all intents and purposes* occupies SpecTP, therefore, the complementiser must not be deleted. This would account for the obligatoriness of *that* in these cases. But note that *that* is obligatory whenever a high adverbial is present, regardless of whether or not the subject is extracted: *I think *(that) for all intents and purposes John Jones is the mayor of this town*. There is nothing that would seem to account for this. And besides, the presence of (non-agreeing) high adverbials such as *for all intents and purposes* in the specifier of the complement of C raises 'the labelling problem' again: how would the presence of a phi-featureless PP in the specifier of C's complement facilitate the labelling of that complement?[13]

All things considered, it seems to me highly unlikely that an approach to the 'EPP' and the 'ECP' (or, at least, the subject/object asymmetry emerging under extraction from CP) couched in terms of the labelling algorithm holds much promise. The desideratum motivating the labelling approach (viz., a unification of 'EPP' and 'ECP') is an important one, however.

In the next section, I will develop a novel analysis that ties 'EPP' and 'ECP' together in an integrated account of the spectrum of facts that need to be explained. The central role in this analysis will be played by an overarching licensing condition on specifiers, a condition which derives both the EPP (the structural subject requirement) and the subject/object asymmetry encoded in the disjunctive ECP, and in addition provides an account of the distribution of PRO and null operators.

## 5.3    A new analysis

### 5.3.1   *The theoretical core*

In a nutshell, my proposal is to unify the Case Filter and the EPP (the requirement that the structural subject position be filled), and to integrate the ECP and the EPP. It seeks to do so with the aid of the principle in (49).

(49)          *Specifier Licensing Principle (SLiP)*
              a specifier is licensed in an Agree relation with a probe

SLiP is a general principle governing the licensing of specifiers — specifically those of *functional* categories (which are the only kinds of specifiers under close scrutiny in this chapter). Its rationale is that, unlike their complements (whose presence is ensured by extended projection, *à la* Grimshaw 1991), specifiers of functional categories are never independently guaranteed to be present. Specifiers hence need to be legitimated: they can be present only when they are licensed to be there.

---

13       Note that labelling obviously cannot be done via matched phi-features: the high adverbial does not share any phi-features with T.

For our purposes in the following discussion, we will be particularly interested in two specific manifestations of SLiP — (50a), for A–specifiers, subsuming the bulk of the EPP and the Case Filter; and (50b), my restatement of the ECP in the realm of arguments.[14]

(50)  a.  *A–SLiP*          an *A*–specifier is licensed in an *Agree* relation with a *case*-probe
       b.  *T–SLiP*          a *trace* in a specifier position is licensed under *Downward Agree*

Agree plays a central role in the licensing of A–specifiers and traces in specifier positions. In conjunction with SLiP, I present the hypothesis in (51) regarding the workings of agreement relations in the tree.

(51)          *Agree*
       a.  RELATOR heads can establish two types of agreement relations at the same time
           *(i)*      *Upward Agree* (a.k.a. Spec–Head agreement) with their specifier
           *(ii)*     *Downward Agree* with their complement
       b.  non-RELATOR heads can only establish a *Downward Agree* relation with their complement

Before demonstrating how a theory incorporating (49)–(51) can derive the distribution of subjects as well as that of subject extraction, let me provide some background for (51), which will play a major role in the account.

### 5.3.2   *On Agree in a top-down grammar*

The structural relation called 'Agree' is the successor, in current minimalist syntax, to the relation of 'government' of earlier principles-and-parameters syntax. For government, definitions abounded, due to the fact that *(a)* the definition of c-command, an integral part of that of government, was unsettled, and *(b)* there was no consensus on the opacity factors reining in the extent of government relations. In today's literature, we see very much the same kinds of debates raging anew, this time about Agree: again it is unclear how far up the tree a head can look when searching for Agree-goals, and again there is a lack of consensus on the delineation of opaque domains ('phases'). There are those who maintain that Agree is altogether insensitive to the kinds of opacity factors constraining filler–gap dependencies (see e.g. Bošković 2003); and while many only allow Agree-probes to see potential goals in their complement domain (their strict c-command domain), for others specifiers are game as well. This time around, there are even syntacticians who want to confine the Agree-relation strictly to the head–specifier relation — i.e., *Upward Agree* as the only option (see e.g. Bjorkman & Zeijlstra 2014 *vs* Polinsky & Preminger 2015). The theoretical spectrum is bewildering, no doubt because the empirical palette of facts relating to the Agree relation is very colourful.
       In a top-down model of syntactic structure building, one might expect a head to try to establish Agree relations as soon as it emerges in the structure, and therefore to privilege as Agree-goals those categories that are already located in the tree at the point at which the head is merged. In a top-down model, specifiers are merged before complements are. So if Agree proceeds on a 'first come,

---

14      In the realm of non-arguments, 'ECP effects' are derived entirely independently in my theory. See chapter 3.

first served' basis, *Upward Agree* might be expected to be the norm in a top-down grammar. But if we operated on this logic, we would make all sorts of woefully inaccurate empirical predictions. Imagine, for instance, that we are building a root *wh*-question and have, at an early point in the structure-building process, constructed (52) as the partial structure of the root of the tree:

(52)          [$_{CP}$ which students [$_{C'}$ C

If the grammar is eager to establish an *Upward Agree* relation between C and its specifier, *which students*, at this point in the structure-building process, it will predict a plurally inflected finite verb in C, valuing its phi-features against the *wh*-phrase. The grammar will subsequently be at a loss accommodating a singular nominative subject in SpecTP, as in (53).[15]

(53)          which students is/*are he saying __ are coming to his party?

          In order not to fall into this trap, we will want the top-down grammar NOT to establish Upward Agree relations early and eagerly. Hasty establishment of Upward Agree relations is dangerous. We have to carefully restrict such relations. If the specifier of a head is a filler that cannot be interpreted as the subject or predicate of the complement of that head, it must not be allowed to engage in an Upward Agree relation with that head. Only the specifier of a head that mediates a direct predication relation with the complement of the head can be an Upward Agree-goal. Heads that mediate a direct predication relation are called RELATOR heads in Den Dikken (2006a). So a top-down grammar only allows Upward Agree relations between a RELATOR head and its specifier:

(54)          [$_{RP}$ XP [RELATOR [YP]]]

          A RELATOR head by its very nature has a structural relationship with both of the terms in its projection: its specifier and its complement. It can also show *agreement* with XP and YP at the same time. This we see in many environments — for instance, in an English finite clause such as (55), where the RELATOR (here 'T') agrees in phi-features (number and person) with the structural subject and the finite verb; or in an Italian small-clause construction such as (56), where the RELATOR agrees in phi-features (number, gender) with the subject *i ragazzi* and the secondary predicate *intelligenti*.

---

15          I made the subject of (53) explicitly nominative by picking the pronoun *he*. Kimball & Aissen (1971) point out that in north-eastern varieties of American English, a plural *wh*-phrase in SpecCP, or even a subpart thereof (see (ib)), can control phi-agreement with the finite auxiliary in C. As Kayne (1989) notes, such agreement is impossible when the subject of the clause is pronominal: (ia,b) become ungrammatical with *he* instead of *Muttonhead*.

(i)          a.          $^{\%}$what people do Muttonhead address his words to?
             b.          $^{\%}$to what people do Muttonhead address his words?

A popular approach to (ia) (following Kayne 1989) is to take agreement between C/T and the occupant of SpecCP to be the reflex of Spec–Head agreement (Upward Agree) in CP. Note that such an approach does not cover (ib), in which the controller of agreement is not the specifier of C: the *to*-PP is. It is likely, in light of (ib), that the Kimball/Aissen facts involve 'agreement attraction' of the type in (57b), below (Kimball & Aissen 1971:243 themselves note this link). Agreement attraction is not a(n ordinary) case of Upward Agree. Exactly what it does involve will be left open here.

(55)           [$_{TP}$ *he* [T=RELATOR [$_{VP}$ *smells*]]]

(56)           ritengo     [$_{RP}$ *i ragazzi* [RELATOR [$_{AP}$ *intelligenti*]]]
               consider.1SG the boys.M.PL          intelligent.M.PL
               'I consider the boys intelligent'

The agreement relation between a RELATOR and its specifier is customarily referred to in the principles-and-parameters literature as 'Spec–Head agreement'. Following the more recent literature (see e.g. Bjorkman & Zeijlstra 2014), I will refer to it here as *Upward Agree* (also sometimes called 'downward valuation'; cf. e.g. Polinsky & Preminger 2015). The agreement relation between a head and its complement, usually just called 'Agree', will be referred to here as *Downward Agree*, for appropriate contrast. Upward Agree is restricted to a closed class of heads: it is the prerogative of RELATOR heads. This a reflex of the special function that RELATOR heads have in the syntax: the fact that they are mediators of a relationship between the two phrasal categories that are its direct dependents (its complement and its specifier). It is thanks to their mediator function that these heads can establish an Agree relation with both of their dependents.

For any other functional head, the occupant of its specifier position is a filler that is not fully interpretable in this position: it must be placed on hold, on a stack, and will eventually be linked to a gap located in a position in which the constituent can be properly interpreted as an argument or predicate. With fillers that are placed on hold, no Upward Agree relations should be established, or else we run the risk of derailing the grammar in such cases as (53). Only in the case of RELATOR heads can we be safe with Upward Agree. As recovering from the wrong Agree relation is costly, the top-down grammar proceeds conservatively and takes only specifiers that are fully interpretable *in situ* (i.e., are not fillers placed on hold) to be Upward Agree-goals. RELATORs allow their specifiers to be interpreted *in situ*, so RELATORs can Upward Agree. Other heads cannot Agree with their specifiers.

### 5.3.3   The locality of Agree relations

The two kinds of Agree relations defined in (51) are both very local ones. Even for the most liberal kinds of heads, viz. RELATORs, Agree is strictly confined to their two *relata*: their specifiers and complements. And for other heads, their only possible Agree-goal is their complement.

The idea that a RELATOR can Upward Agree only with its specifier and not with a subconstituent of that specifier is entirely standard. Though 'agreement attraction' phenomena wherein the finite verb agrees in phi-features with a subpart of its subject certainly do occur frequently in the data, they are usually considered to be performance errors (whence the '!' in front of the example in (57b)) . It is likely that grammatical considerations come into the picture, making certain 'agreement attraction' cases more likely than others (see Kayne 1998, Den Dikken 2001). But under no circumstances will 'attraction' deliver an output that is superior to garden-variety specifier–head agreement, as in (57a).

(57)   a.       the key to the cabinets is missing
       b.       !the key to the cabinets are missing

For Downward Agree, on the other hand, it is by no means standard to think of it as being confined to the head–complement relation: 'Agree', as the successor of 'government', is usually thought of as a relationship that can look inside complements. An important ingredient of (51) is that it does not allow Downward Agree to go beyond the head–complement relation. But case and phi-agreement relations often seem to involve a relation between a functional head and something that is not the head's complement. In the realm of case, the relevant environments are routinely referred to as *exceptional case-marking* (ECM) environments; in the phi-domain, they come under the rubric of *long-distance agreement* (LDA). We will want to model these kinds of relations in our theory.

*5.3.4 Co-projection*

I allow for a way in which Downward Agree can involve a non-complement relation: *co-projection* of a functional head and a RELATOR head in its complement makes the latter's specifier a Downward Agree-goal for the higher functional head. Consider the structure in (58):

(58)         $[_{FP}$ F $[_{RP}$ XP [RELATOR [YP]]]]

When F and the RELATOR project independently, F can Downward Agree with RP, and the RELATOR Upward Agrees with XP and Downward Agrees with YP. F is unable to engage in an Agree relation with XP. Imagine now that F and the RELATOR in (58) join forces and co-project. The F+RELATOR combo Downward Agrees with RP. The RELATOR inherently Agrees with XP and YP, and F Downward Agrees with RP, which dominates XP (i.e., XP is below, *down*stream from, RP). So by transitivity the F+RELATOR co-projection complex now Downward Agrees with XP and YP.

Co-projection is obligatory for structures of the type in (58) when one of the two functional heads (F or the RELATOR) is unable to support an independent syntactic projection of its own, for instance because it cannot provide a label for its own projection. When it co-projects with a head in its immediate vicinity, a label for the co-projection results. Co-projection is a last-resort operation.

Co-projection is functionally similar to incorporation, and covers the effect of Baker's (1988) Government Transparency Corollary, given in (59).

(59)         *Government Transparency Corollary*
             a head that has an item incorporated into it governs everything which the incor-
             porated item governed in its original structural position

Consider, for instance, the 'possessor ascension' effect of incorporation of the possessed noun into a verb — as seen in the pair *they cuffed* [*his hands*] and *they* [*handcuffed*] *him* (in English such incorporation *cum* possessor ascension is highly lexically restricted, but this simple example makes the general point in an accessible way). The possessed noun would, in the absence of incorporation, govern the possessor, and license a particular case on it. But when the possessed noun incorporates into a structural-case assigning verb, one finds that its possessor is typically marked with the case that the verb would otherwise assign to the possessed noun phrase: the possessor 'ascends' to direct object, case-wise, because as a result of noun incorporation, the V+N complex ends up governing everything that the noun would otherwise govern by itself, including the possessor.

An Agree-based update of Baker's Government Transparency Corollary formulated in such a way that it stays maximally close to Baker's original (59) could read as in (59′):

(59)    *Agree Extension Corollary*
         a head that has a RELATOR co-projecting with it Agrees with everything which the RELATOR can Agree with by itself

In an Agree-based account of the link between noun incorporation and possessor ascension, what one would say is that the RELATOR that establishes the possession relation co-projects with *v*–V, as a result of which *v*–V can Downward Agree with the possessor in the RELATOR's specifier, and assign it accusative case. Whenever the possessum projects no more structure than a root-NP, as in (60), the RELATOR must co-project with *v*–V to license the root as part of the complex verb. Possessor ascension is the immediate result of such co-projection.

(60)    [$_{v/VP}$ *v*/V [$_{RP}$ POSSESSOR [RELATOR [$_{NP}$ POSSESSUM]]]]

Co-projection of *v*–V and the RELATOR in its complement also provides one of the ways in which 'exceptional case-marking' can be accounted for in the present system. For 'ECM accusatives' sitting in the specifier position of the small clause in the verb's complement, as in (61a), co-projection of the RELATOR-head of the small clause with *v*–V (cf. Stowell 1991 on 'small clause restructuring') delivers a Downward Agree relation between *v*–V and the small-clause subject, in constructions such as *I consider him clever*. For resultative constructions and ditransitives, it may sometimes be better to place the 'ECM accusative' in one of the object positions in the VP: the specifier-of-VP position, with the resultative secondary predicate occupying the complement-of-V position, as in (61b) (see the discussion in section 2.3.1, in chapter 2). Here, co-projection of *v* and V results in the subject of the resultative predicate being marked for accusative case by the *v*/V complex under Downward Agree.

(61)    a.    [$_{v/VP}$ *v*/V [$_{RP}$ SUBJECT [RELATOR [$_{XP}$ SECONDARY PREDICATE]]]]
         b.    [$_{vP}$ *v* [$_{VP}$ SUBJECT [V [$_{XP}$ SECONDARY PREDICATE]]]]

Within the Germanic language family, English and the Scandinavian languages are the only ones known to productively feature ECM in *to*-infinitival constructions — sentences of the type in (62a). In Dutch and German, this is strictly impossible (see Dutch (62b)). The Romance languages pattern with Dutch and German in this regard: as Kayne (1984) pointed out, unless some syntactic operation removes the ECM subject from the structural subject position of the infinitive, the Romance languages generally disallow ECM infinitives (see French (62c)).

(62)    a.    I considers John (to be) capable of doing such a thing
         b.    ik acht Jan in staat (*te zijn) om zoiets te doen          (Dutch)
               I consider Jan in state to be COMP such.a.thing to do
         c.    je crois Jean (*être) capable de faire une telle chose          (French)
               I believe Jean be capable of do.INF a such thing

If ECM is the result of co-projection of the functional head that harbours the subject in its specifier with the *v*–V above it, we can envisage a generalisation covering the distribution of ECM that makes reference to a particular RELATOR, viz., T: in some languages, T can co-project with *v*–V above it, in which case ECM-infinitives are grammatical; in others it cannot, which causes ECM to be restricted to environments in which the subject is the specifier of a small clause complement.

In this chapter, I will postulate one central typological hypothesis, regarding co-projection:[16]

(63)            *the T Co-Projection Parameter*
    a.        T can co-project: YES/NO
    b.        if YES,
              *(i)*      T co-projects with a head *above* it, or
              *(ii)*     T co-projects with a head *below* it

From (63), the cross-linguistic distribution of 'ECM infinitives' falls out as follows. English and the Scandinavian languages are among the languages that allow T to co-project with *v*–V above it, thereby facilitating the construction of ECM infinitives. Dutch, German, and the Romance languages, by contrast, do not allow T to co-project with *v*–V, and as a result exclude ECM infinitives that are structurally as large as TPs (for 'bare' infinitival small-clause constructions under perception and causative verbs, I assume that they are smaller than TP, although they may include functional material other than T).

Co-projection thus provides straightforward opportunities for accusative case-assignment to a non-sister — 'ECM' (assignment of accusative case to a niece) falls out as a corollary of co-projection, and variation in the realm of 'ECM' can be derived from (63). Long-distance agreement (LDA) relations might seem harder to model in this system because their name suggests that these are rather less local in nature. But as we saw already in the discussion in chapter 3 (section 3.3.1), it seems that LDA can be accommodated without compromising on absolute locality.

Co-projection involving T plays a role in a wide variety of contexts. One of them is complementiser omission, as we will see later in this chapter. Interestingly, for the set of languages reviewed above with respect to the licensing of ECM infinitives, complementiser omission in finite complement clauses makes essentially the same cut: again, English and the Mainland Scandinavian languages are on the 'YES' side,[17] and Dutch, German, and the Romance languages are in the 'NO' camp. This follows if complementiser omission (technically, the licensing of a null C-head) depends on co-projection of C and T: allowed in the former set of languages, but not in the latter. The distribution of *that*-trace effects also fits into this picture.[18]

---

16       The T Co-Projection Parameter is stated here as a 'macroparameter' (in the sense of Baker xxxx). It relates to the 'strength' or morphological 'richness' of T, which may facilitate a reworking of (63) in microparametric terms.

17       Roberts (2004:303): 'English and the Scandinavian languages allow for complementizer deletion in certain environments. (See Stowell 1981 for a discussion of the English environments; an anonymous reviewer points out that they are the same in the Mainland Scandinavian languages.)' On complementiser omission in Icelandic, see for instance Holmberg & Platzack (1995:110).

18       For discussion of the subtleties of Scandinavian *that*-trace effects (which lie beyond my expertise), see Bentzen's (2014) empirical survey.

### 5.3.5   *The EPP and cross-linguistic variation in the filling of SpecTP*

The T Co-Projection Parameter in (63) translates into the present model the common insight that there is a difference between English-type and Italian-type languages with respect to the licensing of the subject: in English-type languages (incl., besides English, at least the Scandinavian languages, but not Dutch or German), the subject must typically be licensed in SpecTP (i.e., the EPP is in effect); in Italian-type languages (incl. the Romance pro-drop languages and possibly also Modern Greek), the subject is licensed inside the complement of T (i.e., the EPP does not seem to be in effect). Both English-type and Italian-type languages answer 'YES' to (63a), but the two language types each pick a different option for (63b): English-type languages choose (63b.*i*) whereas Italian-type languages work with (63b.*ii*).

For the placement of the subject, the general result of (63b.*i*) in conjunction with A–SLiP in (50a) is that the SpecTP position must systematically be occupied — 'the EPP'.[19] The result of (63b.*ii*) in conjunction with A–SLiP is that the structural subject position (SpecTP) may remain unfilled. To see this, consider the structures below (where lexical subject is annotated as 'SU'):

(64)   a.       C         $[_{TP}$ <SU> [T $[_{RP}$ <SU> [R ...
                 *that*                 *is*
       b.       C         $[_{TP}$ <SU> [T $[_{RP}$ <SU> [R ...
                 ∅                      *is*

The subject in SpecTP, in both the a– and b–structures, is in an Agree relation with a case-probe, viz., T. For the licensing of the subject as such, this is sufficient: the orientation of the Agree relation does not matter for A–SLiP. All languages should allow the subject to be licensed in the structural subject position, SpecTP. But if the subject were placed in the complement of T, in the specifier of the RP (= *v*P or VoiceP[20]), it could be licensed only if T and R can co-project: R is not a case-probe for the subject, so an Upward Agree relation between the subject in SpecRP and R will not license the A–specifier of RP; but with T and R co-projecting, T is enabled to Downward Agree with the subject in SpecRP. The possibility of co-projection involving T is parametrically restricted: for languages that allow co-projection involving T in the first place, only a subset allows T to co-project with the head of its complement. Italian-type languages do; English-type languages do not. For the placement of the subject this means that in Italian-type languages, the subject can be below T, whereas in English-type languages, the overt subject must always be in SpecTP. This captures the 'EPP' and the cross-linguistic variation in this realm.

T is usually a RELATOR, and RELATORs need a specifier in order to be able to do their job of relating two phrasal terms. The requirement that SpecTP be projected and occupied — i.e., the classic 'EPP' — is not in effect when T co-projects with a RELATOR head: when such co-projection happens, the two co-projecting heads can share a specifier, either above or below T. This gives us precisely the kind of flexibility of subject placement desired for pro-drop languages such as Italian, where the subject can be positioned either in SpecTP or in the specifier in T's complement.

---

19      I will return to the EPP in control clauses, concentrating for now on clauses with an overt subject.

20      The choice between these two options is immaterial here, which is why I did not pick a specific label in (64).

### 5.3.6   Complementiser omission

The difference between the a– and b–structures in (64) is about the filling of the C position: it is overt in the a–structures and silent in the b–structures. Assume, as is plausible, that (65) is true:

(65)              a silent C cannot provide a label for its projection by itself

In order for the projection of a silent C to be labelled, therefore, C needs to co-project with a structurally adjacent functional head that can provide a label for the co-projection. In (64b), that functional head is T. So when C is silent, C and T must co-project, forming a C/T co-projection complex. When C is overt, by contrast, C has its own label and projects its own phrase; there is no co-projection in this case: co-projection is never our first resort.

      Going back once again to the parameter in (63), we see that we now derive a second difference between English-type and Italian-type languages as well: correlated to the fact that the latter do but the former do not allow the subject to be placed in the complement of T is the fact that former but not the latter allow the complementiser to be 'dropped'. Complementiser omission involves the silence of C, hence requires co-projection of C and a structurally adjacent functional head. The only candidate is T. Since T is structurally below C, co-projection of C and T will be possible only in languages in which T is allowed to co-project with a functional head *above* it. English-type languages are such languages — in English and the Scandinavian languages (recall fn. 17), complementisers can generally be omitted. Italian-type languages do not allow T to co-project with a head above it, and hence will systematically require overt complementisers that can label their projections by themselves.

      The T Co-projection Parameter in (63) typologically correlates the placement of the subject within the complement of T and the omissibility of complementisers. This correlation cuts across language families in an interesting way. Whereas English and the Scandinavian languages allow complementiser omission and generally obey the EPP very strictly, Dutch does not allow the complementiser to be omitted from their finite[21] non-root clauses and, concomitantly, has no trouble with SpecTP remaining unfilled: the subject of a Dutch clause can readily be placed in a position further downstream. The pair in (66a,b) illustrates the inverse correlation between complementiser omission and the requirement that SpecTP be occupied.

(66)   a.       ik denk *(dat) hij intelligent is
                   I think that he intelligent is
                   'I think that he is intelligent'
        b.       ik denk *(dat) <veel studenten> zich in het gebouw <veel studenten> verschansten
                   I think that many students REFL in the building many students entrenched
                   'I think that many students were entrenched in the building'

The fact that it derives this inverse correlation is an interesting bonus of the analysis, which was designed primarily to account for the distribution of subjects and subject traces.

---

21        The complementiser *om* of Dutch infinitival clauses is generally freely omissible. This is not the result of C/T co-projection (which Dutch lacks). I assume instead that it is a reflex of the variability in the size of infinitival clauses.

### 5.3.7   The generalised complementiser-trace effect

Let us return to the subject in the structures in (64a,b), repeated below. In (64a), the structural subject in SpecTP is 'only' in an Upward Agree relation with T; in (64b), where T and silent C must co-project in order for the latter to be able to label its projection, the structural subject in SpecTP is a Downward Agree-goal for the C/T co-projection complex.

(64)   a.       C        [$_{TP}$ <SU> [T [$_{RP}$ <SU> [R ...
                *that*                    *is*
       b.       C        [$_{TP}$ <SU> [T [$_{RP}$ <SU> [R ...
                ∅                         *is*

While for the licensing of the subject *per se* (i.e., for A–SLiP) the directionality of the Agree relation that the structural subject is engaged in is immaterial, it makes a crucial difference with respect to the *extractability* of the subject. The distribution of traces in specifier positions is governed by T–SLiP (50b), which says that a trace in a specifier position must be licensed under Downward Agree. For the structural subject in (64a), where T does not co-project, this means that it cannot be licensed as a trace at all: no trace is licensed in the structural subject position, SpecTP, when the complementiser is overt. This derives the classic *that*-trace effect, for both argument and non-argument traces in SpecTP: recall from section 5.2.4.1 that long-distance filler–gap dependencies involving predicate nominals behave exactly like similar such dependencies involving notional subjects when it comes to the *that*-trace effect. T–SLiP generalises over argument and non-argument dependencies footed in a specifier position, as desired.

As we saw in section 5.2.4.1, the *that*-trace effect has a direct counterpart in English infinitival clauses with the complementiser *for*. The *for*-trace effect in (28)–(29) follows from the theory in the same way as the *that*-trace effect. In the structures in (67), the occupant of SpecTP is allowed to be a trace only if it can be turned into a Downward Agree-goal — i.e., only when C and T co-project. Co-projection, in turn, is sanctioned only when C is silent. The general complementiser-trace effect has now been derived.

(67)   a.       C        [$_{TP}$ <SU> [T [$_{RP}$ <SU> [R ...
                *for*                     *to*
       b.       C        [$_{TP}$ <SU> [T [$_{RP}$ <SU> [R ...
                ∅                         *to*

English-type languages that have complementiser omission can get around the problem posed by the a–structures by resorting to silent C-heads, as in the b–structures. Here, C and T co-project, and as a result, the structural subject position is a Downward Agree-goal for C/T, allowing a trace to be postulated in this position.[22] Italian-type languages do not allow C and T to co-project, and hence cannot use the b–structures at all.

---

22      Recall that when T co-projects with another head, it does not have to project a specifier. But it always retains the right to project one. So co-projection of C and T does not preclude the exploitation of SpecTP, the structural subject position.

But Italian-type languages, which set (63a) to YES and (63b) to option *(ii)*, do allow T to co-project with the RELATOR in its complement, licensing the subject under Downward Agree inside the complement of T. And because the subject in this low position is a Downward Agree-goal, it can be a trace. So in Italian-type languages, a subject outside CP can bind a trace inside the complement of T. The result of Rizzi's (1982) original proposal has now been fully derived.

### 5.3.8 *The* do-*trace effect*

Koopman (1983) was the first to explicitly pursue a parallel between the complementiser-trace effect found in long-distance subject fronting constructions and the fact that in highest-subject root *wh*-questions no *do*-support obtains. (See also Pesetsky & Torrego 2001 for much further investigation, bringing novel data from Belfast English to bear on this parallel.)

(68)    a.       who *(did) John kiss?
        b.       who (*did) kiss Mary?

To understand the distribution of *do*-support in root *wh*-questions, consider (69):

(69)           $[_{CP}$ *wh* $[C [_{TP} T ...]]]$

In non-highest-subject root *wh*-questions, the occupant of SpecCP in (69) is not in an agreement relation with T. Hence T and C must not co-project in such constructions: when C and T co-project, SpecCP becomes an Upward Agree-goal for T, which results in a crash whenever what occupies SpecCP is not the structural subject of the clause that T heads. To make sure that co-projection does not take place, C must be overt. In a root clause lacking a finite auxiliary, the only way to ensure this is by inserting a form of the dummy *do* in C.[23] This accounts for *do*-support in non-highest-subject *wh*-questions in root contexts.

For highest-subject root *wh*-questions, the structure in (69) is also available. But unlike in the case of other *wh*-questions, co-projection is not only legitimate (after all, T and the occupant of SpecTP necessarily have matching features here) but in fact optimal: recourse to co-projection, which is a UG option, allows us to avoid exploiting a language-particular strategy (*do*-support).[24]

---

23    The dummy *do* is *inserted* in C in Subject–Aux Inversion constructions, and in this respect (which is the one that is relevant for us here) behaves just like a complementiser. The fact that the dummy *do* must be linked to T for its phi- and tense-inflection is irrelevant in the context at hand. Importantly, this linkage is not a case of co-projection. Co-projection takes place only between a featured head and a featureless one (such as between T and C$_\varnothing$). A C-head filled with the dummy *do* is not a featureless head: the dummy *do* has formal features. The T-head that the dummy *do* in C is linked to is likewise not a featureless head. So in a structure in which C=*do* is linked to the T-head below it, we are dealing with binding (possibly of a trace, but whether a trace is necessary in such highly local contexts is a question that deserves further discussion), not with co-projection.

24    In assessing the relative cost of co-projection and *do*-support, I follow the approach taken in Chomsky (1995: Chapter 2), where language-particular processes always come at a greater cost than operations that are part of UG.
    I note for completeness' sake that the structure in (69) is not the only logically possible one for highest-subject *wh*-questions: (i), in which the *wh*-subject is placed in the structural subject position (SpecTP) and does not occupy an Ā-position in the C-domain, is a logical candidate as well.

### 5.3.9    *The effect of a high adverbial on complementiser omission and subject extraction*

In the structures in (64), the structural subject finds itself, in English-type languages, in the structural subject position, SpecTP. We discovered in sections 5.3.7, for subordinate clauses, and 5.3.8, for root clauses, why (again, in English-type languages) this position can be occupied by a trace only when the C-position is silent. But in embedded finite clauses, insertion of a high adverbial (such as *for all intents and purposes*) to the immediate right of the finite complementiser changes this: now the complementiser is obligatory, and extraction of the subject across the complementiser is allowed. The structures in (70) sum this up.

(70)    a.        C          $[_{RP}$ AdvP $[R [_{TP}$ SU $[T$ ...
                  *that*                   ∅              *is*
        b.        *C         $[_{RP}$ AdvP $[R [_{TP}$ SU $[T$ ...
                  ∅                        ∅              *is*

        In (70), the subject or its trace occupies the structural subject position, SpecTP, as is customary (and forced by A–SLiP). But this time around, the structural subject position finds itself not in the immediate vicinity of C but instead below a silent RELATOR head that relates the high adverbial to the TP. This silent RELATOR has no label to give to its projection, hence must find some other head to co-project with. For English-type languages, T can play the role of co-projector for R.[25] With T and R co-projecting, the structural subject once again finds itself in a Downward Agree relation with T. This opens up the interesting prospect of grammatical subject extraction. And indeed, (70a) comes out grammatical in English: this is the Bresnan/Culicover effect of high adverbials, illustrated above in (27c,d).

        For Italian-type languages, there is not expected to be any special effect of a high adverbial on subject extraction from a finite clause: T will still co-project only downward, so the subject has to extract from within the complement of T, as always. There is indeed no equivalent of the Bresnan/Culicover facts for Italian-type languages. The analysis correctly derives this, too.

(i)                    $[_{TP}$ *wh* $[T$ ...]]

Chomsky (1986) couches a discussion of highest-subject questions (and relatives) in terms of the choice between (79) and (80), within the context of the Vacuous Movement Hypothesis. Chomsky (1986:50) suggests that the 'the language learner assumes that there is syntactic movement only when there is overt evidence for it', which may favour the non-movement derivation in (i); but he hastens to add that 'the unmarked case for a language with overt *wh*-movement is that it always takes place at S-Structure, so that nonmovement of subject ... would have a somewhat marked character' (p. 50). As this passage from Chomsky (1986) suggests, the question of whether the *wh*-constituent in a highest-subject *wh*-question is in SpecTP or *ex situ* is a theoretically and empirically highly complex matter. For our purposes here, making a choice between (69) and (i) is not strictly necessary. Regardless of which structure we assign to a highest-subject *wh*-question, the theory never expects such a question to feature *do*-support: (69) with SpecCP occupied by the subject of the clause gives rise to co-projection, hence silence of C; (i) treats the highest-subject *wh*-question as an ordinary subject-initial TP, which likewise lacks *do*-support (unless, of course, there happens to be a negation present). Either way, we derive the desired outcome. I will assume (69) because it treats all English root *wh*-questions uniformly.

25        The silent RELATOR in (68a) could in principle be licensed by co-projection with C; but C/R co-projection would not get a trace in SpecTP licensed: C cannot Agree with the subject without T's involvement.

The structure in (70b) is ungrammatical, regardless of whether the structural subject position is occupied by an overt constituent or a by a trace. Whenever a high adverbial occupies the SpecTP position, C must be overt. This follows straightforwardly from the theory. The silent C-head can only survive if it finds a structurally adjacent partner to co-project with. The only candidate in the b–structure is the RELATOR in its complement. But this head itself lacks a label, so it cannot help C out. The story ends there, at least when it comes to finite subordinate clauses.[26]

In *for-to* infinitival clauses, the counterpart to (70b) is of course ungrammatical as well, for the same reason that finite (70b) is bad. But in infinitival contexts, it turns out that placement of a high adverbial to the left of the structural subject is ungrammatical even in the presence of the overt complementiser *for*: (71a) contrasts sharply with (70a).

(71)   a.   *C        [$_{RP}$ AdvP [R [$_{TP}$ SU [T ...
                 *for*              ∅              *to*
       b.   *C        [$_{RP}$ AdvP [R [$_{TP}$ SU [T ...
                 ∅                 ∅              *to*

A–SLiP accounts for the ungrammaticality of (71a), via case. Whereas the T-head in finite clauses introduces a nominative case feature all by itself, the T-head of infinitives is not inherently capable of case-licensing its subject: a T-head spelled out as *to* has a latent case feature that can only be activated in a relation with a local C that has a case feature. In the structures in (67), the infinitival C is indeed local to T=*to*, thereby activating T's latent case feature. But in the structures in (71), a head intervenes between C and T. The intervention of this R-head obstructs the relation between C and T that is necessary in order for T to get its case-assigning capacity activated. As a consequence of the fact that the infinitival complementiser *for* cannot itself Agree with the case-dependent subject (because non-RELATOR heads can only establish sisterhood Agree relations) and the T-head below C cannot be activated for case, (71a) leaves the subject without case, in violation of the Case Filter (or the Visibility Condition, subsumed under the Principle of Full Interpretation) — which is now incorporated into A–SLiP (50a), the very same condition which also takes care of the EPP.

### 5.3.10  The effect of a high adverbial on C–T chain formation

In root questions, high adverbials are also banned categorically — as in infinitival clauses with an overt subject, but quite unlike what we found in finite subordinate clauses. The illegitimacy of insertion of a high adverbial between C and TP holds for all root questions: yes/no- as well as *wh*-questions, and subject as well as non-subject questions. In root highest-subject questions, it manifests itself regardless of whether the dummy auxiliary *do* occurs in C (to the left of the adverbial), in R (to its right), or in T (to the right of the subject). The facts below illustrate:[27]

---

26      Note that, in the top-down approach to syntactic derivation pursued here, co-projection of R and T cannot benefit silent C because it comes too late: silent C needs to be licensed by the head of its complement, which is itself unlabelled at the point at which it first enters the structure.

27      In assessing the status of some of the strings in (72), one should control for parenthetical uses of the adverbial.

(72)  a.       for all intents and purposes he didn't know this
      b.       *didn't for all intents and purposes he know this?
      c.       *what <didn't> for all intents and purposes <didn't> he <didn't> know?
      d.       *who <didn't> for all intents and purposes <didn't> know this?

In non-root questions, insertion of a high adverbial between C and TP is generally possible, though two *caveats* need to be made: (73b) is somewhat worse than (73a) and (73c), probably due to the fact that *wh*-dependencies across a fronted topic are also slightly degraded (*?I have no idea what to Mary we should give for her birthday*); and for (73c) it is impossible to tell where exactly *in all likelihood* is inserted (in view of the fact that this adverbial can also occur between a subject in SpecTP and the finite auxiliary, as in *they say that he for all intents and purposes didn't know this*).

(73)  a.       I have no idea if/whether for all intents and purposes he didn't know this
      b.       ?I have no idea what for all intents and purposes he didn't know
      c.       I have no idea who for all intents and purposes didn't know this

It appears that the insertion of RP in between C and TP is ungrammatical in root questions — that is, the structure in (74) is ill-formed when unembedded.

(74)          $[_{CP}$ (*wh*) $[C_Q$ $[_{RP}$ AdvP $[R$ $[_{TP}$ SU $[T$ ...
              → * in root questions; ✓ in non-root questions

Since the ban on insertion of a high adverbial between C and TP in questions is apparently restricted to root contexts, it cannot be that C's interrogative illocutionary force *per se* blocks (74). A perspective on the picture in (74) emerges when we realise that in root questions, unlike in non-root ones, head movement from T to C takes place — or, put differently (in a way that is more in line with a top-down approach), a chain is formed between C and T.[28] I assume (in line with Chomsky 1995: chapter 3, in his discussion of Holmberg's Generalisation) that head chains can have at most two members. But in structures of the type in (74), with the high adverbial signalling the presence between C and TP of an RP whose head must co-project with T, the head chain between C and T would inevitably have *three* members: C, R, and T. On the hypothesis that such a chain is ill-formed, the ungrammaticality of (74) in root questions of all stripes can be understood.[29]

---

28       I assume that such a chain is formed in *all* root questions, including highest-subject ones.

29       In blaming the ill-formedness of (74) in root contexts on the formation of a chain between C and T, the text proposal aligns itself in essence with what Kayne (1984) suggested as an account for the ungrammaticality of T-to-C movement in the presence of adjunct material to the left of the subject.
         Note that the contrast between (72) and (73) is likely to be problematic for Pesetsky & Torrego's (2001) analysis of the complementiser *that* as the product of T-to-C movement. In the presence of familiar T-to-C movement cases, insertion of a high adverbial between C and the subject in SpecTP is systematically ungrammatical. The fact that (72b–d) are ungrammatical while the examples in (73) are well-formed is compatible with Pesetsky & Torrego's treatment of *that* only if the deviance of (72b–d) is unrelated to T-to-C movement. With T-to-C excluded as a central ingredient in the account of (72b–d), it is unclear what the cause of their ungrammaticality might be.

*5.3.11  On PRO*

The EPP appears to be systematically violated in control infinitives, which have a silent subject (customarily identified as 'PRO') for which Baltin (1995) argues cogently that it does not raise to SpecTP. Baltin's argument to this effect is based on the important observation that while (75b) and (75c) are both grammatical with *all* to the left of *to*, (75a) is not:

(75)    a.        these people tried <*all> to <all> leave
        b.        these people seemed <all> to <all> leave
        c.        I would prefer for these people <all> to <all> leave

The subject of a control infinitive, in contradistinction to the subject of a raising infinitive or a *for-to* ECM infinitive, cannot be associated with a 'floating quantifier' to the left of the infinitival marker *to* (though placing the quantifier to the right of *to* is grammatical, in all three cases). Baltin's plausible interpretation of these facts is that while the infinitival SpecTP position is occupied (at some point in the derivation) in (75b,c), making it possible for *all* to show up to the left of *to*, in the control infinitive in (75a) it remains unfilled, causing pre-*to all* to be unlicensed.

        These facts raise the important question of how the EPP can allow for systematic exceptions. The current theory recasts the EPP partly via SLiP, given in (49) (repeated below), and partly as a result of the fact that T, a RELATOR, requires a specifier whenever it does not co-project.

(49)            *Specifier Licensing Principle (SLiP)*
                a specifier is licensed in an Agree relation with a probe

I propose that it is PRO's very nature that makes it appear precisely in specifier positions that are unlicensed in the sense of SLiP:

(76)            PRO is an unlicensed argument

The statement in (76) is the equivalent of the PRO Theorem of earlier principles-and-parameters theory, in which 'government' was a key player. With 'government' replaced with Agree, and SLiP being the successor of the Extended Projection Principle, (76) is the PRO Theorem's natural successor, though the theory in its present state does not yet derive (76) as a theorem.

        PRO by definition occurs in unlicensed positions. Any argumental noun phrase occupying a *licensed* specifier position will be either overt (in which case A–SLiP must be obeyed) or a trace (subject to T–SLiP). Precisely because PRO is by definition an unlicensed argument, SLiP and (76) collaborate to ensure that PRO is never found in a licensed specifier position: placing PRO in a licensed specifier will 'blow it up'. The SpecTP position, whenever projected, is always a licensed specifier position: it is engaged in an Agree relation with T, a probe. Since SpecTP is always a licensed specifier in the sense of SLiP, PRO is strictly prevented from occurring in SpecTP positions. This now derives Baltin's (1995) facts in (75). While the EPP would at best *allow* PRO *not* to raise to SpecTP, it cannot straightforwardly *prevent* PRO from raising there. SLiP in conjunction with (76), on the other hand, ensures that PRO cannot raise to the structural subject position.

The fact that in control infinitives, the structural subject position cannot be occupied entails that the structural subject position must not be projected: if it were projected but would remain un-occupied, T qua RELATOR would be unable to do its work. Since T must not have a specifier in con-trol infinitives, it must co-project with C. A non-co-projecting T must project a specifier, harbouring the structural subject. But when T, which usually serves as the RELATOR of the primary predication relation in the clause, co-projects with a *non*-RELATOR, and the latter is higher up the tree (i.e., comes in first, in the top-down structure-building process), T can forgo a specifier — succinctly put, the C/T co-projection complex can behave like C (and have no RELATOR function) or like T (and intro-duce a specifier), as needed. So in the case of control infinitives, where T wants no specifier because nothing could legitimately occupy it, C/T co-projection is called for. And since C/T co-projection is legitimate only when C is silent, this means that C must be structurally present in English-type control infinitives but has to be silent.[30] On an analysis based on the PRO Theorem, control infinitives must also be null-headed CPs, in order for PRO to be 'shielded' from government by an outside head. In the analysis of control infinitives presented here, their structure (a null-headed CP dominating a TP without a specifier, with PRO within the complement of T) follows from the combination of the definition of PRO in (76), SLiP in (49), and the general role of T as a RELATOR of the primary predication relation in the clause: T projects and licenses a specifier unless it co-projects with $C_\varnothing$.

### 5.3.12 Null operator relatives and the 'anti-that-trace effect'

The null operator of *that*-relatives, *tough*-movement constructions, and parasitic gap constructions is best identified as PRO: because null-operator subject relatives can be infinitival (*a man to fix the sink*), with no case available for the null operator, there is really no other suitable candidate.[31] Assume, therefore, that the null operator is PRO. With this assumption in place, an explanatory perspective emerges on one of the most perplexing facts from the realm of Ā-dependencies: the fact that while subject extraction from a finite clause into a higher clause resists the presence of an overt complementiser (the '*that*-trace effect', analysed above), highest-subject null-operator relatives force the complementiser to be overt (the 'anti-*that*-trace effect', illustrated in (77a)).

(77)    a.      the man *(that) introduced me to her is John Jones
         b.      the man  (that) she introduced to me is John Jones
         b.      the man  (that) she introduced me to is John Jones

---

30      Note that Dutch is not an English-type language for the T Co-Projection Parameter: Dutch T looks downward rather than upward for co-projection purposes. So non-projection of SpecTP does not lead to obligatory silence of C in control infinitives in Dutch. Indeed, the Dutch infinitival complementiser *om* does occur in control infinitives.

31      A treatment of null operators as PRO (i.e., as an unlicensed argument) also accounts straightforwardly for the well-known fact that null operators strongly resist being non-arguments (*the best candidate is not easy Op (for anyone) to be*), non-nominal arguments (*on government support is not easy Op (for anyone) to depend*), pied-pipers of P (*government support is not easy [on Op] (for anyone) to depend*), or indirect objects (*these people are easy Op to sell this book *(to)*). See Den Dikken (1995) for an explanation for the pied-piping and indirect object restrictions on null-operator dependencies in which the status of the null operator as PRO and a particular analysis of the dative alternation are instrumental.

In the structure in (78), the null operator finds itself in SpecCP, its usual position. When C is filled with *that*, SpecCP is an unlicensed specifier in the sense of SLiP: it is not an Agree-goal because C, which is not a RELATOR, cannot establish an Upward Agree relation with its specifier. For PRO, which does not want to be in a licensed specifier position (by (76)), the fact that SpecCP is an unlicensed specifier when *that* is present is a godsend: it is precisely thanks to the fact that the specifier of an overtly-headed CP is an unlicensed specifier that PRO is allowed to occur there.[32]

(78)              $[_{CP}$ *Op*=PRO $[C=\{that/*\varnothing\}$ $[_{TP}$ ...]]]

We have now ensured that the representation in (78) is grammatical with *that* in C. When C instead remains silent, it has to be licensed as such. Co-projection is the way to license $C_{\varnothing}$ when the occupant of SpecCP has phi-features that match those of T (i.e., when the occupant of SpecCP is the subject of the clause itself): the fact that T and the operator in SpecCP have the same phi-features makes it legitimate for T to co-project with C and to establish Upward Agree relation with the operator. However, for CPs whose specifier is occupied by a *null* operator (i.e., PRO), co-projection of C and T causes the derivation to crash: as a result of C/T co-projection, PRO ends up in a licensed specifier position, precisely where it does not want to be. This draws the curtain on the version of (78) in which C is silent.

We do not, of course, want the grammar to rule out null-operator relatives that lack the complementiser *that* across the board. Non-highest-subject relatives with a silent operator allow *that* to be omitted freely, as shown in (77b,c). Moreover, there are varieties of English in which highest-subject relatives exhibit no 'anti-*that*-trace effect', allowing 'subject contact relatives'. Appalachian English is a case in point:

(79)    a.        at first, you wouldn't believe the characters come knocked on my door
         b.        but he tied the company up some way to get a royalty off the timber was cut for the
                  mines                                                              (Appalachian English)

The structurally minimal representation of contact relatives involves a TP, to which non-subject operators can adjoin, as in (81a). Non-subject contact relatives like those in (77b,c) are fine with the null operator (PRO) adjoined to TP: the adjunction position is not in an Agree relation with T; PRO is not in a licensed specifier position. The subject contact relatives in (79) likely have an even barer structure. In (81b), the null operator occupies an unlicensed specifier position in the complement of T (Spec*v*P, on common assumptions).

(80)    a.        $[_{TP}$ *Op*=PRO$_{NON\text{-}SU}$ $[_{TP}$ SU $[T$ ... *ec* ...]]]
         b.        $[_{TP}$ T $[_{vP}$ *Op*=PRO$_{SU}$ $[v$ $[_{VP}$ ...]]]]

---

[32]      In a highest-subject *that*-relative, the null operator arguably does not bind a gap lower in the structure. It is interpreted as the subject of predication thanks to C serving as the RELATOR of subject and predicate, and the relative CP as a whole serving as the predicate of the 'head' of the relative (with which the null operator is co-indexed).

While (80a) is grammatical throughout the English-speaking world, a 'bare TP' structure with the subject below T, as in (80b), is ruled out for standard English, which generally does not allow subjects below TP in the absence of an expletive. But Appalachian English has been argued by Tortora & Den Dikken (2009) to allow the subject to be in a position in the c-command domain of T, not engaging in a φ-Agree relation with T (as in 'anti-agreement' constructions such as *those boys <u>is</u> plumb foolish* and *the potatoes <u>looks</u> awful*).

What is left to deal with in connection with contact relatives is the fact that *infinitival* highest-subject contact relatives (*a man to fix the sink*), unlike finite subject contact relatives, are grammatical in standard English. Law (1991) argues that infinitival subject relatives are TPs, with PRO in SpecTP. In light of the discussion of Baltin (1995), this exact analysis is not available to us. But the structure in (80b) may well be available. A non-co-projecting *finite* T necessarily formally licenses a specifier in standard English, so (80b) is ungrammatical there when T is finite. But it may be *infinitival* T (perhaps *non-finite* T more generally) is only latently a RELATOR, and can only be activated as one by the head that takes TP as its complement (the complementiser *for* in *for-to* infinitives, which have the subject in SpecTP). If this is right, (80b) converges in standard English as an infinitival relative.

### 5.3.13 Long null-operator dependencies

In the previous subsections, I laid out a theory of the licensing of subjects and their traces that provides an account for complementiser-trace effects and their suspension. It addresses the core facts presented in section 5.2.4.1 — that long extraction of the structural subject is sometimes contingent on the silence of the complementiser (as shown by the a– and b–examples in (27) and (29), repeated below), and that silence of C is not a necessary condition on long extraction of the structural subject (as demonstrated by the c– and d–sentences in (27)) — and presents an explanatory perspective on them.

(27)    a.    John Jones, who everybody thinks (*that) is the mayor of this town, must be a hero
         b.    the mayor of this town, which everybody thinks (*that) is John Jones, must be a hero
         c.    John Jones, who everybody thinks *(that) for all intents and purposes is the mayor of this town, must be a hero
         d.    the mayor of this town, which everybody thinks *(that) for all intents and purposes is John Jones, must be a hero

(29)    a.    John Jones, who everybody prefers/wants (*for) to be the mayor of this town, must be a hero
         b.    the mayor of this town, which everbody prefers/wants (*for) to be John Jones, must be a hero

But in order to get a comprehensive sense of the licensing restrictions on subjects and their traces so as not to fall into the trap of devising an account for what might very well be the tip of an iceberg of restrictions on long subject extraction from CP, we ought to amplify the empirical picture to include cases in which such extraction fails even in the absence of an overt complementiser — cases, in order words, in which silence of C is not a sufficient condition for extraction of the subject.

Such cases are rarely mentioned, let alone explained, in the literature on complementiser-trace effects. There are three types of them, found in *(a)* long null-operator movement constructions (Schachter 1981, Taraldsen 1981, Chomsky 1982:53, Cinque 1990:105–6, Stowell 1986, 1991), *(b)* sentences in which a complement clause occurs alongside a nominal dependent of a particular class of matrix verbs (Stowell 1981), and *(c)* cases of subject extraction from an island (Pesetsky 1982: 581, Cinque 1990:105–6). I will begin my discussion with *(a)*, the simplest case.

Null operators in *tough*-movement and parasitic gap constructions strongly resist binding a trace in the subject position of a finite clause. Thus, see the contrast between (81a), an *it*-expletive construction not featuring *tough*-movement, and (81b), its *tough*-movement counterpart:

(81)    a.        a painting which *it* is easy to think/believe could be by Leonardo da Vinci[33]
         b.        *a painting which is easy to think/believe (that) could be by Leonardo da Vinci[34]

The ungrammaticality of (81b) is not due to some general constraint barring long-distance null-operator dependencies across a finite clause boundary: (81b) is much worse than (82).[35]

---

33        This example was modelled on a sentence from Samuel Butler's (1888) *Ex voto*. That the possibility of long extraction of the subject of a finite clause from an *it*-expletive construction is by no means a thing of the past is shown by the following attestations of this pattern on the internet:

(i)    a.        this is a bit of a problem, and one that it's easy to think is a below-the-belt way of causing problems for competing companies
       b.        one of those problems that it is easy to think is only happening somewhere else
       c.        Word shows you a sample sheet that it is easy to think is the first sheet of the merged data
       d.        this is the kind of customer service that most companies only dream of (and that it's often easy to think is nowhere to be found today)
       e.        staying aware of what's going on behind you and taking a look besides or behind what it is easy to think is the main subject
       f.        Bonaventure allows for a theory of knowledge which it is easy to think is diametrically opposed to that which comes from Plato

34        Sentences of this type are not non-existent (see e.g. the examples below, culled from the internet). But native-speaker judgements indicate that they are highly marked relative to (84a). Stowell (1986) gives examples of the type in (81b) two stars (his example is *Betsy is easy to expect fixed the car*); Stowell (1991:204) gives them one star (here his example is *John is impossible to believe could have written this book*). I will follow him in taking sentences of the type represented by (81b) to be ungrammatical.

(i)    a.        an auto-upgrade process is one of those things that would be easy to think is essential when you're paying for a product
       b.        it helps me cut through all the nonsense and pettiness that is so easy to think is important
       c.        why is domestic violence something that is so easy to think is someone else's problem?
       d.        the 2nd one would be easy to think is real
       e.        one of the things that is easy to think is going wrong but has just not been done correctly is setting the phone to emergency mode
       f.        when a player tries to have his arms behind him as well as the club, which is easy to think is right in trying to get it back there and behind you

35        Stowell (1991:204) points out that null-operator dependencies are quite generally reluctant to cross a finite clause boundary, but they are marginally acceptable when the finite clause contains a modal. So (82) deserves at least a question mark. But the relevant point is that it is quite a bit better than (81b).

(82)                   [?]a painting which is easy to think/believe Leonardo da Vinci could have painted

Nor, as it turns out, is the problem with (81b) indicative of a blanket ban on long subject extraction in null-operator movement constructions: though the literature has not noticed this before to my knowledge, the Bresnan/Culicover effect that famously alleviates subject extraction across the finite complementiser *that* is in evidence in *tough*-movement constructions:

(83)                   [?]a painting which is easy to think/believe that in all likelihood is by Leonardo da Vinci

The *tough*-movement facts do not stand on their own: parasitic gap constructions, which also involve null-operator dependencies, behave the same way (see Taraldsen 1981, who, like Stowell, does not mention the effect of a high adverbial):

(84)    a.    [?]which painting did you buy without thinking that Leonardo da Vinci could have painted?
        b.    *which painting did you buy without thinking (that) could have been painted by Leonardo?
        c.    [?]which painting did you buy without thinking that in all likelihood was painted by Leonardo?

In (84) we see that parasitic gaps can fairly easily be objects of subordinate finite clause but that they can be subjects of finite clauses only as long as there is a high adverbial present.

Taken together, the pair in (81b) and (83) and the pair in (84b,c) present a generalisation covering null-operator movement constructions (*tough*-movement and parasitic gap constructions). We now need to find a theoretical expanation for this empirical generalisation. The central question is why null-operator dependencies generally resist involving the subject of a finite clause.

Stowell (1986, 1991) presents an ECP-style account for the fact that null operators cannot be subjects of finite clauses. The proposal astutely exploits the disjunctive ECP (lexical government versus antecedent-government), and assumes that because their antecedents are null, traces of null operators must be lexically governed: null antecedents are incapable of serving as proper antecedent governors. Not only does such an account beg the question of why null antecedents cannot be proper antecedent governors, it also banks rather too heavily on the disjunctive ECP. And the account also suffers from the empirical shortcoming that it does not capture the Bresnan/Culicover effect seen in (83) and (84c).

Cinque (1990), like Stowell, argues that the null operator environments under discussion do not allow a *wh*-trace to be left behind by movement. But for Cinque, that is not the end of the story — for he allows Ā-dependencies to be established, in principle, via a binding relation between an operator and a *pro* in the position of the gap. To radically rule out the examples in (81b) and (84b), Cinque must in addition ensure that this *pro*-binding strategy is blocked here as well. He does so by assuming that C 'is not an appropriate licenser' of resumptive *pro* (Cinque 1990:120). This could help explain the contrasts between (81b) and (84b), on the one hand, and (83) and (84c), on the other, if the resumptive element in the latter examples is not in the structural subject position but instead in a position in which this *pro* can successfully be licensed. But the idea that resumptive *pro* cannot be licensed in the structural subject position of a finite clause itself remains an explanandum.

We plainly could not claim in general that *pro* cannot be licensed in the subject position of a finite clause: otherwise, the quintessential locus for referential *pro* in null-subject languages would suddenly become *terra prohibita* for it. So the claim that resumptive *pro* cannot be licensed in the structural subject position of a finite clause does not follow from anything we know independently, and makes resumptive *pro* suspiciously different from familiar instances of *pro*, depriving the account of explanatory adequacy. We are in need, therefore, of a different, more descriptively and explanatorily adequate account.

A vital step towards a new account is the treatment of null operators as PROs, put forward in section 5.3.11. This treatment of null operators, in conjunction with the hypothesis that PRO is an unlicensed argument (i.e., (76), repeated below), prevents null operators from being highest subjects of finite clauses lacking a complementiser (already discussed in section 5.3.11), and also explains and the contrasts between (81b) and (83) and between (84b) and (84c). To see how these contrasts fall out, we need to be a more precise about the nature of null-operator chains.

(76)         PRO is an unlicensed argument

Operator–variable chains are standardly modelled, in the minimalist programme, as sets of identical copies of the filler, with the 'trace' position occupied by a full copy of the operator. This copy usually remains entirely silent at PF: it is reduced to naught. But at LF as well, the copy in the 'trace' position usually needs to be reduced: otherwise we would end up with an overly rich semantic representation. What we need is for the copy in the operator position to be reduced to just the operator, and the one in the 'trace' position to be reduced to just the restriction. This copy reduction process is necessary for semantic convergence in the case of operator–variable dependencies headed by an operator that has internal structure — basically, any such dependency involving an *overt* operator will have to go through it. But the *null* operator lacks internal structure altogether: it is a bare pronoun, PRO. The syntactic representation of a chain involving null operator movement, viz., (PRO, PRO), can be preserved exactly as is and be handed over to the semantic component of the grammar, with the lower copy interpreted as a *pronominal variable*. Since conversion of the lower copy into a trace would be costly (because it would involve the application of an operation), and since such conversion is entirely superfluous in the case of null-operator dependencies, it follows that whenever PRO is in an Ā–specifier position (a 'null operator'), PRO's lower copy is not reduced or converted into a trace; so a PRO-headed chain looks like this: (PRO, PRO).

This has four important consequences. First, like Cinque's (1990) proposal, it correctly predicts the absence of weak cross-over effects with null-operator movement (Lasnik & Stowell 1991): while (85a,b) are famously ungrammatical with coindexation of *his* and *who/everyone*, (86a,b) (featuring *tough*-movement and a parasitic gap, respectively) are fine with such coindexation.

(85)   a.      *who$_i$ does his$_i$ mother love?
       b.      *his$_i$ mother loves everyone$_i$
(86)   a.       who$_i$ will be easy for us to get his$_i$ mother to love?
       b.       who$_i$ did you meet before his$_i$ mother loved?

In this regard, null-operator movement constructions behave just like (87a,b), where *him* is a pronominal variable, and coindexation of *who/everyone*, *his* and *him* is perfectly fine.

(87)  a.    who$_i$ thinks that his$_i$ mother loves him$_i$?
      b.    everyone$_i$ thinks that his$_i$ mother loves him$_i$

The parallel behaviour of (86) and (87) suggests that the empty category bound by a null operator is like more like a pronoun than like a trace. The analysis of PRO-chains according to which the copy in the extraction position is not converted to a trace but remains PRO (a pronominal empty category) ensures this.[36]

A second consequence of the hypothesis that the gap bound by a null operator is PRO rather than a trace is that null-operator dependencies should not be allowed to occur in what Postal (1998) calls 'anti-pronominal contexts' — contexts from which pronouns are excluded. Here again, the present analysis of null-operator dependencies is neatly aligned with Cinque's (1990). Cinque points out that null operator dependencies, surprisingly at first blush, resist the direct object function when the indirect object is a full noun phrase (Lasnik & Fiengo 1974, Cinque 1990:108–9):

(88)  a.    *$^?$books are not easy to give that man
      a′.   $^?$books are not easy to give 'm
      b.    *$^?$the book that we filed instead of giving that man
      b′.   $^?$the books that we filed instead of giving 'm

The fact that the gap in null-operator dependencies (*tough*-movement or parasitic gap constructions) cannot be located in the direct object position of a double object construction can be related directly, on Cinque's (1990) assumptions as on mine, to the fact that overt (unstressed) pronouns are generally difficult in the direct object position of (English) double object constructions (cf. *he gave his friends it*), except if the indirect object is also a pronoun (cf. $^?$*gimme it*, $^{?(?)}$*he gave 'm it*), in which case null-operator movement dependencies also improve (see (88a′,b′)). Assuming that what is true of overt (unstressed) pronouns is true of Cinque's resumptive *pro* and my PRO as well, (88) follows.

---

36    The fact that PRO does not normally occur in object positions is due to the fact that PRO cannot check case. This in turn can be made to follow from (76): the positions that PRO is allowed to appear in are unlicensed positions; unlicensed positions are positions in which no phi- or case-features can be checked. For the object, this is the complement-of-V position in a structure in which *v* and V do not co-project: *v*, the case licenser, cannot Agree with PRO in this configuration. In Ā filler–gap dependencies, the case requirements that are imposed by the predicate in the absence of movement are often suspended. We see this, for instance, in (i) (Postal 1974, Kayne 1984) and in (ii) (less widely discussed).

(i)   a.    I wager this man to be the winner
      b.    which man do you wager to be the winner?
(ii)  a.    I'm not sure *(of/about) this thing
      b.    the (only) thing I'm not sure (of/about) is this

While the matrix verb in (ia) is incapable of case-licensing the subject of the *to*-infinitive in its complement, and the adjective (iia) requires a preposition to case-license the object, we see in the b–examples that no case effects emerge when the noun phrase in question occurs in an Ā-position. (The pattern in (iib) is common in spoken English (a search for the string "thing I'm not sure is" delivers an abundance of hits on the internet); it would probably be best classified as an error, but the important point about (ii) is that in the a–example omission of the preposition would be entirely impossible.) In light of the suspension of the case requirements in (ib) and (iib), I hypothesise that PRO can occur in object positions as long as it is Ā-bound.

Thirdly, the fact that the lower copy of a null operator is not converted into a trace means that T–SLiP, the principle governing the distribution of traces, is entirely inactive in the case of null-operator dependencies: since the lower copy of the null operator is not converted into a trace, there is nothing that T–SLiP has jurisdiction over in the case of null-operator constructions.

The fourth consequence of the fact that the lower copy of the null operator is not converted into a trace but remains as 'PRO' is that *no member of the chain of null-operator movement can be in a licensed specifier position*: it is not just the copy of PRO in the operator position that must make sure that its specifier position is not a licensed one; the copy of PRO in the extraction position also must not be in a licensed specifier position. Put succinctly, (76) applies to *all* members of a PRO-headed chain. This entails that *(a)* PRO's Ā–specifier position must not be an Agree-goal (SLiP) and *(b)* since all the members of the chain are PRO's (none of them is converted into a trace), PRO's chain must not have any member in an Agree relation with a case-probe (A–SLiP).[37]

This helps us account for the data presented at the beginning of this subsection: the fact that null operators can serve as subjects of subordinate finite clauses only in the presence of a high adverbial. A null-operator dependency footed in a SpecTP immediately below *that* is bad because a non-co-projecting T requires SpecTP to be occupied (recall the discussion of the EPP earlier in this chapter: section 5.3.5), but the postulation of a lower copy of PRO in SpecTP would not be in conformity with (76).[38] In the absence of *that*, T co-projects with C. The C/T co-projection complex includes a RELATOR (viz., T), and RELATORs must, by their nature, relate two phrasal terms. One of the relata is the complement of T, the predicate. The other must be the subject of this predicate. It is immaterial whether we project this subject as the specifier of CP or as the specifier of TP in the C/T co-projection complex: either way, PRO winds up in a licensed specifier position of the C/T complex, and is 'blown up'. This derives that in null-operator constructions, the operator generally cannot serve the subject function of a finite clause.[39]

But then how come (83) and (84c), the high-adverbial examples, are grammatical? This time around, the complement of C is a silent-headed RP with the high adverbial in its specifier:

(89) $[_{CP}$ C=*that* $[_{RP}$ *in all likelihood* [RELATOR $[_{TP}$ T $[_{RP}$ <PRO> [RELATOR=$v$/Voice ...]]]]]]

---

[37] In this respect, (76) is very different from the PRO Theorem, which it serves to supplant. It was never true, in the days of the PRO Theorem, that no member of a chain headed by PRO was allowed to be in a governed position: requiring of PRO that all the members of its chain be in ungoverned positions would be transparently impossible. But (71) can and does apply over entire PRO-chains.

[38] Recall from fn. 32 that in the case of highest-subject null-operator relatives, PRO binds no gap, thanks to the fact that it is locally linked both to the predicate of the relative clause and to the 'head' of the relativised noun phrase. The derivations currently under discussion are all long-distance dependencies, so PRO *must* bind a gap.

[39] Null operators can be subjects in short-distance environments — this has been taken care of: recall (73) and the previous footnote. They can also be small-clause subjects and subjects of ECM-infinitives, as Stowell points out. These subject positions are specifier positions. But arguably they are not necessarily *licensed* specifier positions: the subjects of small clauses and ECM-infinitives must, when they are overt noun phrases, raise to a licensed specifier position in the matrix clause (recall the discussion of Lasnik & Saito 1990 etc.). As long as the lower copy of the null operator is not in that position but instead in the specifier position of the small or ECM-clause, (71) is satisfied. This seems plausible (though it is hard to verify). We also have to allow null operators to bind lower copies in object positions. As long as these are either in non-specifier positions (the traditional 'complement-of' position) or in unlicensed specifier positions, there should be no problem here either. (Recall the discussion in fn. 36, above.)

T can now co-project with the silent RELATOR above it, and through this co-projection, finite T can satisfy the requirement that it have a specifier via the adverbial in SpecRP. PRO is then free to bind a copy in a specifier position below T. That low position is an unlicensed specifier: it is an A–specifier that is not in an Agree relation with a case-probe (A–SLiP). So in (89), the copy of PRO in the subordinate clause is legitimate. The derivation converges, as desired. Stated in classic principles-and-parameters terms, what we would say is that the structure in (89), representing the sentences in (83) and (84c), is grammatical because in this structure the EPP is satisfied by the high adverbial, with the subject extracting from a lower position in the tree.

### 5.3.14 Convince-*class clausal complementation*

At the beginning of the previous subsection, I mentioned three cases in which silence of C is not sufficient for the establishment of a long-distance dependency involving the subject of a finite clause. I have now dealt with one of these environments: *tough*-movement and parasitic gap constructions. It is time now to turn next to the second case: finite clauses in the complement of a particular class of matrix verb.

### 5.3.14.1 The empirical lie of the land

Stowell (1981:410, 413) discusses two classes of simplex verbs that can take a nominal dependent plus a finite clause — the *tell* class and the *convince* class.[40] These two verb types behave alike with respect to the distribution of the complementiser *that* in the absence of extraction from the subordinate clause: *that* is optional in both (90a) and (90b).

(90)   a.      they told/showed him (that) she had done it
       b       they convinced/persuaded/reminded him (that) she had done it

The two verb classes also pattern alike with respect to long extraction of the object: this delivers perfectly fine results in both cases.

(91)   a.      what did they tell/show him that she had done *e*?
       b.      what did they convince/persuade/remind him that she had done *e*?

---

40      The *tell* class also contains *ask* and *show*; *ask* is clearly irrelevant with respect to the question of whether long-distance *wh*-dependencies can be established across the boundaries of its clausal complement: it takes a question as its complement, and questions are generally opaque. The *convince* class also features *persuade* and *remind*. Stowell also includes *advise*, but this verb is tricky because, alongside the 'give advice' reading that Stowell is after, it also supports a reading in which it is equivalent to 'notify' (cf. (i)). This muddies the waters: *advise* on its 'notify' reading seems fairly liberal when it comes to long subject extraction, esp. in the passive: cf. (ii). (As we will see in the main text below, the *convince* class proper continues to resist long subject extraction even in the passive.)

(i)             the crew advised the passengers that there would be a 20-minute delay
(ii)    a.      only the payments that the board were advised should be paid were paid
        b.      the figure they were advised should be achieved

But as Stowell points out, the two verb classes behave differently when the subject of the subordinate clause is fronted into the matrix clause: with verbs of the *tell* class, this gives rise to the familiar *that*-trace effect (see (92a)); but with verbs of the *convince* class, long extraction of the subject in his judgement fails regardless of whether *that* is pronounced or left out, as shown in (92b).[41]

(92)    a.    who did they tell/show him (*that) *ec* had done it?
        b.    *who did they convince/persuade/remind him (that) *ec* had done it?

Before proceeding, I should note that long subject extraction of the type represented by (92b) is by no means unattested: as long as the complementiser *that* is omitted, sentences of this type occur in reasonable numbers on the internet. Below are some samples, for the strings "convince/persuade/remind us is". Since I confined my searches to strings in which the pronoun is *us* and the finite verb immediately following the pronoun is *is*, these samples underrepresent what is attested; they serve illustration purposes only.

(93)            "*convince us is*"
        a.    Animal Flesh Is Simply the Most Profitable Substance Businessmen Can Convince Us Is Edible [a headline]
        b.    anything you can convince us is a christmas present
        c.    it gives us a way to distinguish between what is true and what powerful people might wish to convince us is true
        d.    it's what society tries to convince us is desirable and acceptable
        e.    the stuff marketing geniuses are trying to convince us is better than "regular" sugar
        f.    who they try to convince us is a totally randomly selected bar-goer
        g.    maybe it was one of those UFO's The Examiner is trying to convince us is there
        h.    the prevailing view about the history of early modern philosophy, which the author dubs "the early modern tale" and wants to convince us is really a fairy tale, ...
        i.    the Napthine Government has a great big project that it has to convince us is a good idea
        j.    it can mean just about anything that you can convince us is interesting
        k.    athletes inspire admiration, both for what they do and for what they convince us is possible

---

41    Stowell's (1981:410) original examples, reproduced in (i) and (ii), involve a number of confounds that need to be eliminated in order to get a clean picture. One confound, noted by (a reviewer of) Lasnik & Bošković (2003), is that (ii) is prone to garden-path the reader/listener into parsing *his parents* or *Dan* as the subject of the subordinate clause. By replacing the matrix object with an explicitly non-nominative pronoun, as in (92), this problem is averted. Secondly, the source examples in (i) strongly favour readings in which the subject of the subordinate clause is coreferential with an argument of the matrix clause — the matrix object in (ia) and the subject in (ib). Readers/listeners who carry this coindexation over onto the *wh*-fronting examples in (ii) will reject these sentences as Strong Crossover cases, entirely irrelevantly. The examples in the main text make a coreference reading highly unlikely, and thereby eliminate the crossover confound.

(i)    a.    Jim advised his parents they should move to Canada
       b.    Carol convinced Dan she didn't want a cat
(ii)   a.    *who did Jim advise his parents *t* should move to Canada?
       b.    *who did Carol convince Dan *t* didn't want a cat?

(93)        "*persuade us is*"
- a.     the first Thing which they would persuade us is not genuine, is the first Commission for Parker's Consecration
- b.     the very liberty which these men would persuade us is destroyed
- c.     a heavy shining yellow sand, which they would persuade us is gold
- d.     it is still a powerful ideal that great novels like *Les Misérables* can persuade us is true
- e.     in the current economic recession — which the newspapers are desperately trying to persuade us is an economic recovery — the situation is getting worse
- f.     it's not just an interpretive approach they're trying to persuade us is best; it's one that the Supreme Court has explicitly asserted
- g.     even now, it is almost impossible to distinguish what Rumsfeld knows, what he thinks, and what he is trying to persuade us is true
- h.     we are individuals who always, without exception, choose what our combined reasonings persuade us is the choice that we prefer the MOST
- i.     real glycerine, and other synthetic agents that cause the excessive foaming that manufacturers try to persuade us is a good thing via TV advertising
- j.     summer's must-have colour, "blush", which of course the fashion media is trying desperately to persuade us is somehow different from beige
- k.     the true nature of what the establishment is trying to persuade us is a 'crisis'

(94)        "*remind us is*"
- a.     who does the narrator remind us is calling all the shots?
- b.     they encourage us in this belief because they have made it appear so difficult, so complex, so confusing to what they constantly remind us is a "layman's" eye
- c.     Peking (which all the Chinese characters remind us is pronounced "Beijing")
- d.     and then, of course, there's Adam, who the showrunners remind us is supposed to be an "asshole"

But the relative contrast between (92a) and (92b) seems real. Bošković & Lasnik (2003) also endorse the judgements Stowell (1981) reports, both for *tell*-type verbs and for *convince*-type ones. I will proceed on the assumption that the contrast is genuine.

     Neither Stowell nor Bošković & Lasnik ask themselves whether there is a Bresnan/Culicover effect in *show*/*tell*-type and *advise*/*convince*-type cases. It turns out that there is, for both verb classes:

(95)   a.     who did they tell/show him that in all likelihood *ec* had done it?
         b.     ?who did they convince/persuade/remind him that in all likelihood *ec* had done it?

     It is also important to note that passivisation of the *convince*-class verb does not make a difference when it comes to long-distance filler–gap dependencies involving the subject of the subordinate clause: long subject extraction remains impossible unless a high adverbial is present.[42]

---

42     Since *be convinced/persuaded* can also be parsed as a copular sentence with an adjectival predicate (cf. *remain unconvinced/unpersuaded*), this should be carefully controlled for when judging (96) with these verbs. For this reason, I used the progressive.

(96)   a.    *who was he being convinced/persuaded/reminded (that) *ec* had done it?

         b.    ?who was he being convinced/persuaded/reminded that in all likelihood *ec* had done it?

This suggests that the root of the problem with (91) does not lie in some 'defective intervention' effect introduced by the accusative: if 'defective intervention' were at stake, taking the intervener out of the object position might be expected to take away the intervention effect.[43] [44]

      Stowell (1981) does not explicitly talk about non-argument filler–gap dependencies in *convince* constructions. Bošković & Lasnik (2003) claim that adjunct extraction works here. Their examples, reproduced here as (97), are indubitably grammatical. But the trouble with these sentences is that *(a)* unlike in the typical bridge-verb cases, readings in which *how* modifies the *convince/persuade* predicates are quite salient (convincing and persuading can be done in all sorts of ways, and it is often highly relevant to find out exactly how the convincing or persuading was done), and *(b)* the embedded clauses do not in any way require manner modification (*Mary fixed the car* is fine as it is), so nothing forces the *wh*-adverb to be interpreted low.

(97)   a.    how did Carol convince Dan Mary fixed the car?

         b.    how did Kevin persuade Roger he should fix the car?

      To fix the second confound, I have constructed examples whose downstairs clause would be very awkward in the absence of an adverbial modifier. As Gross (1979:864) and, in much greater detail, Grimshaw & Vikner (1993) have pointed out, (98a) is good only with the material in parentheses included in it. Grimshaw & Vikner give the version of (98) lacking any of the parenthesised adverbials an asterisk. I have marked it with a '#' diacritic to indicate that the sentence is syntactically well-formed without the material in the parentheses, but pragmatically infelicitous: it is a truism that the house was built; houses do not emerge or grow by themselves. The same is true, *mutatis mutandis*, for (98b). And we can make this point also for so-called 'lexically selected adverbials': *the letter was worded* in (98c) is uninformative without a manner adverbial being included.

(98)   a.    the house was built #(very well / with great craftmanship / by a genius / in 1575)

         b.    the baby was born #(with a rare disease / on an airplane / on 18 March 1983)

         c.    the letter was worded #(carefully / sloppily / with great care)

---

43     For a pointed critique of 'defective intervention', see Bruening (2015).

44     In this connection, I should also point out that certain adjectival matrix predicates give rise to very much the same subject-extraction pattern as the one Stowell noted for *convince*-class verbs.

(i)    a.    they were aware/certain (that) he had done it

       b.    *who were they aware/certain (that) *t* had done it?

       c.    who were they aware/certain *(that) for all intents and purposes *t* had done it?

As in the case of Stowell's original examples with *convince*-class verbs, attestations of long subject extraction with *aware/certain*-type adjectives can be found on the web: see (i). But such cases are very rare indeed.

(ii)    a.    there's only two companies in this conference who I'm aware did this

       b.    the only study that I'm aware did that, is Alain Crobier's study

       c.    the person who I am certain did it was very quiet all day

When we now consider the sentences in (99) and (100), we see that in the former, with a garden-variety bridge verb upstairs, there is no sense of awkwardness at all but in the latter, featuring a *convince*-type verb in the matrix clause, the listener's/reader's immediate reaction is one similar to that elicited by (98) without the modifiers. This suggests that association of *how* with the lower clause in *convince*-type constructions is in fact obstructed, *contra* Bošković & Lasnik (2003).[45]

(99)  a.      how do you think that the house was built?
      b.      how do you think that the letter was worded?
(100) a.      #how did they convince him that the house was built?
      b.      #how did they convince him that the letter was worded?

One would also like to verify whether non-arguments other than adverbials can establish filler–gap dependencies across the boundaries of the clause embedded under *convince*-class verbs. In (101) I present a couple of examples of predicate dependencies. Though perhaps not quite as difficult as the subject dependency in (92b), these are degraded.

(101) a.      ??how much did they convince him that this costs/weighs?
      b.      ??he wanted a strong leader, which they convinced him that she is

Let me close this descriptive subsection on a comparative note. The extraction facts for *convince*-class verbs reviewed above are not specific to English. This is demonstrated by the following quintuplet from Dutch. Here we see that Dutch *convince*-type verbs behave exactly like English. Systematically, we find that object extraction is successful, that extraction of the subject of the lower clause fails unless a high adverbial of the *in all likelihood* type is present,[46] and that a *wh*-adverbial or predicate cannot be interpreted as a constituent of the downstairs clause.

(102) a.      wat hebben ze je overtuigd dat hij fout gedaan heeft?          (Dutch)
              what have they je convinced that he wrong done has
              'what have they convinced you that he has done wrong?'
      b.      *wie hebben ze je overtuigd dat iets fout gedaan heeft?
              who have they you convinced that something wrong done has
      c.      ?wie hebben ze je overtuigd dat naar alle waarschijnlijkheid iets fout gedaan heeft?
              who have they you convinced that to all likelihood something wrong done has
      d.      #hoe hebben ze je overtuigd dat het huis is gebouwd?
              how have they convinced you that the house is built
      e.      *hoeveel hebben ze je overtuigd dat dit kost/weegt?
              how.much have they you convinced that this costs/weighs

---

45      Christopher Piñón (p.c.) tells me that in his judgement, long adverbial dependencies are also impossible in *show/tell*-class constructions. I have not investigated this further. If the effect is robust, it will be a challenge to derive from the theory the apparent fact that extraction of non-arguments from the clause under *show/tell*-type verbs is impossible whereas all arguments, including subjects, allow filler–gap dependencies across the boundaries of that clause.

46      Because Dutch never allows the complementiser of a declarative complement clause to be omitted, there is only one version of (92b) that can be tested for this language.

There is no analysis available in the literature that manages to account for the entire conglomerate of facts surveyed above. In the next subsection, I will briefly review what I believe is the only attempt in the minimalist programme to come to terms with Stowell's original examples, after which I will proceed to develop my own, more inclusive and (I hope) less stipulative account.

### 5.3.14.2 Previous scholarship

Bošković & Lasnik (2003:section 5) present an account of *convince* constructions that is essentially a 'minimalist' update of Stowell's (1981) tentative approach to these cases. It capitalises on the assumption that the C-head of the CP complement of *convince*-type verbs cannot be an 'affix C' of the kind legitimate under bridge verbs.[47] Since C cannot be an affix, whenever *convince*-type verbs take a CP complement, its head must be a non-affix C. But when *that* is absent, the clausal complement can just be an IP. If the IP option is chosen, subject extraction is blocked on 'the plausible assumption that a *wh*-trace in SpecIP must be licensed by an agreeing C'. If the non-affixal C approach is taken, extraction of the subject of the clause can be excluded 'if we assume that the C heading the embedded clause cannot have the EPP feature': if extraction of the subject must proceed via an intermediate stopover, and if C must be able to possess the EPP property in order for successive-cyclic movement via SpecCP to be legitimate, the hypothesis that the non-affixal C-head of the clausal complement of *convince*-type verbs lacks the EPP property rules out extraction of the subject via SpecCP. Extraction of the subject via intermediate adjunction to IP (instead of substitution for SpecCP) is also ruled out since this would violate 'anti-locality' (Grohmann 2003, Abels 2003). Though they do not mention this themselves, this latter assumption predicts the Bresnan/Culicover effect seen in (95) quite effortlessly. The presence of *in all likelihood* remedies the anti-locality problem, on the assumption that the presence of such a high adverbial makes intermediate adjunction legitimate for the subject.

The Bošković & Lasnik approach to *convince* constructions relies on a couple of stipulative ingredients: the assumption that a *wh*-trace in SpecIP must be licensed by an agreeing C, and the assumption that the C heading the embedded clause of *convince*-type constructions cannot have the EPP property. It also banks heavily on two general assumptions that are very much at the centre of debate in this book: that all *wh*-extraction must proceed successive-cyclically, and that successive-cyclic movement via SpecCP can happen only if C is specified for 'EPP'. Finally, their analysis makes it too easy for adjuncts to extract from *convince* complements: as we have seen, long adjunct extraction is in fact quite hard here. We cannot subscribe to Bošković & Lasnik's proposal, therefore.

### 5.3.14.3 What makes the two verb classes different

To arrive at a better account, it is important first of all to understand what makes *tell*-type verbs and *convince*-type verbs different. Key here is that the former are double-object verbs whereas *convince* and its ilk are not. Dutch gives us a useful clue in this regard. Passives of Dutch *convince*-type verbs with clausal objects force the nominal dependent to be promoted to subject, as in (103).

---

47    If we take 'affix C' to be equivalent to Stowell's incorporating C, the basic assumption underlying both accounts is the same. I will not discuss Stowell's approach here because it is less explicit than Bošković & Lasnik's, and ultimately has the same problems as the latter.

(103)  a.      ze overtuigden mij dat S                                                     (Dutch)
                they convinced me that S
       b.      ik werd overtuigd dat S
                I got convinced that S
       c.      *mij werd overtuigd dat S              /        *er werd mij overtuigd dat S
                me got convinced that S                         there got me convinced that S

Passives of *tell*-type verbs with clausal complements, on the other hand, always allow the nominal object to retain its non-nominative case, with the promotion pattern in (104b) giving rise to speaker variation and prescriptive scorn. The non-standard promotion of the indirect object of a ditransitive with a clausal object need not concern us here: what matters for us in the context of the present discussion is that (104c) is grammatical, in sharp contrast with (103c). In this regard, (104) behaves like double-NP ditransitives such as (105).[48]

(104)  a.      ze vertelden mij dat S                                                       (Dutch)
                they told me that S
       b.      %ik werd verteld dat S
                I was told that S
       c.      mij werd verteld dat S                 /        er werd mij verteld dat S
                me was told that S                              there got me told that S
(105)  a.      ze gaven mij twee boeken                                                     (Dutch)
                they gave me two books
       b.      *ik werd twee boeken gegeven
                I got.SG two books given
       c.      mij werden twee boeken gegeven         /        er werden mij twee boeken gegeven
                me got.PL two books given                       there got.PL me two books given

Passives of the c–type are unique to ditransitives. The fact that *overtuigen* 'convince' in (103) does not allow such a passive and forces promotion of the nominal dependent to subject suggests strongly that it does not project a ditransitive syntax.

        English also presents evidence that while *tell*-type verbs are ditransitive, *convince*-type verbs are not. To see this for English, we have to look beyond examples in which the nominal dependent is accompanied by a clausal object, and look instead at cases in which both dependents are nominal. When verbs of the *tell* class take two nominal dependents, we get a double-object construction, as in (106). But when verbs of the *convince* class take two nominal dependents, one of them must be contained in a PP headed by *of*, as we see in (107).[49]

---

48      Ordinary double-NP ditransitives ban promotion of the indirect object categorically, probably because such promotion would deprive the direct object of its case. A clausal object, esp. an infinitival one (which is formally a PP), may escape the demands of the Case Filter, and thus allow promotion of the indirect object more readily.

49      In this context, we find another indication that *advise*, which Stowell (1981) places in the *convince* class, deserves a separate treatment. For *advise* (in its intended 'give advice' reading), it is possible to participate in what appears to be a true double-NP pattern when the direct object is a bare noun like *treatment*, as in (i):

(i)      the doctor advised him treatment

(106)　　　　they showed/told him the news that Islamic State was preparing an attack
(107)　　　　they advised/convinced/persuaded/reminded him *(of) the danger of Islamic State

　　　　In light of this, the hypothesis that I would like to advance is that with *tell*-type verbs, the nominal dependent is an *indirect object* and the clause is the direct object,[50] whereas in *convince*-type verbs, the nominal dependent is the accusative *subject* of an abstract secondary predicate that takes the clause as its complement. I take the predication relation between the secondary predicate 'PredP' and *him* to be mediated by V (i.e., V is the RELATOR, in the sense of Den Dikken 2006a). The secondary predicate is most likely adjectival in nature, which is why, when the second object is not clausal but nominal, the preposition *of* must be inserted (cf. *afraid *(of) heights*). The hypothesis that 'Pred' is adjectival will also help us make sense of the fact that *that*-omission is grammatical in *convince*-class constructions (recall (90b)): adjectives likewise allow the complementiser of their clausal complement to be omitted (*afraid/certain/convinced (that) it will rain*). I will return further below to the finer details of the structure — in particular, to the question of what to do with *of*.

(108)　　　　$[_{vP} \ v \ [_{VP} \ him \ [_{V'} \ tell/show \ [_{CP} \ that \ S]]]]$
(109)　a.　　$[_{vP} \ v \ [_{VP} \ him \ [_{V'} \ convince/persuade/remind \ [_{PredP} \ Pred_{\varnothing} \ (*of) \ [_{CP} \ that \ S]]]]]$
　　　b.　　$[_{vP} \ v \ [_{VP} \ him \ [_{V'} \ convince/persuade/remind \ [_{PredP} \ Pred_{\varnothing} \ *(of) \ [_{DP} \ the \ danger]]]]]$

　　　　The approach to *convince*-type verbs reflected by (109) can derive the fact, noted by Stowell (1981:409), that the object clause cannot be topicalised. The examples in (110) are Stowell's.

(110)　a.　　*[that his hamburgers were worth buying], Kevin persuaded Roger
　　　b.　　*[that tigers are dangerous], Eric reminded the teacher

The ill-formedness of these examples plausibly has its roots in the licensing of the gap bound by the topicalised clause. We can hold the fact that the closest c-commanding head in (109) is a silent predicate head responsible for precluding the licensing of this empty category.[51]

---

50　For the reader's convenience, I present the structures as 'Larsonian' VP-shells. Since nothing in what follows depends in any way on the way ditransitives and secondary predication constructions are represented structurally, I have chosen a format that will be easily recognisable to most readers. I will systematically abstract away from *v*+V amalgamation, and represent the verbal root in the 'V'-slot.

51　One might reasonably ask why the examples in (110) cannot be derived by topicalisation of the entire zero-headed predicate phrase, so that issues concerning the licensing of the CP-trace would not arise (because in fact CP would not bind a trace on such a derivation). Bošković & Lasnik (2003) point out that the complementiser-less object clause of *convince* constructions can be extraposed, as shown in (i). (They point out in a footnote that there is some speaker variation on how good these sentences are, but they note that are clearly better than a similar attempt at extraposing the complementiser-less clausal complement of a verb like *seem*, as in (ii).) The bracketed constituents in (i) probably are the projections of the null predicate head. In extraposed position, this null predicate head can apparently be licensed: one might imagine that its licensing can be taken care of at PF, and succeeds because the extraposed constituent can be prosodically integrated into what precedes it. For topicalisation, by contrast, PF licensing via prosodic integration is clearly not an option. So preposing the entire null-headed predicate phrase is not an option.

(i)　　a.　　$^{?}$Kevin persuaded Roger yesterday [his hamburgers were worth buying]
　　　b.　　$^{?}$Carol convinced Dan at that time [she didn't want a cat]
(ii)　　　　*it seemed at that time [John had left]

Not only can the silent predicate head itself not license the trace of CP-topicalisation, the clausal complement of *convince*-type verbs also cannot be turned into an Agree-goal for the verb. This is confirmed by the facts of passivisation. Passives have the well-known property of exhibiting promotion to subject of the object with which the active verb entertains an Agree relation. Some prepositional objects can become derived Agree-goals for the verb, in so-called 'pseudo-passives' — which sometimes are grammatical even when a nominal object intervenes between V and P (*his gullibility was taken advantage of*, *these facts have not been paid sufficient attention to*). But the prepositional object of *convince*-type verbs cannot be promoted to subject in a pseudo-passive, regardless of whether it is nominal (and preceded by an overt *of*) or clausal:

(111)        they convinced/persuaded/reminded people of the danger of Islamic State
     a.        people were convinced/persuaded/reminded of the danger of Islamic State
     b.      *the danger of Islamic State was convinced/persuaded/reminded people of

(112)        they convinced/persuaded/reminded people that Islamic State posed a serious threat
     a.        people were convinced/persuaded/reminded that Islamic State posed a serious threat
     b.      *that Islamic State posed a serious threat was convinced/persuaded/reminded people
     c.      *it was convinced/persuaded/reminded people that he should leave

The ungrammaticality of (111b) and (112b) is much sharper than what we see in double-object constructions with *give*-type verbs, where the effect of promoting the direct object to subject in a passive evinces a considerable amount of speaker variation (see Postal 20xx for a clear demonstration that the grammaticality of things like *the book was given him* is not just a matter of dialectal variation). The acute degradation of (112b) also cannot be blamed on the fact that it features a sentential subject: sentential subjects are by no means impossible in the general case (*that he did this was unfortunate* is perfectly fine); moreover, extraposing the sentential subject and filling the structural subject position with proleptic *it* (analogously to *it was unfortunate that he did this*) does not save (112b), as the ungrammaticality of (112c) shows.

The aggregate of (107), (110) and (112) indicates clearly that in (90) the matrix verb does not entertain an Agree relation with the clausal object. CP is the complement of the silent secondary predicate in the structure in (109a) (repeated below); this silent predicate stands in the way of an Agree relation between the verb and the CP.

(109a)        $[_{vP}\ v\ [_{VP}\ him\ [_{V'}\ convince\ [_{PredP}\ \text{Pred}_{\varnothing}\ (*of)\ [_{CP}\ that\ S]]]]]$

The silent predicate head in (109a) does not Agree with the clausal object: the fact that this predicate is probably adjectival in nature (recall the text above (108)) makes it impossible for it to engage in an Agree relation with its complement; adjectival predicates do not show object agreement. Although the matrix *v* is in an Agree relation with the small clause in its complement and probably also with the head of the predicate of the small clause (licensing the latter's silence thereby), the fact that the predicate head does not Agree with its complement renders this CP opaque: it is an absolute island.[52]

---

52        On the other hand, in the ditransitive examples with *tell*-class verbs, whose structure was given in (108), there is an Agree relation between the *v*/V co-projection complex and the object clause. For the CP in the complement of *tell*-type verbs, this means that it is transparent for extraction. Since the filler finds itself in the clause of this triadic verb, and

*5.3.14.4 Long-distance* wh-*dependencies in* convince-*class constructions:* Wh-*control*

The impossibility of establishing subject and non-argument dependencies across the boundaries of the CP below *convince*-class verbs (as shown for English in (92b), (100) and (101), and for Dutch in (102b,d,e)) now falls out directly from the absolute islandhood of this CP. But given that subject extraction is facilitated by the insertion of a high adverbial in the object clause (the Bresnan/Culicover effect, seen in (95b) and (102c)), and that object extraction is entirely unobstructed (see (91b) and (102a)), identifying the finite CP as an absolute island in *convince*-type constructions would seem to be too blunt a tool. How do we allow the grammatical dependencies to be established across the absolute island?

 Absolute islands are absolutely impenetrable for the creation of *direct* filler–gap dependencies across their borders. But if a filler has been downloaded onto the edge of the first predication structure in the matrix clause and a trace can be postulated for it in that predication structure, the filler will have been successfully anchored in the matrix clause — and if it can then find a way to link up to a gap inside the opaque domain without binding a trace there directly, the derivation can still land on its feet. What we are on the look-out for, therefore, is a way to cut up the dependency between the matrix filler and the gap inside the opaque subordinate CP into two separate dependencies, each local, and connected to the other.

 At this juncture, it will be informative to have a look at the way Hungarian deals with long filler–gap dependencies in *convince*-class constructions. To set the stage, let us first examine what *convince*-class constructions are like in Hungarian. We see this in (113a): the verb *meggyőz* 'convince' takes an accusative dependent (the one who is being convinced) and a delative PP (cf. the *of*-phrase of English *convince somebody of something*) containing a proform linked to the subordinate clause. Since Hungarian is a language allowing for object drop, the accusative pronoun *engem* and delative *arról* can both be omitted, as in (113b).

(113)  a.   arról meggyőztek engem, hogy János szereti Marit       (Hungarian)
      it.DEL VM.convinced.3PL.INDEF me.ACC that János.NOM love.3SG.DEF Mari.ACC
   b.   meggyőztek, hogy János szereti Marit
      VM.convinced.3PL.INDEF that János.NOM love.3SG.DEF Mari.ACC
      both: 'they convinced me that János loves Mari'

 When we now try to build a filler–gap dependency for *kit* 'who.ACC' *qua* object of the lower clause, we see that the result is sharply ungrammatical: (114a). This confirms our earlier conclusion that the clause in the complement of *convince*-type verbs is an absolute island. But though (114a) is not the way, Hungarian does have two alternative strategies at its disposal to establish a dependency for the object of *szeret* 'love' across the boundaries of the *hogy*-clause. The first (and preferred) strategy for constructing the Hungarian equivalent of English *who did they convince you that János loves?* is illustrated in (114b):

since the CP containing its gap is an Agree-goal of this verb, a filler–gap dependency can be established across the boundaries of the CP in (108). No constituent type should be prevented categorically from extraction, which is indeed the case: the object-CP in *tell* constructions allows subject, object and adverbial filler–gap dependencies to be created across the its boundaries.

(114)  a.    *kit győztek meg (téged) hogy János szeret?        (Hungarian)
            who.ACC convinced.3PL.INDEF VM you.ACC that János.NOM love.3SG.INDEF

        b.    miről győztek meg (téged) hogy kit szeret János?
            what.DEL convinced.3PL.INDEF VM you.ACC that who.ACC János.NOM love.3SG.INDEF

        c.    kiről győztek meg (téged) hogy János szereti?
            who.DEL convinced.3PL.INDEF VM you.ACC that János.NOM love.3SG.DEF
            'who did they convince you that János loves?'

(114b) is a *wh*-scope marking construction, with *miről* as the [+WH] counterpart to the *arról* seen in (113a), and the 'real' *wh*-phrase located inside the subordinate clause (as *kit*, which controls indefinite agreement downstairs). The alternative exploits what (following Salzmann 2006) I have called 'resumptive prolepsis': delative *kiről* in (114c) binds a silent copy selected by the secondary predicate of the matrix *convince* clause; the embedded clause contains a silent pronoun (*pro*) which controls definite object agreement with the lower verb. The pattern is the same for the subject: once again, a direct long-distance filler–gap dependency (this time involving *ki* 'who.NOM' in the matrix clause) is ungrammatical; but *wh*-scope marking and resumptive prolepsis both deliver grammatical outputs, as shown in (115b,c).[53]

(115)  a.    *ki győztek meg (téged) hogy szereti Marit?        (Hungarian)
            who.NOM convinced.3PL.INDEF VM you.ACC that love.3SG.DEF Mari.ACC

        b.    miről győztek meg (téged) hogy ki szereti Marit?
            what.DEL convinced.3PL.INDEF VM you.ACC that who.NOM love.3SG.DEF Mari.ACC

        c.    kiről győztek meg (téged) hogy szereti Marit?
            who.DEL convinced.3PL.INDEF VM you.ACC that love.3SG.DEF Mari.ACC
           (*)'who did they convince you loves Mari?'

The c–strategy employed by Hungarian, involving prolepsis with a silent resumptive pronoun, is unavailable for English: English is not a pro-drop language, so it cannot license a silent pronoun in the lower clause. *Wh*-scope marking would not seem to come in very handy for English either: after all, adult English does not use this strategy.[54] Yet the strategy of having the *wh*-constituent bind a trace in the matrix clause and linking it to an independently established operator dependency in the subordinate clause does, it turns out, lead us to a solution for the English (and Dutch) facts reviewed above — one that makes precise and accurate predictions about the distribution of long filler–gap dependencies in constructions with *convince*-class verbs.

---

53    Not surprisingly, direct long-distance non-argument filler–gap dependencies are also impossible in *convince*-class constructions in Hungarian: only the *wh*-scope marking construction in (ib) can be used with non-arguments.

(i)    a.    *hány kiló győztek meg (téged) hogy lesz?
           how.many kilo(NOM) convinced.3PL.INDEF PV you.ACC that will.be.3SG
           *'how many kilos did they convince you that (s)he will be?'

        b.    miről győztek meg (téged) hogy hány kiló lesz?
           what.DEL convinced.3PL.INDEF PV you.ACC that how.many kilo(NOM) will.be.3SG

54    Children acquiring English as their first language and also second-language learners of English do produce *wh*-scope marking constructions. See Thornton (1990) and the references given in this context in chapter 4.

What I propose as the derivation for English *convince*-class constructions involving a filler in the matrix clause and a gap associated to it in the subordinate clause is that they have a syntax in which the *wh*-filler in the matrix SpecCP gets associated with the matrix predicate as a dependent of it, alongside the clause — i.e., the *wh*-expression binds a trace in the position that pronominal associates of complement clauses would otherwise occupy. Since the *wh*-expression cannot be fully interpreted in this position, it needs to also bind an empty category inside the subordinate clause. In Hungarian, this can happen via binding of a silent pronoun, as in the b–sentences in (114) and (115). In English, such is impossible, as I pointed out above. But there is a straightforward alternative: a form of CONTROL. A null operator (PRO) is inserted in the specifier position of the subordinate CP, and this null operator is linked to a trace inside the embedded clause, which facilitates full interpretation. We thus have two operator dependencies: one in the matrix clause, headed by an overt *wh*-operator (the filler), and one in the subordinate clause, headed by a null operator (PRO). The schematic structure in (116) depicts this.

(116)     $[_{CP}$ *wh*$_i$ *did they convince me t*$_i$ $[_{CP}$ PRO$_i$ ... *e*$_i$ ...]]

This derivation is in a sense a blend of what we see in the Hungarian resumptive prolepsis and *wh*-scope marking constructions in (114b,c) and (115b,c). With the resumptive prolepsis construction, (116) shares the presence of a pronominal element in the lower clause (though the identity and position of the pronominal elements are different: *pro* in an A-position in Hungarian, PRO in an Ā-position in English); with the *wh*-scope marking construction it has in common the fact that there are two operators, one in each clause (though in English the lower operator is null: PRO).

In section 4.4, I used the term '*wh*-control' for the dependency between the matrix *wh*-filler and the PRO in the subordinate clause in (116). Note that the absolute islandhood of the subordinate clause is in no way an impediment to the establishment of *wh*-control: control relations are well known to be establishable across island boundaries (control into adjunct clauses, in particular).

When we compare the control relation that (116) depicts for a sentence like (117b) with what we find in the ordinary control construction in (117a), where *me* controls a PRO in the subject position of the infinitival complement clause, we see two differences: *(a)* the position occupied by the controller (SpecVP in (117a), SpecPP in (117b)), and *(b)* the position occupied by PRO (in the former an A-position, immediately relatable to its predicate; in the latter an Ā-position, necessarily linked to a trace further downstream). The two control relations can be combined in a single sentence, as in (117c).

(117)  a.     they convinced/persuaded/reminded him to do it
       b.     what did they convince/persuade/remind him that he should do?
       c.     what did they convince/persuade/remind him to do?

Schematic representations of the three examples in (117) are given in (118). Underscoring is used to keep track of the control relations.

(118)  a.     $[_{vP}$ *convince <u>him</u>* $[_{CP}$ C $[_{TP}$ <u>PRO</u> [T ...]]]
       b.     <u>*wh*</u> ... $[_{vP}$ *convince him t*$_{wh}$ $[_{CP}$ <u>PRO</u> [C $[_{TP}$ *he* [T ... <<u>PRO</u>>]]]]]
       c.     <u>*wh*</u> ... $[_{vP}$ *convince <u>him</u> t*$_{wh}$ $[_{CP}$ <u>PRO</u> [C $[_{TP}$ <u>PRO</u> [T ... <<u>PRO</u>>]]]]]

Before returning to the central question that this discussion of *convince*-class constructions is about (viz., the question of why long *subject* dependencies are bad here regardless of the presence or absence of a complementiser), let me dwell a little bit further on the idea that the relation between the *wh*-chain in the matrix clause and the one in the subordinate clause in these constructions is a CONTROL relation. It turns out that this idea delivers a simple account of a gap in the Hungarian paradigm for long-distance dependencies with *convince*-class verbs that is otherwise quite difficult to understand. In (114), repeated below, I presented two ways in which Hungarian can build a dependency between a matrix *wh*-filler and a gap in object position in the clause embedded under the *convince*-class verb: the b– and c–sentences. Interestingly, the third logically possible strategy, given below in (114d) (which now completes the paradigm), is ungrammatical.

(114) a.    *kit győztek meg (téged) hogy János szeret?                                    (Hungarian)
            who.ACC convinced.3PL.INDEF VM you.ACC that János.NOM love.3SG.INDEF
      b.    miről győztek meg (téged) hogy kit szeret János?
            what.DEL convinced.3PL.INDEF VM you.ACC that who.ACC János.NOM love.3SG.INDEF
      c.    kiről győztek meg (téged) hogy János szereti?
            who.DEL convinced.3PL.INDEF VM you.ACC that János.NOM love.3SG.DEF
      d.    *kiről győztek meg (téged) hogy János szeret?
            who.DEL convinced.3PL.INDEF VM you.ACC that János.NOM love.3SG.INDEF

The contrast between (114c) and (114d) is morphophonologically subtle, but very sharp to native speakers. The only surface difference between the two sentences is the presence or absence of the *-i* on the lower verb. This minimal morphophonological distinction signals a robust difference in syntactic derivations: in (114c) the definite inflection on *szereti* indicates that the object of this verb is a silent pronoun (*pro*), as in the 'resumptive prolepsis' derivation; the ungrammaticality of (114d) tells us that it is impossible for the object of *szeret* to be a trace bound by a null operator — null operators serving as objects in Hungarian give rise to indefinite agreement on the verb. Put differently, (114d) shows that there is no grammatical derivation for the Hungarian equivalent of English *who did they convince you that János loves?* in which *wh*-control is involved. Why?

        An answer suggests itself when we consider the contrast between (119a) and (119b):

(119) a.    akarom/elvárom tőled, hogy menj haza                                    (Hungarian)
            want/expect.1SG.DEF you.ABL that go.2SG.SUBJUNC home
            'I want/expect from you that you go home'
      b.    *akarok/elvárok tőled hazamenni
            want/expect.1SG.INDEF you.ABL home.go.INF
            'I want/expect from you to go home'

What the ungrammaticality of (119b) shows is that an oblique dependent of the matrix verb cannot serve as a controller for a PRO in an infinitival complement clause.[55] When we now look back at (114d), and approach it as a case of *wh*-control, we see an immediate connection between its ungrammaticality and that of (119b): *kiről* 'who.DEL' is an oblique matrix constituent; it tries but

---

55      I will come back in section 5.3.14.7 to the question of why control fails with oblique matrix constituents.

fails to control the *wh*-PRO (a.k.a. null operator) in the subordinate clause. The English *wh*-control cases in (117b,c), which are grammatical, differ precisely from Hungarian (114d) in not featuring an oblique matrix *wh*-element; so *wh*-control is successful in the English sentences.

*5.3.14.5 The ban on long subject dependencies in* convince-*class constructions explained*

With the *wh*-control now identified as the active ingredient in the creation of long filler–gap dependencies in English *convince*-type constructions, we can return to the ill-formedness of (92b) and the grammaticality of (95b) (repeated below as (120a,b), resp.).

(120)   a.      *who did they convince/persuade/remind him (that) had done it?
        b.      ?who did they convince/persuade/remind him that in all likelihood had done it?

This contrast can now easily be seen to follow in the same way as that between (121a) and (121b), from the discussion of null operator movement constructions in section 5.3.13.

(121)   a.      *a painting which is easy to think/believe (that) is by Leonardo da Vinci
        b.      ?a painting which is easy to think/believe that in all likelihood is by Leonardo da Vinci

Recall from section 5.3.13 that since lower copies of PRO are not converted to traces, any null-operator dependency will feature multiple tokens of PRO. For each of these, it has to be the case that it is not in a licensed position. For the a–examples in (120) and (121), this is impossible to achieve. When *that* is absent, and C and T co-project, the PRO in SpecCP is in a licensed specifier position. When *that* is present, the PRO in the specifier of non-co-projecting T is in a licensed specifier position. Either way, there is a PRO in the structure that fails to be in conformity with (76), which causes this PRO and the structure that it is in to be 'blown up'. The fact that the b–sentences in (120) and (121) are grammatical follows from the fact that the high adverbial is introduced as the specifier of a silent-headed RP in the complement of C: with T now co-projecting with the silent RELATOR above it, it satisfies the requirement that it have a specifier via the adverbial in SpecRP, setting PRO free to bind a copy in an unlicensed specifier position below T (recall (89), repeated below).

(89)                    $[_{CP}$ C=*that* $[_{RP}$ *in all likelihood* [RELATOR $[_{TP}$ T $[_{RP}$ <PRO> [RELATOR=*v*/Voice ...]]]]]]

        We have now achieved a uniform account of the two sets of examples discovered by Stowell in which an Ā-dependency involving the subject of a finite clause is prohibited regardless of whether the complementiser of the clause is overt or silent. In both cases, we are dealing with a null-operator dependency; null-operator dependencies anchored in the structural subject position of a finite clause are ungrammatical, in the absence of a high adverbial, because they are at variance with (76), the definition of PRO.

        It is now easy to see why Hungarian and English are different with regard to long subject dependencies in their *convince*-class constructions. The grammatical Hungarian example in (115c), above, involves what I have called 'resumptive prolepsis': there is a *pro*, not a PRO, in the subject position of the embedded clause, linked to the matrix *wh*-element in a binding dependency; (76) is not active. By contrast, English (120a) is a *wh*-control construction, in which the gap in the subordinate clause is a PRO, banned by (76) from occurring in a licensed specifier position.

*5.3.14.6 Non-argument dependencies in* convince-*class constructions revisited*

The *wh*-control approach to the *convince*-class data also explains the fact that for non-argumental long-distance filler–gap dependencies in these constructions, there is no way out. We had already found out why a direct filler–gap dependency between an adverbial *wh*-element in the matrix clause and a trace in the subordinate clause is impossible to establish: the absolute islandhood of the subclause precludes such a direct dependency. Arguments can in principle get around the problem posed by absolute islands via the *wh*-control strategy discussed above. But control is the privilege of arguments: by definition (recall (76)), PRO is an unlicensed *argument*, hence no non-argument dependencies can ever be established with the aid of PRO.

This directly accounts for the fact that in absolute island configurations, non-arguments can never get a second chance via *wh*-control: whenever an *ex situ* non-argument cannot establish a grammatical direct filler–gap dependency with a trace, it perishes. We thus explain the fact that in (100), repeated below, it is impossible to interpretively link the adverbial *wh*-element to the embedded clause. Since the embedded clauses represent uninformative truisms without a modifier, the sentences fail to converge. The same is true for non-argument dependencies across other absolute island boundaries — see (122).

(100)   a.      #how did they convince him that the house was built?
        b.      #how did they convince him that the letter was worded?
(122)   a.      #how were they protestesting because the house was built?
        b.      #how were they protesting because the letter was worded?

(A similar account applies to the ungrammaticality of (101), involving predicates. Again, there is a link with dependencies across adjunct islands: \**how much are they protesting because this costs?*.)

So effectively, the only elements that are successful at building long-distance filler–gap dependencies with *convince*-class verbs are objects. In this regard, *convince*-type constructions are remarkably similar to established cases of null-operator dependencies. The analysis of long Ā-dependencies across absolute islands in terms of what I have called *wh*-control delivers this result.

*5.3.14.7 The location of the gap bound by the matrix* wh-*filler*

A few questions remain to be addressed at this juncture. The first is why *wh*-control, and control more generally, should fail with oblique matrix constituents (recall (114d) and (119b)). At its core, the answer to this question is simple. We have known since Williams (1980) that objects of P cannot serve as subjects for depictive secondary predicates (see (123)), which, on an analysis of depictive secondary predication involving control (see Chomsky 1981:110–11, Stowell 1983, Hoekstra 1988), translates into the generalisation that objects of P cannot serve as controllers for PRO. This generalisation, in turn, is straightforwardly translated structurally, in terms of c-command: the object of P does not c-command PRO; with control requiring a c-command relation between the controller and PRO, the ill-formedness of (123b) follows.

(123)   a.       John ate the meat [PRO raw]
        b.      *John ate at/of the meat [PRO raw]

What I just said carries over directly to (119b) if we treat *tőled* 'you.ABL' as a branching constituent, composed of *-től* (the 'case particle'), inflected for second person singular, and a *pro* in complement position. The same lack of c-command can then also be held responsible for the ill-formedness of (114d), on the *wh*-control analysis proposed here: *kiről* is an internally complex constituent, with *ki* encapsulated within it and thereby being prevented from c-commanding out and controlling the PRO operator in the embedded clause.[56]

The inability on the part of P-objects to serve as controllers of PRO seems at first to be flatly contradicted by the grammaticality of examples such as those in (124) (which may sound archaic or formal):

(124)   a.        I ask/beg of you [PRO to tell me the truth]
         b.        I expect from you [PRO to tell me the truth]

But as I will show, an analysis of the way in which verbs like *beg* and *expect* can accommodate both a PP and a clause in their complement not only deals with the apparent control problem in (124) but also paves the way for a structural analysis of *convince*-class constructions that pinpoints the location of the gap bound by the matrix *wh*-filler.

Putting into a single structure all three ingredients of what follows the verb in (124) — a preposition, a noun phrase, and a control clause whose subject is controlled by the noun phrase following the preposition — poses a challenge for a theory that allows nodes to have at most two daughters (binary branching). We cannot escape the challenge by tossing binarity out the window: a flat structure with V, a PP, and a CP would not make a control relation possible between *you* and the PRO-subject of the infinitival clause, because *you* would still be encapsulated in a PP and be unable to c-command out. For similar c-command problems, Pesetsky (1995) suggested a solution in terms of 'cascade structures', as we saw in chapter 2. While cascades are not very helpful as a *general* solution for recalcitrant c-command challenges, it turns out that in the case at hand, they do actually come in handy and do not need to be paired with a 'layered' structure to get constituency facts right: in this particular context, constituency actually tracks c-command.

The structure that I would like to propose for (124) is one in which the complement position of the verb is occupied by a *p*P, in turn taking a PP as its complement within which the two remaining terms of the structure are accommodated, with *you* as the specifier of PP and the control infinitive (a CP) as P's complement:

(125)   a.        [$_{VP}$ *ask/beg* [$_{pP}$ *p* [$_{PP}$ *you* [P [$_{CP}$ C [$_{TP}$ PRO *to tell me the truth*]]]]]]]
         b.        [$_{VP}$ *expect* [$_{pP}$ *p* [$_{PP}$ *you* [P [$_{CP}$ C [$_{TP}$ PRO *to tell me the truth*]]]]]]]

In (125), P can be thought of as a RELATOR of a possessive-like predication relation between the *to*-infinitival CP and the noun phrase in SpecPP, with 'P$_{HAVE}$', in the sense of Harley (20xx), In (126a), this PP structure is embedded below *v* or the copula, as in (125b), and P$_{HAVE}$ amalgamates with that to form the verbal form *have*.

---

56        The analysis of the Hungarian facts can be directly assimilated to that of English (123) if the 'case particles' are analysed as Ps. The same result falls out if they are K-heads and if K obstructs c-command. For our purposes here, it is immaterial how best to analyse the Hungarian 'case particles'.

(126)  a.      you have to tell me the truth
       b.      [$_v$/BE [$_{PP}$ *you* [P$_{HAVE}$ [$_{CP}$ C [$_{TP}$ PRO *to tell me the truth*]]]]]]

The desired word order for (124) comes about with *of* or *from* spelled out in *p*. With *you* occupying the specifier of PP and the control clause sitting in P's complement, c-command between *you* and PRO is straightforwardly guaranteed. And because *of/from you* in the structures in (125) is not a constituent, it is predicted that fronting of this string should fail, which is indeed the case:

(127)  a.      *of you I ask/beg to tell me the truth
       b.      *from you I expect to tell me the truth
       b′.      from you I expect to hear the truth

It is interesting to note that although *from you, I expect to VP* is not impossible *per se* (as in (127b′)), where *from you* binds a gap in the infinitival clause, what IS impossible in such cases is for *you* to serve as controller of the PRO-subject of the infinitive: in (126b) the presence of *me* precludes (for binding-theoretic reasons) a control relation between PRO and *I*, and when *you* is fronted along with *from*, lack of c-command prevents a control relation between *you* and PRO. The 'cascade'-like structures in (125) thus serve the syntax well not just for the purposes of c-command but also for constituency. Cascades of this sort are entirely unobjectionable. With P treated as a RELATOR of the infinitival clause and the noun phrase in its specifier position, we also get the obligatory object control relation found in the examples in (124) to fall out from predication.

Assuming that this is sufficient recommendation for the structures in (125), we can now return to *convince*-class constructions and proceed to building the structures for them, exploiting the kind of 'cascading' PP structure developed for (124). Consider the paradigm in (128), and the partial structures for these sentences given in (129) (based on the schematic structures given in (116) and (118), above). The structures in (129c,d) are directly parallel to those of Hungarian bridge-verb constructions with a pronominal 'herald' and with *wh*-prolepsis, in chapters 2 and 4, for which the proform *azt* 'it.ACC' and the trace of the accusative proleptic *wh* were argued to occupy SpecVP, with the CP sitting in the complement of V.

(128)  a.      they convinced/persuaded/reminded him that she loved him
       b.      *they convinced/persuaded/reminded him of that she loved him
       c.      they convinced/persuaded/reminded him of it that she loved him
       d.      who did they convince/persuade/remind him that she loved *ec*?
       d′.      *who did they convince/persuade/remind him of it that she loved *ec*?

(129)  a.      [$_{vP}$ *v* [$_{VP}$ *him* [$_{V'}$ *convince* [$_{PredP}$ Pred$_{\varnothing}$ [$_{CP}$ *that* S]]]]]]
       b.      *[$_{vP}$ *v* [$_{VP}$ *him* [$_{V'}$ *convince* [$_{PredP}$ Pred$_{\varnothing}$ [$_{pP}$ *p* [$_{PP}$ P [$_{CP}$ *that* S]]]
       c.      [$_{vP}$ *v* [$_{VP}$ *him* [$_{V'}$ *convince* [$_{PredP}$ Pred$_{\varnothing}$ [$_{pP}$ *p* [$_{PP}$ *it* [P [$_{CP}$ *that* S]]]]
       d.      *who* ... [$_{vP}$ *v* [$_{VP}$ *him* [$_{V'}$ *convince* [$_{PredP}$ Pred$_{\varnothing}$ [$_{pP}$ *p* [$_{PP}$ $t_{wh}$ [P [$_{CP}$ *that* S]]]]

For (128a), we can readily content ourselves with the structure in (129a): merging a *p*+PP in the complement of the silent predicate would be pointless; (128b) is ungrammatical because (128a) is well-formed. But when a proform associated to the subordinate clause is present, as in (128c), a

*p*+PP must be projected in the complement of the silent predicate head, as shown in (129c). Here, the proform is accommodated as the specifier of the PP whose head takes the CP as its complement, and *it* is predicated of the *that*-clause (recall chapter 2, and Moro's 1997 *it's that he talks too much*).

With the aid of the *p*+PP structure, we can now also accommodate the trace of the *wh*-operator in the matrix clause in the position occupied by the proform *it* in (128c) — i.e., in the specifier position of the PP in (129d). Doing so immediately accounts for the fact that the *wh*-operator cannot co-occur with an overt proform for the complement clause: (128d′) is ungrammatical.

In the structure in (129d), for (128d), the *wh*-filler in the matrix clause binds a gap that c-commands the subordinate clause. As a result, *wh*-control is grammatical in English *convince*-class constructions. For Hungarian *convince*-type verbs we had found, however, that *wh*-control (manifested by (114d), above) is ungrammatical. I had blamed this precisely on lack of c-command. It now looks like we have a contradiction on our hands. But that is merely apparent. The contradiction dissolves when we take a closer look at (128c) and Hungarian (130).

(130)       arról meggyőztek engem, hogy János szereti Marit       (Hungarian)
            it.DEL VM.convinced.3PL.INDEF me.ACC that János.NOM love.3SG.DEF Mari.ACC
            'they convinced me that János loves Mari'

In (130) we see that Hungarian treats the proform *az* and the postposition as a constituent. What this means is that all of *arról* is mapped into the specifier position of the *p*+PP structure in (129c). We had already found the same to be true, *mutatis mutandis*, for the Hungarian counterparts to (125): all of *tőled* in (119) occupies the specifier position of the *p*+PP structure in (125). The P-heads of the *p*+PP structures in question are themselves silent in Hungarian. The complex internal structure of the occupants of the PP-specifier positions prevents the establishment of the c-command relationship needed for control. So just as (119b) is ungrammatical because lack of c-command causes control to fail, so (114d) crashes as well. And the English examples in (128d) and (124) are grammatical because in the structures for these sentences, the controller sits by itself in the specifier position of PP in the *p*+PP structure, successfully c-commanding the PRO in the embedded clause in P's complement that it wants to control.

### 5.3.14.7 An *of*-trace effect

Now that we have found a place for the trace of the *wh*-operator in *wh*-control constructions with *convince*-class verbs, one last question remains. In (131a), we see the preposition *of* occurring obligatorily. In the *wh*-control construction in (131b), by contrast, *of* cannot occur.[57] The obligatoriness of *of* in the a–sentence would traditionally be attributed to the need to case-license the proform *it*. Now what about (131b), analysed as in (128d)?

(131)   a.      they convinced/persuaded/reminded him *(of) it that she loved him
        b.      who did they convince/persuade/remind him (*of) that she loved *ec*?

---

57      This is of course in stark contrast to Hungarian (114b,c) and (115b,c), where the *wh*-element in the matrix clause does get case-marked in the same way as the proform associated to the CP in non-*wh* contexts. No case-related questions arise for Hungarian: in both the *wh*-scope marking construction and the resumptive prolepsis construction, the *wh*-element in the matrix clause is fully case-licensed (with delative case).

We might want to say that the *wh*-trace in the matrix clause does not need to be assigned case because the entire matrix *wh*-dependency is associated, via the control relation, with an operator dependency in the subordinate clause — which has a case feature available (the one brought in by the *v*+V co-projection complex of *love*) that does not get 'used up' there (PRO does not check case). In the pursuit of such an intuition, we could perhaps draw inspiration from the thematic realm, via the 'copy raising' construction, exemplified by such English sentences as those in (132).

(132)   a.      he seems like he has missed his flight
        b.      he seems like his flight got cancelled

Here it seems as if the subject of the matrix clause and the subject of the embedded clause (or something inside it, as in (132b)) share a θ-role, perhaps like *who* and the null-operator dependency in the embedded clause in (131b) might share case.

Copy raising constructions have been approached as an instance of predication (see Asudeh & Toivonen 2012): the lower finite clause, necessarily containing a pronoun coindexed with the matrix subject, serves as a predicate for the matrix subject, allowing that subject to be interpreted as an argument. Null-operator dependencies frequently serve as predicates as well: the *tough*-movement construction is a well-known case in point. And I had already pointed out in the text below (126) that in the structures in (125) the obligatory control relation can be understood as the reflex of a predication relation between the CP in the complement of P and the constituent in SpecPP. The same carries over to (128), the structure assigned to *wh*-control in *convince*-class constructions. So if the null-operator dependency in the subordinate clause of *wh*-control constructions allows the lower clause to serve as a predicate of the *wh*-dependency in the matrix clause, *and* if this relationship could somehow get the upstairs *wh* case-licensed, we could kill two birds with one stone: we would directly link the two *wh*-dependencies (via predication), and we would account for the obligatory absence of *of* in the matrix clause.

This line of analysis may well be worth pursuing in more detail. But the 'case sharing' approach faces at least one major problem. One would not want to give the subject of *tough*-movement constructions a blanket licence to occur in caseless environments with an appeal to the fact that the null-operator dependency in the *to*-infinitival clause that this subject is linked to gets case in the lower clause. After all, (133) is sharply ungrammatical.

(133)           *[John to be easy to please] would be a surprise

In our pursuit of a successful account of the obligatory non-appearance of *of* in (128d), it is important to return to our earlier examples in (124), of which the a–examples (featuring *of*) are repeated as (134a). For these cases, I have developed a structural analysis which runs along lines entirely parallel to (128d): there is a PRO in a CP that serves as the complement of P, and its controller occupies the specifier position of PP. When we now look at what happens when the controller of PRO is a *wh*-element, we see in (134b) that the *of* that precedes the controller in the a–sentences must remain absent: (134b) is ungrammatical with *of* pronounced.[58]

---

58      The effect is stronger with *beg* than with *ask*, but it is present in both cases. Note that in the a–sentences, *of* is freely omissible. The point that is relevant is that in the b–examples, *of* cannot occur.

(134)   a.        I ask/beg (of) you [PRO to tell me the truth]
        b.        you are the person who I ask/beg (*of) to tell me the truth

We see here a striking parallel between what I have analysed as a *wh*-control construction (viz., (127d)) and what is traditionally treated as a control construction (viz., (134b)). This may bolster the control approach to the former. And it also suggests that the 'case sharing' approach explored above is unlikely to cover the entire spectrum of cases: after all, there is nothing in the subordinate clause with which the chain of *who* in (134b) might share case; the PRO subject of the infinitival clause is not in a case-licensing environment. So 'case sharing' would not just overgenerate in the case of (133), it would also undergenerate for (134). Let us then pursue a different approach.

        What I would like to propose is that the obligatory non-appearance of *of* in (128d) and (134b) is directly related to the obligatory absence of the infinitival complementiser *for* in (135b) — the well-known '*for*-trace effect', discussed earlier in this chapter.

(135)   a.        I would prefer (for) you to tell me the truth
        b.        you are the person who I would prefer (*for) to tell me the truth

The strings in (134) and (135) are very similar. Their analyses are by no means identical: while the *for* in (135) belongs to the embedded clause, the *of* in (134) in all likelihood does not; I have analysed *of* as the spell-out of a *p* in the structure of the matrix clause. But this difference notwithstanding, in relevant respects we find very much the same constellation in both cases:

(136)   a.        $[_{pP}\ p\ [_{PP}\ t_{wh}\ [\text{P} ...]]]$
        b.        $[_{CP}\ \text{C}\ [_{TP}\ t_{wh}\ [\text{T} ...]]]$

In both configurations, spelling out the higher head (as *of* in the case of *p*, and as *for* in the case of C) is disallowed in the presence of a *wh*-trace in the specifier below it.

        For (136b), I have presented an analysis of the obligatory silence of the C-head couched in T-SLiP, the licensing condition for traces in specifier positions: $t_{wh}$ must be licensed as a trace, and in order for this licensing to come about, the trace must be a Downward Agree-goal; this forces C and T to co-project, and co-projection of C and T is only possible when C is silent. The '*for*-trace effect' is thereby derived.

        For (136a), an entirely parallel account now presents itself. In fact, nothing needs to be changed other than the references to specific labels: if for 'C' in the previous paragraph we substitute '*p*' and for 'T' we read 'P', the silence of *p* (otherwise spelled out as *of*) is guaranteed. What we have here, then, is an '*of*-trace effect' that falls out, from T-SLiP, in exactly the same way as the '*for*-trace effect' familiar from infinitival clauses — or, for that matter, the '*that*-trace effect' in (137b). In all these cases, the syntactic configuration features a *wh*-trace in a specifier position immediately below a functional head that needs to be given the opportunity to license the trace in accordance with the demands of T-SLiP. For this to transpire, the functional head must co-project with the head below it. Such co-projection goes hand in hand with the silence of the functional head, as discussed earlier.

(137)   a.        I think (that) you should tell me the truth
        b.        you are the person who I think (*that) should tell me the truth

The obligatory silence of *p* in (136a) has no consequences for the licensing of case for the matrix *wh* in *wh*-control constructions of the type in (117d): *p* is structurally present; it just does not get a phonological matrix when the specifier in its complement is a trace. This is not different in any way from the situation in (136b). Case licensing is a syntactic, structural matter; the question of whether a head is given a phonological matrix or not is a matter of PF, divorced from the question of case.[59]

Now that we have made our way back to the 'complementiser-trace effect', and identified a further case of the same structural type but not involving a C-element, let us turn to the third environment in which the presence or absence of a complementiser does not seem to make a difference — the third environment, put differently, in which extraction of the subject of a finite clause fails even in the absence of *that*.

*5.3.16 Other long subject dependencies across an absolute island*

Cinque (1990) draws attention to the striking parallelism between null-operator dependencies such as those found in parasitic gap constructions, illustrated in (138), and long-distance dependencies between an overt operator and a gap inside an island, such as the CNPC cases in (139).

(138)  a.    who did you turn to *e* for help [after *Op*$_\varnothing$ talking to [$_{NP}$ *ec*]]?
       b.    *to whom did you turn *e* for help [after *Op*$_\varnothing$ talking [$_{PP}$ *ec*]]?
       c.    *how many pounds does he weigh *e* [without *Op*$_\varnothing$ believing he weighs *ec*]?
       d.    *what did John file *e* [instead of *Op*$_\varnothing$ giving that man *ec*]?
       e.    *who did John expect *e* would be successful [though *Op*$_\varnothing$ believing (that) *ec* is incompetent]?

(139)  a.    $^?$who is John looking for someone [who could talk to *ec*]?
       b.    *to whom is John looking for someone [who could talk *ec*]?
       c.    *how many pounds is John looking for someone [who weighs *ec*]?
       d.    *what is John looking for someone [who could give his friends *ec*]?
       e.    *who is John looking for someone [who says (that) *ec* will help him]?

---

59      Regarding case in *convince*-class constructions, it may be relevant to note that in pseudocleft constructions featuring such verbs, it is not uncommon, in informal English, to see the preposition *of* being 'dropped' in the *wh*-clause. Some attested examples from the internet are provided in (i), for *convince*, and (ii), for *remind*. But dropping *of* is generally poor, and certainly non-standard. So the fact that (131b) requires *of* to remain absent cannot be accounted for with an appeal to colloquial cases such as the ones in (i) and (ii).

(i)     a.      what that convinced me was that they do not know
        b.      what this convinced me was that I wouldn't want to book a trip during the busiest times of the year
        c.      what it convinced me was that I either needed to move up or move out
(ii)    a.      what they reminded me was that the concern with 'peace' of many of our fore-runners emerged from, or combined with, engagement in other social movements
        b.      what they reminded me was is that sometimes it's about the opportunities that arise during placement
        c.      what they reminded me was that the Commission does not use in-house expertise only

The a–examples show that a gap in the complement position of *to* can be bound by a null operator in a parasitic gap construction or by an overt operator (matrix *who*) separated from the gap by an island (a complex NP). The b–sentences show that a PP-gap is impossible in both cases. In the c–sentences we see that non-argument dependencies fail in both contexts. The d–examples illustrate the fact that argument dependencies involving a gap in the direct object position of a double-object construction, while grammatical over short distances with an overt operator (*what did you give that man?*), fail both in null-operator constructions (recall (88)) and in long-distance filler–gap dependencies across an island. And, most relevantly for the discussion of subjects in this chapter, the e–examples indicate that both null-operator dependencies and island-crossing overt filler–gap dependencies resist having the gap in the structural subject position of a finite clause, entirely regardless of whether there is a complementiser present or not.

Cinque (1990) seeks to unify the two sets of examples — and I would like to do the same. We have already seen (recall the discussion below item (84) in section 5.3.13, above) that Cinque's *pro*-based attempt is not the optimal vehicle. The PRO-based account pursued here delivers the desired unified result quite simply, when we bear in mind what has already been established in the foregoing. Let me begin by going through (138).

Recall that the key to the examples in (138) is the hypothesis that the null operator $Op_\emptyset$ in these examples is PRO, and that the gap that PRO binds is and always remains an exact copy that of the head of the chain, never reduced to a trace. So both the head and the foot of the chain feature an element, PRO, which is pronominal, argumental, and by definition resists licensing. With this in mind, the ungrammaticality of (138b) falls out from the fact that PRO is nominal; that of (138c) is a consequence of the fact that PRO is argumental; and that of (138d) is caused by the fact that pronouns generally resist being direct objects in English double-object constructions. Finally, the ill-formedness of (138e) is an illustration of the fact that PRO resists licensing: the lower copy of PRO (the one in the structural subject position of the embedded finite clause) is in a position licensed under the SLiP, which causes the structure to fail.

All of this reasoning can be grafted directly onto the examples in (139) if it can be argued that the gap '*ec*' inside the island is a gap bound by a null operator — i.e., a PRO. Put differently, the facts in (139) fall out if (139a) involves what I have called *wh*-control: the *wh*-constituent in the matrix clause in (139a) binds a gap within that clause, *à la* (129d), and controls a second chain in the embedded clause, headed by the null operator (PRO):[60]

(140)        $who_i$ ... [$_{vP}$ $v$ [$_{VP}$ *looking* [$_{pP}$ $p$ [$_{PP}$ $t_i$ [P [$_{DP}$ *someone* [$_{CP}$ $Op$=PRO$_i$ [*who* ... $ec_i$]]]]]]]]]]

This is, in fact, the only option available to (139a) if it is to survive: after all, a direct filler–gap dependency between *who* and *ec* cannot be established due to the presence of an absolute island boundary in between the two elements. The relative clause is not in a Downward Agree relation with a probe, and hence an absolute island (recall chapter 1). In the particular case in (139a), the *wh*-operator in the matrix clause gets a second chance, as it were, via *wh*-control. But in the other examples in (139), *wh*-control fails because the nature and distribution of PRO are not being respected: the gap is not nominal, or not argumental, or licensed, or in a position in which pronouns cannot sit.

---

60        I will turn in section 5.3.17 to the question of how the relative clause associated to someone can harbour both a null operator and *who*. See the structure in (153) for the answer.

*5.3.17 On the selective transparency of highest-subject relative clauses*

It will be good to linger a bit longer on the grammatical example in (139a). This example is representative of a range of sentences, from English as well as other languages, that have attracted a good deal of attention in the generative literature, dating back at least as far as Erteschik-Shir (1973), who talked about violations of the Complex NP Constraint in Danish. For English, the transparency of certain relativised noun phrases was placed prominently on the research agenda by Chung & McCloskey (1983), whose short paper was cited in Chomsky (1986) in the context of a brief discussion of the Vacuous Movement Hypothesis.

　　　Chung & McCloskey (1983) point out the contrast between (141a) and (141b), and thereby emphasise the fact that extraction from a relativised noun phrase is possible when the relative clause is a subject relative but not when it is an object relative. Cinque (1990), who picks up the syntactic thread, adds an important piece to the puzzle: the fact that it is not the case that highest-subject relatives are transparent to any and all filler–gap dependencies. As the ill-formedness of (141c) shows, a PP-filler outside a relativised noun phrase cannot bind a gap inside it. We had seen this in (139b) already. And from (139c–d) we had learnt that dependencies involving a non-argumental filler (even if a *nominal* one) or a direct-object gap in a double-object construction fail across relativised noun phrases, even if the operator in relative clause is the subject of the relative clause itself.

(141)　a.　　　this is a paper which we need to find someone [who understands *e*]
　　　　b.　　*this is a paper which we need to find someone [who we can intimidate with *e*]
　　　　c.　　*this is a paper about which we need to find someone [who can talk *e*]

Such contrasts are not unique to English (though it is probably fair to say that they are not universal either). Thus, Cinque (1990:143–44) points out that similar effects manifest themselves in Italian:

(142)　a.　　　$^?$l'unica persona [che non troveremo [nessuno [che sia disposto ad ospitare *e* questa notte]]]　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　(Italian)
　　　　　　　'the only person that we won't be able to find anyone who is willing to put up for the night'
　　　　b.　　*l'unica persona [che non è facile trovare [le persone [alle quali poter presentare *e*]]]
　　　　　　　'the only person that it is not easy to find people to whom to introduce'
　　　　c.　　*l'unica persona [su cui abbiamo trovato [qualcuno [che sia disposto a contare *e*]]]
　　　　　　　'the only person on whom we found someone who is willing to count'

　　　From the literature on Scandinavian filler–gap dependencies across a highest-subject relative, we are familiar, in addition, with the fact that the verbal material (mostly in the matrix clause but to some extent also in the relative clause) matters a great deal when it comes to the success or failure of such dependencies. Thus, Allwood (1982:24) presents the following cline of examples from Swedish, with acceptability declining gradually, from perfectly fine to 'much more difficult':[61]

---

61　　　Allwood himself does not give acceptability diacritics, but from his prose the diacritics given in (143) can be distilled.

(143)  a.      de blommorna ser jag en man som vattnar                    (Swedish)
               those flowers see I a man that is.watering
       b.      de blommorna känner jag en man som säljer
               those flowers know I a man that sells
       c.      ?de blommorna tänker jag på en man som sköter
               those flowers think I of a man that tends
       d.      ??de blommorna tänker jag på en man som vattnar
               those flowers think I of a man that is.watering
       e.      *de blommorna talar jag med en man som säljer
               those flowers talk I with a man who sells

For Allwood, one important factor constraining extraction from Swedish relative clauses is the question of whether the matrix verb 'can be directed toward whole states of affairs' (p. 25). Put differently, the result is best when the matrix clause can select a proposition. This suggests a link with the so-called pseudorelative of the Romance languages — see e.g. French (144) (Kayne 1975):

(144)  a.      j'ai vu ton frère qui volait son voisin
               I have seen your brother who robbed his neighbour
               'I saw your brother rob(bing) his neighbour'
       b.      ton frère a été vu qui volait son voisin
               your brother has been seen who robbed his neighbour
               'your brother has been seen robbing his neighbour'
       b′.     (*)ton frère a été vu qui volait son voisin
               your brother has been seen who robbed his neighbour
               [* on a reading corresponding to (143b)]

For (144a) an ordinary restrictive relative reading is available if *qui volait son voisin* is supposed to narrow down the reference of *ton frère*, if the addressee has multiple brothers and only one of them robbed his neighbour. On such a reading, the speaker need not have seen the addressee's brother rob anyone. But (144a) can also be used in a context in which the addressee has just a single brother, making a construal of *qui volait son voisin* as a restrictive relative clause infelicitous. In that case, the speaker must have witnessed the event of the addressee's brother robbing his neighbour. This reading is best rendered in English as a small-clause complementation construction: the string *qui volait son voisin* is the predicate of a small clause in the complement of the perception verb; we are not dealing with a complex noun phrase structure at all. This is further confirmed by the fact that on the second reading for (144a), fronting just *ton frère* in the passive is grammatical, as in (144b), while fronting the entire string *ton frère qui volait son voisin* (as in (144b′), is not. Allwood's (1982) observation that the 'head' of the relative clause can be *wh*-fronted by itself in precisely the kinds of *som*-constructions in Swedish that also support apparent subextraction from the relative clause (see (145)) goes along well with the idea that we may not be dealing with a complex NP in the grammatical cases in (143).

(145)          vem ser jag en man som vattnar de blommorna?
               who see I a man who is.watering those flowers

Kush *et al*. (2009) explicitly present a pseudorelative approach to the Scandinavian examples apparently involving subextraction from a highest-subject relative clause. The fact that the relative must always have a highest-subject gap follows straightforwardly from such an analysis, as does the fact that even non-argumental PP-extraction from these relatives is grammatical in Scandinavian: the examples in (146) are from Allwood (1982:18);[62] established cases of pseudorelatives likewise allow non-argumental PP-dependencies across their borders, as do English small-clause complementation constructions such as those in (147), which allow downstairs construal for the locative adjunct-PP.[63]

(146)  a.      på Röda Torget känner jag till en flicka som gav en pojke en kyss
                on Red Square know I of a girl who gave a boy a kiss
        b.      var känner jag till en flicka som gav en pojke en kyss?
                where know I of a girl who gave a boy a kiss
(147)  a.      on Red Square, I saw a girl giving a boy a kiss
        b.      where did you see a girl kissing a boy?

The pseudorelative approach also captures some of the restrictions on verbs noted by Allwood (1982) — in particular, the fact that (143e) is 'much more difficult' than (143a) falls out straightforwardly. But (143b,c) are not typical pseudorelative environments. More microscopically, the contrast that Allwood says some speakers observe between (143c) and (143d) (as Allwood 1982:24 puts it, '[s]ome speakers who cannot accept [(143d)] can accept [(143c)]') cannot be made sense of based on a pseudorelative analysis.

Kush *et al.* (2009) present the results of several experiments run on English native speakers. They report a contrast between sentences like (148a,b) and (148c). Averaged over the 24 participants in the experiment, the acceptability rating for (148a,b), for which a small-clause construal is possible, is slightly above 3, on a 7-point Likert scale; that for (148c), which supports no small-clause syntax (*meet* does not take propositional complements) is at 2.5. Statistically, this contrast is significant.

(148)  a.      this is the bill that there were no senators who discussed
        b.      this is the bill that I saw no senators who discussed
        c.      this is the bill that I met no senators who discussed

---

62      Engdahl (1982 *et seq.*) also notes this. One of her examples is reproduced in (i) (Engdahl 1982:159). Kush *et al.* (2009) give (ii) as an English example of the Swedish type, which they claim 'do[es] not seem too bad'. But while English does indeed reproduce to some extent cases of nominal-argumental extraction from what appears to be a relativised noun phrase (recall (141) and other examples given in the main text further below), calling (ii) acceptable stretches credibility: Cinque (1990) certainly had a point when he drew attention to the contrast between (141a,c).

(i)             där har jag en kusin som bor
                there have I a cousin that lives
(ii)            *there, I have an aunt that lives

63      Since a typical human's eyesight covers a limited distance, it is highly likely that the observer of the kissing event is in the same space where the kissing event takes place. But natural eyesight can be enhanced, for instance with the aid of binoculars, or a television set. The reading of (147a) that interests us here is one in which the observer is not in Red Square.

This small contrast suggests that a pseudorelative parse may facilitate extraction from highest-subject relatives in English. But above all, what stands out is that none of the sentences in (148) is judged anywhere near the maximum score of 7. Kush *et al.* found this in their other experiments as well. Extraction from highest-subject relative clauses thus never seems to be particularly good in English — but existential and perception-verb contexts ameliorate it to a certain extent.

The active ingredient in this 'extential/perception versus other' contrast in English does not, however, seem to be the (un)availability of a syntactic representation of the string following the matrix verb as a small clause (i.e., a pseudorelative). English simply does not support pseudo-relatives: (149a), the English translation equivalent of French (144a), does not have a reading in which the speaker reports witnessing the addressee's brother robbing his neighbour; and (149b), involving extraction of the lower verb's object, is significantly worse than, say, (141a), which suggests that the familiar specificity effect on extraction from complex noun phrases (Fiengo & Higginbotham 1981) is in effect here — unexpectedly, if pseudorelativisation could be involved.

(149)  a.      I saw your brother who robbed his neighbour
        b.      *the person who I saw your brother who robbed

I will therefore set aside the possibility of treating (marginally) successful cases of extraction from relative clauses in English as instances of pseudorelativisation: the relative clause is a genuine relative clause, part of a complex noun phrase. The effect of the matrix verbal environment on the efficacy of establishing what appears to be a filler–gap dependency across the boundaries of a complex noun phrase, which Erteschik-Shir, Allwood, Engdahl, and Kush *et al.* all prominently draw attention to, is certainly real. The paradigm in (150), for English, provides a somewhat broader spectrum of the facts. This empirical pattern calls for careful examination and explanation.

(150)  a.      these are the flowers that I know someone who sells
        b.      these are the flowers that I am looking for someone who sells
        c.      $^?$these are the flowers that I am trying to find someone who sells
        d.      $^{??}$these are the flowers that I met someone who sells
        e.      $^?$*these are the flowers that I was listening to someone who sells
        f.      *these are the flowers that I photographed someone who sells
        g.      *these are the flowers that I recognised someone who sells

In our quest for an analysis of this pattern, it will be useful to start by comparing the paradigm in (150) to the one in (151), involving resumptive prolepsis. (The proleptic object, which is a *for*-PP, and the resumptive pronoun are italicised for easy recognition.)

(151)  a.      these are the flowers *for which* I know someone who sells *them*
        b.      these are the flowers *for which* I am looking for someone who sells *them*
        c.      $^?$these are the flowers *for which* I am trying to find someone who sells *them*
        d.      $^{??}$these are the flowers *for which* I met someone who sells *them*
        e.      $^?$*these are the flowers *for which* I was listening to someone who sells *them*
        f.      *these are the flowers *for which* I photographed someone who sells *them*
        g.      *these are the flowers *for which* I recognised someone who sells *them*

In the examples in (151) we are not dealing with filler–gap dependencies across a relativised noun phrase at all. Yet we find an acceptability cline that looks just like the one seen in (150). What this suggests is that this cline is not about restrictions on long-distance filler–gap dependencies. Apparently , it is easy to associate a filler outside the relative clause with a gap inside it if the matrix verb is *know* or intensional *look for* or *try to find*. Similarly, a proleptic *for*-phrase linked to a resumptive pronoun inside the relative clause works well with these verbs. At the other end of the spectrum, with *listen to* and *photograph*, both a filler–gap dependency across the relative clause and the alternative resumptive prolepsis strategy yield very poor results. In the middle of the paradigm, in the d–examples, we find that the filler–gap strategy and resumptive prolepsis are both marginal.[64] But importantly, to the extent that the stategies work here at all, they do so only if it is the speaker's objective to acquire a bunch of flowers from the salesperson in question: (152a,b), for which there is no pragmatic link between the matrix and embedded clauses, are distinctly awkward. This recalls Allwood's (1982) observation that it is not just the choice of *matrix* verb but also that of the verb of the relative clause that influences the acceptability of extraction out of a relative clause.

(152)   a.      #these are the flowers that I met someone who ignored
        b.      #these are the flowers *for which* I met someone who ignored *them*

This last observation may suggest a connection with extraction from adverbial adjuncts, another 'strong island' type. Recall from chapter 2 that extraction from adjuncts, while commonly taken to be ungrammatical, yields virtually perfect results when there is a strong pragmatic link between the matrix eventuality and the one denoted by the adverbial phrase. Truswell (2011:157) formulates this link in his Event Grouping Condition: 'an instance of *wh*-movement is legitimate only if the minimal constituent containing the head and the foot of the chain can be construed as describing a single *event grouping*' (original italics). For cases of extraction from and resumptive prolepsis across highest-subject relatives we see a similar picture: it is much easier to envisage an 'event grouping' in (150/151d) than it is in (152); if a particular situation should make a pragmatic link between the two events easier to imagine, the status of the latter should improve.

While in chapter 2 I managed to syntacticise at least a subset of cases falling under the Event Grouping Condition in the realm of extraction across an adjunct boundary, I do not consider it feasible or worth one's while to provide a syntactic account for the acceptability cline in (150) and (151). This pattern is most likely not syntactic in nature: a pragmatic analysis is likely to be called for here. What interests me is not the clines *per se* but rather the fact that they match, and that this gives us a hint towards an analysis of (150), where we do not appear to be dealing with prolepsis on the surface.

We can understand the parallel if the syntax of (150) does in fact involve prolepsis: the operator in the relative clause associated with *flowers* is represented as a proleptic object in that relative clause. This paves the way for an understanding of the grammaticality of apparent long-distance filler–gap dependencies across a relativised noun phrase. Thinking of the analysis of the *convince*-class cases, above, we can represent them as cases of *wh*-control: the operator in the matrix clause controls a null-operator dependency in the embedded clause. (129d), repeated below, is the structure of apparent long-distance dependencies in *convince*-class constructions:

---

64      Recall that Kush *et al.* (2009) had likewise found that (147c) is deemed poor.

(129d)     *who* ... [$_{vP}$ *v* [$_{VP}$ *him* [$_{V'}$ *convince* [$_{PredP}$ Pred$_\emptyset$ [$_{pP}$ *p* [$_{PP}$ $t_{wh}$ [P [$_{CP}$ *that* S]]]]

Rather than featuring a single filler–gap dependency crossing an absolute island (i.e., the clausal complement of *convince*), what we find is a combination of two dependencies, one upstairs (involving a proleptic object) and the other downstairs (headed by a null operator, PRO). For the cases of apparent long-distance dependencies across a relativised noun phrase under current discussion, I presented an entirely similar analysis in (140) (for (139a)).

(140)     *who*$_i$ ... [$_{vP}$ *v* [$_{VP}$ *looking* [$_{pP}$ *p* [$_{PP}$ $t_i$ [P [$_{DP}$ *someone* [$_{CP}$ *Op*=PRO$_i$ [*who* ... $ec_i$]]]]]]]]]

An analysis along these lines makes it unnecessary to find ways to link a filler outside the complex noun phrase to a trace inside it: there simply is no filler–gap dependency established across the complex noun phrase at all; the *wh*-element in the matrix clause binds its trace there, not in the relative clause.

For the *convince*-class cases, finding a spot for the null operator in the embedded CP is not difficult: the SpecCP position of the subordinate clause is obviously available for PRO. For the relative clause cases that are the topic of the present section, things are not as simple. After all, the relative clauses under investigation are themselves introduced by an operator: *who* in all the English examples. This prompts a discussion of the special status of highest-subject relatives — the fact that they are the only ones allowing apparent long-distance dependencies to be established across them.

We can understand this very much in the way that Chomsky (1986) tries to understand the transparency of highest-subject relatives. The leading idea of Chomsky's brief discussion of the Chung & McCloskey (1983) data is that in the grammatical cases, the relative clause is introduced by an operator that can be taken to occupy the structural subject position of the relative clause — put differently, *who* is *in situ*; it finds itself in SpecTP rather than SpecCP, and thereby leaves SpecCP available for other purposes. For Chomsky, the other purpose was successive-cyclic movement of the other operator out of the relative clause, via SpecCP: in the course of the derivation of an example such as (141a), *which* makes an intermediate stop-over in the SpecCP of the relative clause prior to moving into the matrix clause. But landing in SpecCP on the way out of the relativised noun phrase does not actually help alleviate the locality problem that a single-chain derivation of examples such as (141a) incurs: the step from the SpecCP of the relative clause into the matrix VP will still violate Subjacency because it crosses two barriers — the CP (which is an inherent barrier) and the NP (which inherits barrierhood from CP). So leaving *who* in SpecTP inside the relative clause will not deliver an account for the grammaticality of (141a) and its ilk if *which* is to move out of the relativised noun phrase.

But if *which* binds a gap in the matrix clause and controls a null-operator dependency inside the relative clause, on the analogy of the analysis of *convince*-class constructions in the previous section, the hypothesis that *who* inside the relative clause is *in situ* in SpecTP procures exactly the right outcome. Now the fact that *who* does not occupy SpecCP in the relative clause is not providing the matrix-*wh* with an intermediate stop-over opportunity (which, as I showed above, would be fruitless) but gives it the chance to control a null operator, situated in SpecCP and binding a gap inside the relative clause. The structure in (153) now updates (140) and spells this out in detail.

(153)     *who*$_i$ ... [$_{vP}$ *v* [$_{VP}$ *looking* [$_{pP}$ *p* [$_{PP}$ $t_i$ [P [$_{DP}$ *someone* [$_{CP}$ *Op*=PRO$_i$ [C [$_{TP}$ *who* ... $ec_i$]...]

The fact that SpecCP is not 'used up' by the operator associated with the head of the relative provides us with exactly what we need: a place in which to put the null operator controlled by the matrix *wh*-operator. Whenever the relative clause is not a highest-subject relative, the operator linked to the head of the relative clause must itself be in SpecCP, thereby making it impossible for a null operator to be postulated in this position. So for non-highest-subject relatives, *wh*-control fails. A direct filler–gap dependency across the boundaries of the relative clause is also impossible, because the relative clause is an absolute barrier.[65]

The preceding paragraphs sketch a solution of the puzzle of why it is strictly impossible in English to extract anything from a relative clause *unless (a)* the relative clause is a highest-subject relative with an *in-situ* operator, and *(b)* the operator in the matrix clause can be construed as a proleptic object with the projection of the verb selecting the relativised noun phrase. The approach seems empirically effective for English; but for Scandinavian, questions remain — esp. the fact that PP and non-argument dependencies can apparently be established across a relativised noun-phrase boundary in these languages: recall (146). The prolepsis analysis cannot accommodate such examples: as Cinque (1990) already determined more than twenty-five years ago, the binding of a silent pronoun by a *wh*-filler across an island is the prerogative of nominal argument dependencies. For English, this is definitely the right result: the English equivalents of the examples in (146) are unacceptable. The fact that they are good in Swedish suggests that there may be something about *som*-relatives that makes them perhaps more amenable to a pseudorelative construal than English *that*- or *wh*-relatives are. I do not know of a concrete way to explore this further, so I will leave it at this.[66]

---

65      This account of 'extraction' from relativised noun phrases makes a further prediction. The ability to postulate a null operator in SpecCP, controlled by the matrix *wh*-element, is dependent on the availability of SpecCP — i.e., on the operator linked to the head of the relative not occupying this position. For *who* in (153), there is no particular reason why it should be unable to remain *in situ*, in SpecTP: we know that *wh-in-situ* is grammatical in principle in English, with overt *wh*-elements. But recall that null operators are, because of their status as PRO, prevented from occurring in licensed positions: PRO is, by definition, an unlicensed argument. Imagine, then, that we replaced *who* with a null operator. In highest-subject relatives, this null operator would have to be in SpecCP: leaving it in SpecTP would cause a crash because SpecTP is a licensed specifier. But with the null operator occupying SpecCP, in a *that*-relative, this makes it impossible to postulate a second null operator in that very same position. Hence, null-operator relatives are predicted to categorically resist 'extraction': not only can a direct filler–gap dependency not be established across the relative CP, this time around the *wh*-control option is unavailable as well because SpecCP is occupied by a different null operator (the one linked to the head of the relativised noun phrase). Jim McCloskey (p.c.) has confirmed to me that all the attested examples of extraction from highest-subject relatives that he has come across involve *wh*-relatives; but he adds that, for him, replacing *who* with *that* would not render the examples ungrammatical, which suggests that it is not an absolute requirement that the relative be introduced by a *wh*-operator. I do not have an account at this time of successful extraction from highest-subject *that*-relatives. In light of the text discussion, I am certainly committed to the idea that null operators cannot stay in SpecTP. A general possibility that comes to mind is that *that*-relatives can mimic the behaviour of highest-subject *wh*-relatives by analogy. But a question that would inevitably arise for such an approach is how a derivation that is rejected in syntax can somehow make it to PF by way of its grammatical analogue.

66      See also Sichel (2015) for interesting discussion of extraction from relative clauses in Hebrew. In Hebrew, as in Scandinavian, it is not the case that only nominal dependencies can be established across a relativised noun phrase: Sichel's key examples all involve PP-dependencies. Yet unlike in Scandinavian (and English), it does not appear to be the case that only highest-subject relatives are transparent. This suggests that something different is going on in (some of) the Hebrew examples — different not just from what I suggested for English in the main text, but also different from Kush *et al.*'s (2009) pseudorelative approach. I will not attempt to analyse the Hebrew facts here.

*5.3.18 Subject extraction from* wh-*islands*

While objects often manage to get out of *wh*-islands (and other 'weak islands') quite easily (see (154), discussed in chapter 3), subjects cannot build filler–gap dependencies across a *wh*-island at all, as (155) shows. I already presented an account of this '*wh*-trace effect' in chapter 3. Let us remind ourselves how the systematic ungrammaticality of the examples in (155) can be made to fall out from the analysis of filler–gap dependencies developed in this work.

(154)    a.      $^?$what can't you figure out why she read to him?
          b.      $^?$what can't you figure out whether she read to him?
          c.      $^{??}$what can't you figure out when she read to him?
          d.      $^?$*what can't you figure out how she read to him?
          e.      *what can't you figure out who she read to?

(155)    a.      *who can't you figure out why read this poem to him?
          b.      *who can't you figure out whether read this poem to him?
          c.      *who can't you figure out when read this poem to him?
          d.      *who can't you figure out what read to him?
          e.      *who can't you figure out who read this poem to?

     The answer turns out to be quite simple. To see this, consider the derivation in (156). Once the *wh*-filler at the edge of the subordinate CP comes along, it prevents *who* from reaching down into the embedded clause. But since $wh_2$ must be uploaded onto the filler stack, and since a previously downloaded filler can piggy-back on an uploading operation and can get uploaded along the way, it is possible in principle for *who* to join the *wh*-stack that $wh_2$ is a part of. When this happens, we get the stack depicted in (156a), where $wh_2$ is the innermost member of the stack and *who* sits on the outside.[67] When the stack is subsequently downloaded onto the edge of the TP in the subordinate clause, *who* is attached to TP first, followed by $wh_2$, as depicted in (156b). The derivation then continues to build the rest of the embedded clause and to postulate traces for the fillers on the edge of TP as necessary, as in (156c).

(156)    a.      $who_1$ ... [$_{CP}$ $wh_2$
                 STACK:
                 [$who_1$ [$wh_2$]]
          b.      $who_1$ ... [$_{CP}$ $wh_2$ [$_{TP}$ who$_1$ [$_{TP}$ wh$_2$
          c.      $who_1$ ... [$_{CP}$ $wh_2$ [$_{TP}$ who$_1$ [$_{TP}$ wh$_2$ [$_{TP}$ $t_1$ ... ($t_2$)]]]]]

The problem with (156c) manifests itself regardless of the nature of $wh_2$. Because of the fact that $who_1$ is at the top of the push-down stack of the embedded CP, it gets downloaded onto the edge of TP first, in keeping with the LIFO procedure. The other *wh*-element on the stack is subsequently placed in an adjunction position to TP as well, after which the rest of the TP is built. If $who_1$ had been downloaded last (or all by itself) onto the edge of TP, it could have served perfectly well as the

---

67      Here I represent the uploaded *wh*'s as 'wh' with an index rather than as PROs, to keep things transparent.

subject of predication right in its download position. But since the *wh*-stack in (156a) includes a second *wh*-element which must be downloaded onto TP after $who_1$ has already been placed on TP's edge, a direct predication relation between the *v*/VP and the subject on the edge of TP, with T as the RELATOR, is impossible. Postulating a trace of $who_1$ in SpecTP, as shown in (156c), is the only way to ensure a proper predication relation between *v*/VP and the subject (i.e., $who_1$). But $who_1$ cannot legitimately bind a trace in the SpecTP position because the link between the TP-adjoined copy of $who_1$ and $t_1$ is too short: it violates what has been called the 'anti-locality constraint' (Grohmann 2003, Abels 20xx, etc.). Phrased representationally, no element is allowed to be on the edge of the same phrase twice. And $who_1$ is in fact on the edge of TP twice: once after being downloaded onto TP as an adjunct, and then again in the guise of a trace in SpecTP. This rules all examples in (141) ungrammatical, regardless of whether the *wh*-element downloaded onto TP after $who_1$ binds a trace lower down or not: when there is no trace, no predication relation between *who* and the predicate of the lower clause can be established; when there *is* a trace, 'anti-locality' is violated. It is a no-win situation for *who*.

What happens when a Bresnan/Culicover-type high adverbial is inserted right after the *wh*-element introducing the subordinate clause? The fact of the matter is that the sentences in (155) do not improve under the influence of high-adverbial insertion: if anything, the examples in (157) are worse because of added complexity.

(157)   a.     *who can't you figure out why in all likelihood read this poem to him?
        b.     *who can't you figure out whether in all likelihood read this poem to him?
        c.     *who can't you figure out when in all likelihood read this poem to him?
        d.     *who can't you figure out what in all likelihood read to him?
        e.     *who can't you figure out who in all likelihood read this poem to?

In the presence of a high adverbial of the *in all likelihood*-type, a functional projection finds itself between C and TP — a RELATOR phrase, as in (158) (recall (68)).

(158)              $[C [_{RP} AdvP [R [_{TP} SU [T ...$

Since the RELATOR that introduces the high adverbial is silent and has no label to give to its projection, it must find some other head to co-project with. For English-type languages, T can play the role of co-projector for R. So in (158), R and T form a co-projection complex. The *wh*-stack that has *who* and $wh_2$ on it must be downloaded onto the edge of this co-projection complex. I assume that the grammar generally operates in such a way as to download filler stacks as far down the first RP in the tree as possible, to ensure that the downloaded fillers will be as close as possible to their prospective traces. The lowest position on the edge of the R+T co-projection complex is a position adjoined to TP (rather than RP, which is structurally higher). Downloading the *wh*-stack onto the edge of TP is thus preferred from the point of view of the general desideratum of downloading fillers as low in the tree as possible. Downloading the *wh*-stack onto TP's edge has the additional advantage of steering clear of potential intervention effects induced by the adverbial in SpecRP: especially adverbial downloads would be expected to be prevented from reaching across an adverbial in SpecRP when trying to associate with a trace further downstream. For these reasons, the grammar downloads the *wh*-stack onto the edge of TP, which results in (159).

(159)  a.  *who$_1$ ... [$_{CP}$ wh$_2$*
                      STACK:
                      [*who$_1$* [*wh$_2$*]]

  b.  *who$_1$ ... [$_{CP}$ wh$_2$ [$_{RP}$ in all likelihood* [R [$_{TP}$ who$_1$ [$_{TP}$ wh$_2$*

  c.  *who$_1$ ... [$_{CP}$ wh$_2$ [$_{RP}$ in all likelihood* [R [$_{TP}$ who$_1$ [$_{TP}$ wh$_2$ [$_{TP}$ t$_1$ ... (t$_2$)]]]]]

When we compare (159) with (156), it is easy to see that the two derivations create precisely the same problems for the subject-*wh*. It thus follows that insertion of a high adverbial does not alleviate '*wh*-trace effects'.

It is not the case that it is impossible *in general* to *wh*-extract a subject across a *wh*-island. As long as the extractee is not the subject of the *wh*-island itself but is extracted from a subject position further downstream, the result is often much improved:

(160)         $^?$who can't you figure out why John considers the best candidate?

On the top-down approach presented in this book, the fact that (160) presents no 'ECP effect' is perfectly straightforward. Recall that the problem with the derivation in (156) was that *who* was downloaded onto the outer edge of the very same TP in which it is supposed to serve as the subject of predication: it cannot be the subject of predication in its download position, nor can it bind a trace in SpecTP because a single element cannot be on the edge of the same phrase twice. Now consider (161), the schematic representation of (160):

(161)  a.  *who$_1$ ... [$_{CP}$ why$_2$*
                      STACK:
                      [*who$_1$* [*why$_2$*]]

  b.  *who$_1$ ... [$_{CP}$ why$_2$ [$_{TP}$ who$_1$ [$_{TP}$ why$_2$*

  c.  *who$_1$ ... [$_{CP}$ why$_2$ [$_{TP}$ who$_1$ [$_{TP}$ why$_2$ [$_{TP}$ John considers t$_1$ ...]]]]*

Here, the *who* that is downloaded onto the edge of the TP of the clause in the middle (the *wh*-island introduced by *why*) is NOT the subject of predication of this TP; it must establish a *wh*-dependency with a trace much further down the tree. Since *who* this time around is not trying to link up to a trace that is on the edge of the same TP that *who* is downloaded onto, no 'anti-locality effect' arises in (161), and the sentence is well-formed.

The theory of filler-gap dependencies presented in this book thus handles the 'ECP effects' seen in (155) adequately without getting into trouble with (160) or with the persistence of island effects in the presence of a high adverbial. The operative ingredient of the account of the facts discussed in this section has been 'anti-locality': the restriction that forbids an element from being on the edge of the same phrase twice.

### 5.3.19 *Multiple* wh-*fronting and Superiority*

The discussion of the '*wh*-trace effect' in the previous subsection paves the way for a brief look at the last item on our agenda: the fact that in constructions featuring multiple *wh*-dependencies, one of which involves the subject, it is always the subject dependency that gets superior treatment, so to speak: the Superiority Condition.

We see this condition reflected in English in the form illustrated in (162): in English-type languages, only one *wh*-element in a multiple *wh*-question can be in the left periphery, and whenever the subject is one of the *wh*-elements, it is the one that is pronounced there. In languages such as Bulgarian allowing multiple *wh*-fronting in CP, we see the Superiority Condition coming to the surface in the form of an ordering restriction, as shown in (163) (see Rudin 1988, Richards 1997, Bošković 1997 *et seq.*).

(162)   a.      who saw whom?
         b.      *whom did who see?
(163)   a.      koj     kogo    vižda?                                          (Bulgarian)
                 who     whom  sees
                 'who sees whom?'
         b.      *kogo   koj     vižda?
                 whom  who     sees

The derivations for (163) given in (164) and (165) illustrate how the ordering restriction on multiple *wh*-fronting in Bulgarian-type languages falls out from the top-down approach using push-down stacks. In the derivation in (164), with *koj* uploaded onto the stack first and hence, given the LIFO nature of stacks, downloaded last, the subject-*wh* ends up on the edge of the TP in precisely the right spot to serve, directly in its download position, as the subject of predication for the verbal predicate, with T as the RELATOR of the predication relation. No trace needs to be postulated for *koj*, and the other *wh*-element (*kogo*), encounters no difficulty establishing a link with its trace inside TP.

(164)   a.      $[_{CP}$ *koj*$_1$      $[_{CP}$ *kogo*$_2$      ...      $[_{TP}$
                        STACK:         STACK:
                        $[koj_1]$        $[kogo_2 [koj_1]]$
         b.      $[_{CP}$ *koj*$_1$      $[_{CP}$ *kogo*$_2$      ...      $[_{TP}$ kogo$_2$ $[_{TP}$ koj$_1$ $[_{TP}$ T ... $t_2$ ...]]]]

In the derivation in (165), by contrast, it is *koj* that gets downloaded first. Because *kogo* in (165b) obstructs the establishment of a direct predication relation, mediated by T, between *koj* and the verbal predicate,  (165b) fails in the absence of a trace for the subject ($t_2$). But insertion of a subject trace in SpecTP does not ameliorate the problem: with $t_2$ included, (165b) has the subject (*koj*) represented on the edge of the same projection (TP) twice, which is not allowed.

(165)   a.      $[_{CP}$ *kogo*$_1$ $[_{CP}$ *koj*$_2$      ...      $[_{TP}$
                        STACK:         STACK:
                        $[kogo_1]$        $[koj_2 [kogo_1]]$
         b.      *$[_{CP}$ *kogo*$_1$ $[_{CP}$ *koj*$_2$      ...      $[_{TP}$ koj$_2$ $[_{TP}$ kogo$_1$ $[_{TP}$ ($t_2$) ... $t_1$ ...]]]]

So just as in the case of the '*wh*-trace effect' cases discussed in the previous subsection, the subject just cannot do right in (163b).[68]

68      The parallel derived in the present theory between '*wh*-trace effects' and superiority effects involving the subject captures an insight that is also prevalent in Richards' (1997) approach to multiple *wh*-constructions (where the active ingredient is 'tucking in', which makes sense only in a bottom-up derivational approach diametrically opposed to the one pursued here) and in Tree Adjoining Grammar (which derives this parallel very elegantly; see Frank 2002, 2006).

For Superiority effects in English-style *wh-in-situ* languages, the same result as the one just obtained for Bulgarian-style multiple *wh*-fronting languages will emerge if the entire derivation for multiple *wh*-constructions unfolds in a single cycle (recall chapter 3): surface appearances notwithstanding, *who* and *whom* are both fronted to CP; the *wh*-element that occupies the highest position in CP is the one that is pronounced; any and all additional *wh*-elements in CP get their lower copies pronounced because of a PF restriction on the amount of material that the CP periphery tolerates in languages such as English.

From the discussion up to this point in this subsection, the impression might easily have emerged that whenever there are multiple *wh*-fronted constituents, they must always be lined up as in Bulgarian. This is not the case. The theory needs to provide for languages like Serbo-Croatian (see (166); Bošković 1997 and much subsequent work) or Hungarian (167), where we find multiple *wh*-fronting constructions with 'freedom' of ordering in the left periphery.

(166) a.     ko     je     koga   vidjeo?                  (Serbo-Croatian)
               who    is      whom   seen
      b.     koga   je     ko      vidjeo?
               whom   is      who    seen
               both: 'who saw whom?'
(167) a.     ki      mit          vett?                         (Hungarian)
               who    what-ACC     bought
      b.     mit          ki      vett?
               what-ACC     who    bought
               both: 'who bought what?'

To take care of this, we should take into account the fact that not all positions in the tree are created equal, and that the creation of push-down stacks is sensitive to the nature of the positions that fillers occupy. From the literature on multiple filler–gap dependencies, it has become clear that the nature of the positions occupied by the various fillers matters a great deal. So what we need to design is a theory in which sometimes multiple fillers are placed on the same push-down stack and at other times each filler is placed on a stack of its own — with the choice between the two options being based on the nature of the positions occupied by the fillers.

In particular, what we are seeking is a theory in which a single multi-member stack is created for multiple fillers whose positions are of the same type, and separate stacks are created for each filler whenever the positions they occupy are of different types. By 'type' I mean something very much like what Rizzi (1990) meant in his Relativised Minimality approach to the locality of syntactic dependencies. Concretely, a *wh*-operator occupying SpecCP is put on a different stack from the one that harbours *wh*-operators occupying SpecDistP or SpecTopP positions, interpreted as universal quantifiers, which are in turn on a different stack from ones that are interpreted as bare existential quantifiers ('something', 'someone'), etc. For Hungarian (167) we know, from É. Kiss's (1993) work, that the immediately preverbal *wh*-element is in a different type of structural position from the *wh* that precedes it: SpecFocP *vs* SpecDistP. For Serbo-Croatian, the situation appears to be very much the same. In Serbo-Croatian (166) and Hungarian (167), therefore, we have *two* push-down stacks, one for universals and one for foci, and each happens to have a single member, because of the simplicity of the example.

The downloading of the two stacks onto the edge of the first RP in the structure is unordered: though *within* a single pushdown stack, the items on the stack are downloaded in a strict LIFO order, multiple pushdown stacks are unordered *vis-à-vis* one another. Since the two separate pushdown stacks in (166) and (167) each have exactly one member, and since the downloading of the individual stacks is unordered, problems of path containment cannot arise in languages whose multiple *wh*-fronting constructions target discrete positions in the left periphery.[69]

## 5.4   Summary and conclusion

We have now arrived at an account of the distribution of subjects, their traces, and PRO that integrates the ECP and the EPP (and the Case Filter into the bargain) and in addition provides a perspective on the cross-linguistic variation in this realm. The key player in this theory is Agree, defined in such a way that it recognises that this structural relation is non-directional in principle but at the same time distributes Upward and Downward Agree with reference to the nature of functional heads, RELATOR heads being the only ones simultaneously capable of engagement in Agree in both directions. The active ingredients in the analysis of the licensing of subjects are brought together in (168):

---

69      Some more needs to be said. For when we make the Hungarian example a bit more complex such that it contains more than one pre-focal *wh*-element, we find that the *wh*-elements preceding the focus, while all preceding the focus as a block, can freely change places amongst each other. We see this in (i), where the curly brackets indicate that the relative ordering of the elements enclosed within them is free in principle. In this regard, they behave exactly like topics, which are unordered *vis-à-vis* one another. This is illustrated in (ii), the answers to (i). If all topics end up in the same, single push-down stack, we will not manage to allow for this ordering freedom: the LIFO nature of pushdown stacks is such that these will always allow for just one outcome.

| (i) | a. | {ki | kinek} | MIT | vett? | (Hungarian) |
|---|---|---|---|---|---|---|
| | | who | who-DAT | what-ACC | bought | |
| | b. | {kinek | mit} | KI | vett? | |
| | | who-DAT | what-ACC | who | bought | |
| | | both: 'who bought what for whom?' | | | | |
| (ii) | a. | {János | Marinak} | CSAK EGY KÖNYVET | vett | (Hungarian) |
| | | János | Mari-DAT | only a/one book-ACC | bought | |
| | | 'János bought Mari only a/one book' | | | | |
| | b. | {Marinak | egy könyvet} | CSAK JÁNOS | vett | |
| | | Mari-DAT | a book-ACC | only János | bought | |
| | | 'only János bought Mari a book' | | | | |

This problem, however, is to a large extent unique to the topic function. We know that topics are particularly prone to a filler–gap construal that is different from the one resorted to in garden-variety movement dependencies: whereas moved elements routinely bind a trace, topics have a semantic property (viz., specificity) that makes them eminently eligible for binding a resumptive pronoun, which may itself be silent. We will, at some point, need to come to terms with the way pronominals link up with their antecedents; but this relationship is certainly very different from movement-type filler–gap dependencies, and in all likelihood cannot be treated in terms of pushdown automata. I will set pronominal binding dependencies aside in this work. On the assumption that topic–gap dependencies can in principle be treated in terms of silent resumption, the fact that there is freedom in the placement of *wh*-elements preceding the focal *wh* in languages like Serbo-Croatian or Hungarian will not be on our agenda.

(168)  a.        a condition on the licensing of A–specifiers (A–SLiP in (50a))
       b.        a condition on the licensing of traces in specifier positions (T–SLiP in (50b))
       c.        a parameter regarding the typological distribution of co-projection of T (63)
       d.        a definition of PRO as an unlicensed argument (76)

Much of this is already present, in one way or another, in extant theories. No stipulations are added to this toolkit. The formulations of A–SLiP and T–SLiP are very similar in nature and both couched in terms of the one major structural relation in the syntactic tree: Agree. The result is eminently minimalist, and empirically highly adequate.

       By way of a summary, let me go through the list of problems posed by the subject that I provided at the opening of this chapter:

•        I have derived 'EPP' effects [*(i)*] with the help of A–SLiP (which forces a subject into SpecTP for licensing purposes in languages like English), in conjunction with the fact that T, whenever it serves a RELATOR role, must have a specifier by its nature as a RELATOR.

•        I have derived the fact that the 'EPP' is suspended in control infinitives (i.e., the fact that no SpecTP position is projected in such infinitives) [also part of *(i)*] from SLiP in conjunction with a definition of PRO that says that it is an unlicensed specifier.

•        I have derived the 'complementiser-trace effect' and the '*do*-trace effect' [*(ii)*], and more generally the subject/object asymmetry encoded by the ECP, with the help of T–SLiP and obligatory co-projection of C and T.

•        I have derived the most robust case of 'superiority' (viz., the fact that the subject-*wh* cannot remain *in situ*) [*(iii)*] from the general perspective on filler–gap dependencies that informs this work (in particular, the LIFO nature of the mechanism by which *wh*-stacks are downloaded onto the left edge of a predication structure), in conjunction with a condition that bans an element from being on the edge of the same projection twice.

•        I have derived '*wh*-trace effects' for subjects [*(iv)*] in the same way as superiority effects.

•        I have derived the effect of insertion of a high adverbial in the left periphery on the grammaticality of certain long subject dependencies (the Bresnan/Culicover facts) [*(v)*] with an appeal to T–SLiP and co-projection, and have derived the lack of an effect of high adverbials on '*wh*-trace effects' from co-projection and the plausible hypothesis that the grammar downloads a stack as far down the first RP in the tree as possible.

•        I have derived the fact that long subject extraction from finite clauses embedded inside islands is ungrammatical regardless of the presence or absence of a complementiser [*(vi)*] from the '*wh*-control' hypothesis, together with SLiP and the conceptually minimal(ist) assumption that PRO-headed chains do not undergo copy-to-trace conversion.

- I have derived the parallel behaviour of null-operator constructions [*(vii)*] and *convince*-type constructions (and their relatives) [*(viii)*] when it comes to long subject dependencies by treating both in terms of PRO-headed dependencies, and in the process, I have identified an '*of*-trace effect' structurally parallel to '*for*-trace effects' and '*that*-trace effects'.

- I have derived the 'anti-*that*-trace effect' of highest-subject relative clauses [*(ix)*] by analysing them as null-operator constructions and thereby preventing co-projection of C and T.

- I have derived the fact that highest-subject relatives are often (though not systematically) transparent for the establishment of filler–gap dependencies across their boundaries [*(x)*] by chopping up the relevant dependencies into two parts, a matrix dependency involving a proleptic object and a null-operator dependency inside the relative clause, the latter controlled by the former (another case of '*wh*-control').

In conclusion, I think it is fair to say that the SLiP, in conjunction with a definition of PRO, considerations of economy of derivation and representation, and a top-down approach to the building of syntactic structures and the filler–gap dependencies established inside them does a solid amount of work in accounting for the intricate tangle of facts surrounding subjects and their dependencies.