COPiL CAMBRIDGE OCCASIONAL PAPERS *in* LINGUISTICS
Volume 9, Article 7: 184–211, 2016 | ISSN 2050-5949

UNIFYING THE *THAT*-TRACE AND ANTI-*THAT*-TRACE EFFECTS*

JAMIE DOUGLAS
UNIVERSITY OF CAMBRIDGE

ABSTRACT

This article proposes a unified analysis of the *that*-trace and anti-*that*-trace effects in English. Unification of these two seemingly diametrically opposed effects remains an outstanding problem in the literature. Assuming a split CP and an anti-locality condition banning movement from SpecTP to SpecFinP, a novel analysis is set out consisting of a number of postulates that have theoretical implications for phase theory, the mechanics of successive cyclicity, and the distinction between final and intermediate landing sites.

1 INTRODUCTION

The English *that*-trace effect is a well-known and long-standing problem. Numerous analyses have been proposed since it was first observed by Perlmutter (1968, 1971) but no consensus seems to have been reached. It is also well-known, though seldom more than mentioned in *that*-trace analyses, that English short subject relative clauses do not exhibit the *that*-trace effect and that they exhibit a reverse or anti-*that*-trace effect. The anti-*that*-trace effect is thus doubly unexpected from the perspective of the *that*-trace effect, and vice versa. The co-existence in English of these two seemingly mutually incompatible and diametrically opposed effects is a serious outstanding puzzle (Pesetsky 2015). In this article, we will address this challenge and propose a unified analysis of both the *that*-trace and anti-*that*-trace effects.

It is argued that the anti-*that*-trace effect results straightforwardly from an anti-locality condition. Assuming such a condition to be correct, it is then shown that, when the *that*-trace effect paradigm is considered, we are led to several interesting theoretical postulates concerning the mechanics of phases, successive cyclicity, and a distinction between final and intermediate landing sites.

With the advent of Minimalism and the abandonment of the Empty Category Principle (ECP), anti-locality accounts of the *that*-trace effect have been gaining currency in the literature (Erlewine 2014, Ishii 1999, 2004, Pesetsky & Torrego 2001,

* Acknowledgments: My thanks go to Ian Roberts, Theresa Biberauer, Michelle Sheehan, Sten Vikner, the audience of SyntaxLab (27th) January 2015, University of Cambridge), and the audience of LEL Postgraduate Conference (5th June 2015, University of Edinburgh). An extended version of this article is in progress.

2007). The present work continues this trend. However, as far as I am aware, no anti-locality account has been offered that unifies both the *that*-trace and anti-*that*-trace effects.

The structure of this paper is as follows: in section 2, the *that*-trace and anti-*that*-trace effects are introduced and their similarities and differences are evaluated. In section 3, a novel and unified analysis of these effects is proposed and a number of postulates concerning phases and successive cyclicity are formulated. Section 4 explores the implications of the analysis for phase theory in a cartographic context. Section 5 considers some of the benefits of the analysis from the perspective of acquisition. Section 6 concludes.

2 The *that*-trace and anti-*that*-trace effects

Let us start by illustrating the *that*-trace and anti-*that*-trace effects. First, note that the complementiser *that* is generally optional in finite declarative clauses.

(1) a. You said John saw Mary.
b. You said that John saw Mary.

This optionality is preserved under long-distance *wh*-extraction.

(2) a. Who did you say John saw t?
b. Who did you say that John saw t?

Crucially, however, this optionality disappears in cases of long-distance *wh*-extraction of subjects.

(3) a. Who did you say t saw Mary?
b. *Who did you say that t saw Mary?

The presence of *that* in (3b) results in ungrammaticality[1]. This is the *that*-trace effect (Perlmutter 1968, 1971).

Turning now to the anti-*that*-trace effect[2], note that *that* is also generally optional in relative clause (RC) contexts.

(4) a. The woman [John saw t] was tall.
b. The woman [that John saw t] was tall.

[1] For some speakers, though not for all; see Chomsky & Lasnik (1977) and especially Sobin (1987, 2002) for English; Lohndal (2007, 2009) for *that*-trace variation in Scandinavian languages; and Maling & Zaenen (1978) for *that*-trace violations in Dutch and Icelandic. Experimental evidence suggests that extraction with *that* is degraded for both subjects and objects relative to extraction without *that*, and that subject extraction with *that* is degraded for subjects relative to extraction with *that* for objects (Cowart 1997). Therefore, although absolute judgements vary, the relative judgements are statistically robust.

[2] The term 'anti-*that*-trace effect' has also been used to describe the observation that *that* becomes almost obligatory for some speakers when there is fronted material (Pesetsky & Torrego 2001: 376). I use the term 'anti-*that*-trace effect' to refer exclusively to the effect found in short subject relative clauses.

The optionality in (4) also disappears in cases of short subject relativisation. However, in RC contexts, it is the absence of *that* which results in ungrammaticality[3].

(5) a. *The man [t saw Mary] was short.
b. The man [that t saw Mary] was short.

This appears to be the reverse of the *that*-trace effect, hence the name anti-*that*-trace effect.

Numerous accounts of the *that*-trace effect have been proposed over the half-century since it was first described (Abe 2015, Boeckx 2008, Bošković 1994, 1996, 1997, Bresnan 1972, 1977, Brillman & Hirsch 2015, Browning 1996, Chomsky & Lasnik 1977, Deal 2015, Doherty 1993, 2000, Erlewine 2014, Ishii 1999, 2004, Kandybowicz 2009, Lohndal 2009, Perlmutter 1968, 1971, Pesetsky & Torrego 2001, 2007, Pesetsky 1982, Richards 2001, Rizzi & Shlonsky 2007, Rizzi 1982, 2014, Roussou 2002, 2010, Sobin 1987, 2002, Taraldsen 1978), but few (if any) have attempted a *unified* account of both the *that*-trace and anti-*that*-trace effects. One potential reason for this is that, as pointed out above, each effect is apparently doubly unexpected from the perspective of the other, thereby making a unified analysis very difficult to imagine.

Of course, I am making the assumption that these two effects should be unified. If one does not believe they should be, the whole problem of finding a unified account vanishes. However, I believe that a number of considerations suggest that a unified analysis is at least not implausible from the outset.

First, as the names of the effects suggest, both effects involve the presence/absence of *that* in finite subordinate clauses (note that short subject RCs may use a *wh*-relative pronoun instead of *that*). Admittedly, the subordinate clauses are of different types, i.e. RCs and verbal complement clauses. One could argue that *that* in these different clause types is not the same element. Indeed, in many languages, the morphological form of relative markers and clausal complementisers is different. Furthermore, subject relative markers may differ from non-subject relative markers. However, in English, the morphology does not motivate such distinctions. If we take seriously the idea that UG is impoverished (Berwick & Chomsky 2011, Biberauer & Roberts 2015, Boeckx 2014, Chomsky 2005), distinctions made in one language cannot simply be used as arguments for those same distinctions in another. Instead, we must seriously consider the idea that different languages will make different grammatical distinctions based on the Primary Linguistic Data used during acquisition.

[3] There is a small class of exceptions. Consider (i):

(i) There's a man sells bread at the market.

However, McCawley (1998: 460-463) observes that the contexts for such exceptions are 'existential' sentences (broadly construed) and that RCs in these contexts behave differently from normal restrictive RCs in a number of respects (e.g. extraction possibilities, co-occurrence with proper names, ability to insert parentheticals). He thus calls the RC-like structures in 'existential' contexts pseudo-relative clauses (see also, e.g. Cinque 1995: ch.8, Koopman & Sportiche 2014). See Doherty (1993, 2000) for arguments against the pseudo-relative claim.

Second, both effects involve subject extraction. The optionality in the presence/absence of *that* present in complement clauses and RCs disappears in the context of subject extraction. Furthermore, the condition on optionality only affects the *that* which is local to the original extraction site of the subject. Other instances of *that* remain optional.

(6) a. Who did Bill think (that) you said t saw Mary?
b. *Who did Bill think (that) you said that t saw Mary?

This shows that the *that*-trace effect is not tied to subjects per se, but rather to the initial movement from subject position.

Third, whilst short subject RCs exhibit the anti-*that*-trace effect, long-distance subject RCs exhibit the *that*-trace effect (compare (7a)-(7d) where there is no overt complementiser introducing the clause embedded inside the RC, with (7c)-(7d) where the complementiser introducing the clause embedded inside the RC is overt). Furthermore, long-distance subject RCs do not exhibit an anti-*that*-trace effect (the *that* introducing the RC is optional, as (7a) and (7b) show).

(7) a. The man [**that** Bill said t saw Mary] was short.
b. The man [Bill said t saw Mary] was short.
c. *The man [**that** Bill said **that** t saw Mary] was short.
d. *The man [Bill said **that** t saw Mary] was short.

These facts show that the anti-*that*-trace effect is not related to relativisation *per se*, nor to relativisation of subjects *per se*. Instead, the anti-*that*-trace effect arises in cases of relativisation of the highest subject in a clause. In contrast, the *that*-trace effect arises in cases of extraction of a subject embedded in a clause. This is an important point: whilst both effects involve *that* and subject extraction, they differ in that the *that*-trace effect involves extraction across a clause boundary whilst the anti-*that*-trace effect does not. This generalisation does not seem to have received much consideration in the literature. It will be crucial to the analysis developed in what follows.

The striking similarities suggest that the two effects should be unified and the observation that the two effects differ in terms of whether a clause boundary is crossed or not suggests a way of approaching the problem. The major questions for a unified analysis of the *that*-trace and anti-*that*-trace effects are thus: (i) why do both effects involve the presence/absence of *that*? (ii) why do these effects only arise local to the original extraction site? (iii) why do these effects differ in terms of whether a clause boundary is crossed or not? These questions will guide our investigation.

3 An analysis

The structure of this section is as follows: Section 3.1 will make some remarks on previous accounts of the *that*-trace effect and obstacles to a unified account of *that*-trace and anti-*that*-trace effects. Section 3.2 will lay out my assumptions about

clause structure. Section 3.3 presents a hypothesis based on an intuitive analysis of the anti-*that*-trace effect, and Section 3.4 is an exploration of the consequences of this hypothesis when applied to the *that*-trace paradigm. Section 3.5 is a summary of the major findings.

3.1 Some remarks on previous accounts

There are too many previous analyses of the *that*-trace effect in the literature to be able to give a comprehensive review here (see Pesetsky 2015 for an overview). Instead, I will suggest some reasons why previous approaches do not lend themselves to a unified analysis of the *that*-trace and anti-*that*-trace effects (admittedly this is not their primary goal).

The vast majority of *that*-trace analyses assume a single C head, i.e. the C domain is not split. This applies not just to proposals pre-dating Rizzi's (1997) seminal paper, but also to many proposals since (Brillman & Hirsch 2015, Erlewine 2014, Ishii 1999, 2004, Pesetsky & Torrego 2001, 2007, Roussou 2002, 2010), although there are some notable (albeit unsurprising) exceptions (Rizzi & Shlonsky 2007, Rizzi 2014, Lohndal 2009). Economy plays a significant role in minimalist accounts of the *that*-trace effect. Although the details and implementation vary, there is something of a consensus that the presence of *that* renders movement of a *wh*-subject uneconomical in some way. Pesetsky & Torrego (2001), for instance, propose that C has unvalued Wh and T features, uWh and uT respectively, each associated with an EPP property. The EPP property of C's uWh feature is satisfied by moving a *wh*-element to SpecCP. The EPP property of C's uT feature, however, is satisfied in one of two ways: either T moves to C, resulting in C being spelled out as *that*, or the nominative subject moves from SpecTP to SpecCP, nominative case being a uT feature on a nominal. Now, if the EPP properties of both the uWh and uT features on C can be satisfied by one operation rather than two, Economy will dictate that only the one operation option is permitted. A *wh*-subject, by virtue of having both Wh and T features, can satisfy both EPP properties of C's uWh and uT features. This is more economical than moving T to C (resulting in *that*) to satisfy uT's EPP property and moving the *wh*-subject to satisfy uWh's EPP property. Consequently, movement of the *wh*-subject to SpecCP is incompatible with *that* by Economy.

If there is only one C head, it is very difficult to imagine how an Economy-based approach can account for both the *that*-trace and anti-*that*-trace effects in a unified way. Whilst it is conceivable that C may have multiple exponents depending on its featural specification, and that these may impact on a *wh*-subject's extraction possibilities, it is no longer clear that we are dealing with the 'same' derivations. If this is the case, Economy has no role to play. The same problem arises on a strongly lexicalist approach, since in these cases we would presumably be dealing with different numerations. For this reason, it seems more promising to me to pursue the intuition underlying Economy-based proposals, but to adopt a split CP, which is independently motivated, so as to capture both the *that*-trace and anti-*that*-trace effects in a unified way.

Another hurdle to a unified analysis relates to the major concern of many of the

previous approaches. Approaches based on the Nominative Island Condition (e.g., Pesetsky 1982, Taraldsen 1978) and the Empty Category Principle (e.g., Doherty 1993, 2000, Rizzi 1982) both focussed on the licensing conditions of the empty subject element in SpecTP. As pointed out above, however, there is a significant difference between the *that*-trace and anti-*that*-trace effects in terms of whether a clause boundary is crossed or not. By focussing on the licensing of the empty subject in SpecTP, attention is diverted away from what happens to the *wh*-extracted subject after moving from this position. A similar objection could be mounted against a Subject Criterion approach (Rizzi & Shlonsky 2007, Rizzi 2014). According to such approaches, the *that*-trace effect arises because of a special property of SpecTP, namely the hypothesis that it is a criterial position. As will be seen, we will argue that the *that*-trace and anti-*that*-trace effects have more to do with a distinction between final and intermediate landing sites in the C domain than with any special property of SpecTP.

Overall the tendency seems to have been to ignore or set aside the anti-*that*-trace effect as an outstanding problem or curiosity (Chomsky & Lasnik's (1977) well-known *that*-trace filter simply includes an exception clause for it). Where the anti-*that*-trace effect has been considered in more detail, the tendency has been to give *that* in RCs (or possibly only *that* in short subject RCs) a different analysis from the one given to *that* elsewhere (e.g., Doherty 1993, 2000, Rizzi & Shlonsky 2007, Roussou 2002, 2010). Whilst this is certainly a possibility (and highly plausible for languages that make a lexical distinction in such cases), if we take the syncretism in English complementisers seriously (see, e.g., Baunaz & Lander submitted, Baunaz 2014) from an acquisition perspective, it is not so clear that having different analyses is *a priori* plausible for English. Furthermore, it would leave unexplained the striking similarities between the two effects pointed out above.

3.2 Assumptions concerning clause structure

Before turning to the analysis proper, it is necessary to clarify and justify my assumptions concerning clause structure.

For RCs, I adopt the Matching Analysis following Salzmann (2006). According to this analysis, the RC head, i.e. the nominal modified by the RC, is external to the RC itself. The RC head is matched by a copy inside the RC itself. The RC-internal copy moves to the edge of the RC, thereby creating a predicate (Heim & Kratzer 1998). The RC predicate is then predicated of the RC head (only the RC-external copy of the RC head is pronounced).

(8) the man that John saw
$[_{DP}$ the $[_{NP}$ [man$]_k$ $[_{CP}$ [man$]_i$ that $[_{TP}$ John saw $t_i]]]]$

As will be seen, our analysis of the *that*-trace and anti-*that*-trace effects takes place entirely inside CP. Consequently, I will remain agnostic concerning exactly how the predication relation of the Matching Analysis is to be captured syntactically. Furthermore, for simplicity, only the CP structure of RCs will be illustrated in what follows.

More generally, I assume that the CP is split into at least two distinct functional projections: ForceP and FinP, headed by Force and Fin respectively, where Force is higher than Fin (Rizzi 1997). Whilst Rizzi (1997) assumes that Force and Fin are split only if the Topic/Focus field is activated, I assume that they are always split. Also following Rizzi (1997), I take *that* to be in Force and the null complementiser, *Ø*, to be in Fin[4]. In other words, *that* and *Ø* are independent heads in the C domain; they are *not* phonological variants of the same functional head. This applies to both complement clauses and RCs. Treating *that* as a complementiser head in RCs is a widespread idea (Bhatt 2002, Bianchi 1999, 2000a,b, De Vries 2002, Kayne 1994), although some have recently argued that that is a relative pronoun (Arsenijević 2009, Kayne 2014, Manzini 2014).

A number of authors have proposed that *that*-less finite declarative clauses and *Ø*-RCs are actually TPs (=IPs), not CPs (Bošković 1994, 1996, 1997, Doherty 1993, 2000, Grimshaw 1997, Weisler 1980). I adopt the conclusion that such clauses are not full CPs, i.e. ForcePs, but I propose that what the authors above identified as TPs are actually FinPs.

Starting with RCs, it has been observed that *that*-RCs and *Ø*-RCs exhibit no discernible interpretive difference (Doherty 1993, 2000). Syntactically, I take this to mean that both involve Ā-movement (creating an Ā-chain) in the RC (Chomsky 1977). This is interpreted as lambda abstraction and turns a proposition into a predicate (Heim & Kratzer 1998). In *that*-RCs, the target of Ā-movement is SpecForceP, whilst in *Ø*-RCs, the target of Ā-movement is SpecFinP. The ForceP or FinP is then predicated of an external copy of the RC head in line with the Matching Analysis of RCs (Salzmann 2006). In what follows, I will only illustrate the structure of the RC itself (the bracketed constituent). RC-internal copies of the RC are not pronounced; the pronounced copy of the RC head is external to the RC itself. Traces will be used throughout.

(9) a. the man [that John saw]
[$_{\text{ForceP}}$ man$_i$ [$_{\text{Force'}}$ Force=that [$_{\text{FinP}}$ Fin=*Ø* [$_{\text{TP}}$ John [$_{\text{T'}}$ T saw t$_i$]]]]]
b. the man [John saw]
[$_{\text{FinP}}$ man$_i$ [$_{\text{Fin'}}$ Fin=*Ø* [$_{\text{TP}}$ John [$_{\text{T'}}$ saw t$_i$]]]]

If *Ø*-RCs were TPs, there would be no available target for Ā-movement and hence no clausal predicate could be formed.

Turning now to verbal complement clauses, observe that those introduced by *that* allow adverbial fronting, as in (10). Fronted adverbials cannot appear to the

[4] Boeckx (2008) and Lohndal (2007, 2009) take *that* to be in Fin when *that*-trace effects obtain but propose that *that* is in Force when *that*-trace effects are obviated (see especially Boeckx (2008: Section 5.5.2, 185-190)). Rizzi & Shlonsky (2007) propose that *that* is either syncretic (expressing both Force and Fin) or is merged in Fin and moves to Force. They suggest that in dialects which lack the *that*-trace effect, Force and Fin are separate with *that* realising Force. They also propose that this option "is invariably available for subject relatives, presumably because of the functional need to have a device to express relatives on all major argument positions, subjects in the first place" (Rizzi & Shlonsky 2007: 149). This raises non-trivial problems for the relation between function and form. Furthermore, there are questions concerning how the differences in the position of *that* are acquired (see Boeckx 2008: 189).

left of *that* with an embedded construal, as in (11). (10) and (11) are from Doherty (2000: 15):

(10) a. She prayed that *next Wednesday* the check would arrive.
b. We concluded that *in the future* he should be closely watched.
c. We maintain that *in Dublin* good coffee is hard to find.
d. John claims that *during the party* Ted squirted water at Eric.

(11) a. *She prayed *next Wednesday* that the check would arrive.
b. *We concluded *in the future* that he should be closely watched.
c. *We maintain *in Dublin* that good coffee is hard to find.
d. *John claims *during the party* that Ted squirted water at Eric.

These data suggest that fronted adverbials are positioned lower than Force, consistent with Rizzi's (1997) fine-grained left-periphery. Doherty (2000: 16) claims that fronted adverbials in *that*-less finite declarative clauses are "robustly ungrammatical" (see also Pesetsky & Torrego 2001, Rizzi 1997, 2014). However, whilst these examples are degraded, I do not think they are completely ungrammatical and I am not alone in this (see Sobin, 1987). Consider the following, adapted from (10) and (11).

(12) a. She prayed next Wednesday the check would arrive.
b. We concluded in the future he should be closely watched.
c. We maintain in Dublin good coffee is hard to find.
d. John claims during the party Ted squirted water at Eric.

With the appropriate intonation, the fronted adverbials in (12) can be interpreted with the embedded clause, contrary to Doherty's claim. Nonetheless, there does seem to be a contrast.

Returning to RCs, *that*-RCs also allow fronted material such as fronted adverbials, topicalisation and negative preposing. The following examples and judgements are from Bianchi (1999: 177).

(13) a. This is the kind of car that [for my son] I wouldn't even have considered buying.
b. ?I saw a dress that [under no circumstances] would I have considered buying for my daughter.

The data in (13) resemble those in (10). Furthermore, the subject-auxiliary inversion in (13b) provides evidence for the assumption that there is a functional head, which I equate with Fin, in the C domain lower than Force. It also shows that the fronted material in (13) targets a position lower than Force but higher than Fin.

In contrast, Ø-RCs are incompatible with fronted adverbials, topicalisation and negative preposing (Bianchi 1999: 177).

(14) a. *This is the kind of car [for my son] I wouldn't even have considered buying.
b. *I saw a dress [under no circumstances] would I have considered buy-

ing for my daughter.

The contrast between *that*-RCs and Ø-RCs follows if *that*-RCs contain more structure than Ø-RCs such that Ø-RCs lack the structure present in *that*-RCs that is required to host fronted material. Specifically, Ø-RCs lack any projections above FinP (this includes the Topic/Focus field of Rizzi 1997) and so they cannot host fronted adverbials, topics or preposed negative phrases in their left periphery. However, *that*-RCs are ForcePs and so can include the Topic/Focus field.

Now recall that a similar contrast is observed in complement clauses, except that *that*-less complement clauses do seem to allow fronted material in some cases[5].

(15) a. ?*I met the man at a party last night you were about to marry.
b. I met the man at a party last night that/who you were about to marry.

(16) a. *I met the man you wanted to marry Sue likes.
b. I met the man you wanted to marry that/who Sue likes.

FinP clausal complements, however, have no such predication requirement and so Topic/Focus projections may be present above FinP. In other words, Ø-RCs are FinPs whilst *that*-less complement clauses may be FinPs or something slightly larger, e.g. TopicPs. Nevertheless, both are smaller than ForcePs.

To summarise, I have argued that clauses introduced by that and *that*-RCs are ForcePs, whilst *that*-less clauses and Ø-RCs lack a ForceP projection. Henceforth, *that*-clauses and *that*-RCs will be modelled as ForcePs, and Ø-clauses and Ø-RCs will be modelled as FinPs (note that the argumentation thus far says that that lexicalises Force and Ø lexicalises Fin, but it says nothing about why, i.e. it does not explain why *that* should lexicalise Force rather than Fin. I will return to this point in section 5).

*3.3 The anti-*that*-trace effect*

With our assumptions concerning clause structure in place, we turn now to the analysis of the *that*-trace and anti-*that*-trace effects. In a break from tradition, let us first consider the anti-*that*-trace effect. Short subject *that*-RCs have the following structure (for convenience, I show the subject originating in SpecTP):

(17) The man [that saw me] was tall.
$[_{\text{ForceP}}\ [_{\text{DP}}\ \text{man}]_i\ [_{\text{Force'}}\ \text{Force=that}\ [_{\text{FinP}}\ \text{Fin=Ø}\ [_{\text{TP}}\ t_i\ [_{\text{T'}}\ \text{T saw me}]]]]]$

In (17), the subject *man* moves from SpecTP (the canonical subject position) to SpecForceP. This establishes the Ā-dependency (the operator-variable relation) required in RCs (Chomsky 1977).

[5] I suggest that the difference between (12) and (14) arises from a requirement that the FinP in Ø-RCs must be directly predicated of the RC head (recall that we are assuming a Matching Analysis of RCs). Projections within the C domain that are higher than FinP, e.g. Topic and Focus projections, would block this direct predication relation. Such a requirement may also account for the inability of Ø-RCs to extrapose, as in (15), or stack, as in (16) (note that *that*-RCs and *wh*-RCs can be extraposed and can stack, as in the (b) examples. I will return to this in the next subsection).

Now consider the structure of the ungrammatical short subject Ø-RC:

(18) a. *The man [Ø saw me] was tall.
b. *[$_{FinP}$ [$_{DP}$ man]$_i$ [$_{Fin'}$ Fin=Ø [$_{TP}$ t$_i$ [$_{T'}$ T saw me]]]]

(18) exhibits the anti-*that*-trace effect. The subject *man* moves from SpecTP to SpecFinP to establish the Ā-dependency, but the result is ungrammatical. Why should this be? For now, we will simply postulate that movement from SpecTP to SpecFinP is banned, a type of anti-local movement (see also Erlewine 2014, 2015).

(19) Movement from SpecTP to SpecFinP is prohibited.

In section 4.3, I will attempt to derive the ban on movement from SpecTP to SpecFinP via labelling and economy.

(19) specifically rules out movement to SpecFinP from SpecTP; it does not rule out movement from another position such as SpecvP (hence the reason for calling (19) a type of anti-local movement). The ban in (19) thus captures the fact that the anti-*that*-trace effect applies to highest subjects only. Other RCs do not involve such anti-local movement. The structures for a direct object *that*-RC and Ø-RC are shown below (I assume that internal arguments transit through SpecvP, i.e. the left edge of the v phase domain, on their way to the C domain, as shown).

(20) The man [that John saw] was tall.
[$_{ForceP}$ [$_{DP}$ man]$_i$ [$_{Force'}$ Force=that [$_{FinP}$ Fin=Ø [$_{TP}$ John [$_{T'}$ T [$_{vP}$ t$_i$ saw t$_i$]]]]]]

(21) The man [Ø John saw] was tall.
[$_{FinP}$ [$_{DP}$ man]$_i$ [$_{Fin'}$ Fin=Ø [$_{TP}$ John [$_{T'}$ T [$_{vP}$ t$_i$ saw t$_i$]]]]]

Note that SpecForceP or SpecFinP is a final landing site in these cases. The *wh*-extracted RC head does not move out of the RC, i.e. no clause boundary is crossed.

Before moving on to the *that*-trace effect, a word on *wh*-RCs, i.e. RCs introduced by a *wh*-relative pronoun, is in order. *Wh*-RCs and *that*-RCs pattern alike (and unlike Ø-RCs) in a number of respects: (i) they permit fronted material in the RC (such material follows the *wh*-relative pronoun); (ii) they can extrapose (recall (15b)); and (iii) they can stack (recall (16b)). This suggests that *that*-RCs and *wh*-RCs are both larger than FinP. This correctly predicts that short subject *wh*-RCs are possible.

(22) a. the man who saw me
b. the house which collapsed

However, as is well-known, *wh*-relative pronouns and *that* cannot co-occur.

(23) a. *the man who that saw me
b. *the man who that John saw

(24) a. *the house which that collapsed
b. *the house which that John destroyed

These examples exhibit the so-called Doubly Filled COMP effect (Chomsky & Lasnik 1977). More generally, it seems to be impossible for both the specifier and head

of a projection to be overtly pronounced (for proposals on how to derive this, see, e.g., Koopman 2000, Neeleman & Van de Koot 2006, Starke 2004). This suggests that *wh*-RCs are ForcePs, and that the Force head is null as a result of the Doubly Filled COMP effect. Therefore, it is not the absence of *that* which indicates the absence of a Force projection, but rather the absence of any material that would typically occupy some position within the Force projection.

3.4 The that*-trace effect*

If we assume that movement from SpecTP to SpecFinP is prohibited (as stated in (19)), what does this tell us about (cross-clausal) extraction of a subject?

First, let us consider successful extraction of an embedded subject, which requires *that* to be absent. As argued above, the embedded clause thus lacks a ForceP. There are two possible derivations to consider.

(25) Who did you say Ø saw me?

a. Who$_i$ did you [$_{vP}$ t$_i$ say [$_{FinP}$ Fin=Ø [$_{TP}$ t$_i$ [$_{T'}$ T saw me]]]]?

b. *Who$_i$ did you [$_{vP}$ t$_i$ say [$_{FinP}$ t$_i$ [$_{Fin'}$ Fin=Ø [$_{TP}$ t$_i$ [$_{T'}$ T saw me]]]]]?

In (25a), the subject moves directly from the embedded SpecTP to the matrix SpecvP, SpecvP being the phase edge (Chomsky 2000, 2001, 2004, 2008), without transiting through the edge of the embedded clause. In (25b), movement proceeds via the edge of the embedded clause, i.e. via SpecFinP. Given our hypothesis that movement from SpecTP to SpecFinP is prohibited, as stated in (19), we can rule out (25b). By process of elimination, (25a) is thus the correct derivation for successful extraction of embedded subjects.

We can also assume that, in the absence of *that*, i.e. a ForceP, non-subjects can move straight from the embedded SpecvP to the matrix SpecvP without transiting through SpecFinP, i.e. the edge of the embedded clause.

(26) Who did you say Ø John saw?
Who$_i$ did you [$_{vP}$ t$_i$ say [$_{FinP}$ Fin=Ø [$_{TP}$ John [$_{T'}$ T [$_{vP}$ t$_i$ saw t$_i$]]]]]?

Variants of the proposal that there is no movement via the edge of the embedded clause in the absence of *that* can also be found in work by Ishii (1999, 2004) and Erlewine (2014), though these authors do not assume a split CP.

We now turn to unsuccessful extraction of an embedded subject. In such cases, *that* is present, i.e. there is a Force projection. There are three possible derivations to consider, all of which result in ungrammaticality.

The first is that a subject cannot move directly to matrix SpecvP across both FinP and ForceP. This is stated in (27) and exemplified in (28).

(27) Movement to matrix SpecvP across both FinP and ForceP in one step is prohibited.

(28) *Who did you say that saw me?
*Who$_i$ did you [$_{vP}$ t$_i$ say [$_{ForceP}$ Force=that [$_{FinP}$ Fin=Ø [$_{TP}$ t$_i$ [$_{T'}$ T saw me]]]]]?

The distance between embedded SpecTP and the matrix SpecvP is simply too great if there is a ForceP as well as a FinP in between. (27) presumably also applies to non-subjects.

The second option that is ruled out is the availability of SpecForceP as an intermediate landing site. This is stated in (29) and exemplified in (30).

(29) SpecForceP is not an available intermediate landing site.

(30) *Who did you say that saw me?
*Who$_i$ did you [$_{vP}$ t$_i$ say [$_{ForceP}$ t$_i$ [$_{Force'}$ Force=that [$_{FinP}$ Fin=Ø [$_{TP}$ t$_i$ [$_{T'}$ T saw me]]]]]]?

If SpecForceP were an available intermediate landing site, it should be possible for an extracted subject to move across FinP to SpecForceP, and then from SpecForceP to matrix SpecvP. Recall that there is nothing preventing a subject moving from SpecTP to SpecForceP (this was our derivation for short subject *that*-RCs). (29) also presumably applies to non-subjects.

The postulate in (27) implies that movement via the edge of the embedded clause is required when *that* is present, and the postulate in (29) states that movement does not take place through SpecForceP. The third and final option left to us is that movement proceeds through SpecFinP. However, this means that the subject will be forced to move from SpecTP to SpecFinP. This violates the anti-locality condition in (19), hence this derivation is ruled out.

(31) *Who did you say that saw me?
*Who$_i$ did you [$_{vP}$ t$_i$ say [$_{ForceP}$ Force=that [$_{FinP}$ t$_i$ [$_{Fin'}$ Fin=Ø [$_{TP}$ t$_i$ [$_{T'}$ T saw me]]]]]]?

Since non-subjects move to SpecFinP from a position lower than SpecTP, non-subjects can be extracted from *that*-clauses via SpecFinP without violating anti-locality.

(32) Who did you say that John saw?
Who$_i$ did you [$_{vP}$ t$_i$ say [$_{ForceP}$ Force=that [$_{FinP}$ t$_i$ [$_{Fin'}$ Fin=Ø [$_{TP}$ John [$_{T'}$ T [$_{vP}$ t$_i$ saw t$_i$]]]]]]]?

The generalisation that emerges is stated in (33).

(33) *Wh*-extracted elements only transit through the C domain (more specifically through SpecFinP) when ForceP is present. If ForceP is absent, *wh*-extracted elements move across the C domain without transiting through it.

We can now see that the *that*-trace effect arises when transition through or across a CP-domain is at stake. In contrast, the anti-*that*-trace effect arises when the CP-domain hosts the final landing site of Ā-movement. Furthermore, both effects arise primarily from the ban on movement from SpecTP to SpecFinP. This captures the fact that only extracted subjects yield *that*-trace and anti-*that*-trace effects.

3.5 Summary

Starting with our anti-locality condition, repeated as (34), we derived (35a)-(35b) and (36), as well as the generalisation in (37).

(34) Movement from SpecTP to SpecFinP is prohibited.

(35) a. Movement to matrix SpecvP across both FinP and ForceP in one step is prohibited.
b. Movement across FinP only is permitted.

(36) SpecForceP is not an intermediate landing site.

(37) *Wh*-extracted elements only transit through the C domain (more specifically through SpecFinP) when ForceP is present.

These postulates allow us to derive both the *that*-trace and anti-*that*-trace effects in a unified way. This works technically but a deeper, more principled explanation is desirable. We will attempt to derive these postulates in the next section.

4 Some theoretical implications

In this section we will consider the results of section 3 from the perspective of phase theory. Phase theory is typically associated with skeletal phrase structures of the form C-T-v-V (Chomsky 2000, 2001, 2004, 2008, 2013, Citko 2014, Gallego 2010, 2012, Richards 2007). The cartographic enterprise, on the other hand, deals in much more fine-grained phrase structures. The problem of how phase theory and cartography fit together is a real one (Shlonsky 2010). In standard phase theory, T and V are non-phase heads, C and v are phase heads, and SpecCP and SpecvP are phase edges. However, it is unclear what these correspond to in cartographic terms. This problem arises even if we only assume a single split in the C domain between Force and Fin. What counts as the phase head? What counts as the phase edge? Is the escape hatch part of the phase edge? The analysis developed in section 3 has a number of implications for questions such as these, and thus suggests a way in which phase theory and cartography may 'fit together'.

4.1 The phase head

Our analysis suggests that Force is a phase head whilst Fin is not. In other words, Ø-clauses are not phasal (see also Ishii, 2004).

If Fin is not a phase head, movement across FinP is permitted, as stated in (35b). If Force is a phase head, movement across both FinP and ForceP counts as movement across a phase boundary, which is ruled out, as stated in (35a). Movement across a phase boundary is typically ruled out by the Phase Impenetrability Condition (PIC) (Chomsky 2000: 108):

(38) Phase Impenetrability Condition

> In a phase φ with head H, the domain of H is not accessible to operations outside φ, only H and its edge are accessible to such operations.

According to the PIC, an element within the domain of a phase head H is only accessible to operations outside the phase if that element is in the phase edge. Movement of an element across a phase boundary thus requires that element to transit through the phase edge. In contrast, if there is no phase head present, there will be no movement through the phase edge. This is effectively what is stated in (37). However, our analysis from section 3, particularly the results in (36) and (37), suggests that the typical characterisation of the phase edge and phase escape hatch needs to be modified.

4.2 The phase edge and phase escape hatch

Our analysis suggests that the phase edge and the (phase) escape hatch need to be distinguished, more so than is typically assumed. For a phase head H, the phase edge includes H and all material in HP higher than H. The phase escape hatch is the position through which elements within the complement of H move to escape the phase. In standard phase theory, the escape hatch is part of the phase edge. However, we proposed in section 3 that the escape hatch is SpecFinP, the specifier of the complement of the phase head Force, hence not part of the phase edge at all. Nonetheless, recall that movement through SpecFinP is only forced when the phase head Force is present. The relation between the phase head and the phase escape hatch is thus preserved, albeit in a different configuration.

Why might this configuration obtain? We could derive this by combining Chomsky's (2000) idea that the complement of a phase head is spelled out upon completion of that phase with Fox & Pesetsky's (2005) idea of cyclic linearisation. Cyclic linearisation is also invoked in the *that*-trace analyses of Erlewine (2014) and Abe (2015). The present analysis is closer in spirit and implementation to the former. However, I will show below that the present analysis offers improvements in several respects over Erlewine's analysis.

Suppose the phase head Force is present. Since it is a phase, it will trigger spellout of Force's complement, FinP, upon completion of ForceP. There is some debate about exactly what it means for structure to undergo spellout. If we adopt Fox & Pesetsky's (2005) cyclic linearisation idea, we could say that spellout of structure means that that structure is linearised. In the case at hand, this would mean that the elements within FinP are linearly ordered with respect to each other. Crucially, on their account, these elements are still accessible in principle to further syntactic operations (cf. Chomsky 2000), i.e. elements within FinP can still move out of FinP, but with one important proviso: such operations cannot change the linear ordering relations established between the elements within FinP.

Fox & Pesetsky use this idea to derive escape hatch effects. Following their reasoning, in simplified terms, an element X moving from within FinP must be linearised so as to precede, >, all other elements in FinP, call them A and B. In other words, X must move to the edge of FinP before FinP is linearised, as in (39b). In this way, when X moves out of FinP and later gets linearised with respect to FinP, as in (39c), there will be no ordering contradictions because X precedes FinP and X precedes all elements in FinP (traces do not count for linearisation according to

Fox & Pesetsky).

(39) a. $[_{\text{FinP}}$ A X B]
b. $[_{\text{FinP}}$ X A t_{X} B] X > A > B
c. X …$[_{\text{FinP}}$ t_{X} A t_{X} B] X > FinP; X > A > B

If X did not move to the edge of FinP before linearisation of FinP, X would be ordered as following at least one element, A, within FinP, as in (40b). X then moves out of FinP and later gets linearised with respect to FinP, as in (40c). But now there is an ordering contradiction because X precedes FinP but X follows A, which is inside FinP.

(40) a. $[_{\text{FinP}}$ A X B]
b. $[_{\text{FinP}}$ A X B] A > X > B
c. X …$[_{\text{FinP}}$ A t_{X} B] X > FinP; but A > X > B

This would account for why *wh*-phrases move to SpecFinP, i.e. the left edge of FinP, in the presence of the phase head Force. Force causes its complement FinP to be linearised. If a *wh*-phrase does not move to SpecFinP, it cannot move out of FinP without resulting in an ordering contradiction. The way this account is phrased gives the impression that movement to SpecFinP takes place to avoid some future difficulty. This is a look-ahead problem. However, it would be a mistake to think of avoidance of ordering contradiction causing a *wh*-phrase to move to SpecFinP in the syntax (see Fox & Pesetsky, 2005: 39, on this point).

For this account to work, all elements within FinP (except traces) must be linearised during linearisation, even if some of those elements ultimately end up having no overt morpho-phonological exponence. In other words, a *wh*-phrase must still be linearised so as to precede Fin even though Fin is *Ø* in English. This ensures that *wh*-subjects move from SpecTP to SpecFinP in the presence of Force, thereby resulting in the anti-locality violation underlying the *that*-trace effect. The structure would be as in (41) (showing only the embedded ForceP):

(41) $[_{\text{ForceP}}$ Force $[_{\text{FinP}}$ *wh*-SUBJ Fin $[_{\text{TP}}$ $t_{\textit{wh}\text{-SUBJ}}$ T]]] (anti-local)

That Fin should participate in linearisation even though it is *Ø* in English can be understood in terms of Distributed Morphology (DM) (Halle & Marantz 1993, 1994). According to DM, vocabulary insertion takes place after the syntactic computation has been completed. In other words, the elements manipulated by the syntax have no phonological content, and hence syntax is insensitive to morpho-phonological exponence.

When Force is not present, FinP is not a spellout domain and is thus not linearised until later in the derivation (when it will be part of a larger spellout domain). Consequently, in the absence of Force, there is nothing forcing elements to move to SpecFinP and so *wh*-subjects are not in danger of violating anti-locality.

Erlewine (2014) was the first to combine Spec-to-Spec Anti-locality with cyclic linearisation to account for the *that*-trace effect. However, his implementation differs from mine in several crucial respects. First, Erlewine does not assume a split

CP. In other words, *that* and *Ø* are phonological variants of the same C head (see above for my arguments against this position). Erlewine proposes that the phonological overtness of the C head is crucial in determining whether a *wh*-subject needs to move to SpecCP (the edge of CP) in order to be linearised as preceding all other elements within CP. In other words, movement is sensitive to morpho-phonological exponence (see my critique above). If C is null, i.e. *Ø*, the *wh*-subject can remain in SpecTP when CP is linearised since the *wh*-subject will already be phonologically leftmost in the CP. If C is overt, i.e. *that*, the *wh*-subject must move so as to be phonologically leftmost. However, movement from SpecTP to SpecCP is anti-local according to Spec-to-Spec Anti-locality. This accounts for the *that*-trace effect.

I pointed out some conceptual problems with making syntactic movement sensitive to the morpho-phonological exponence of elements, but there are empirical problems too. First, as mentioned earlier, there are varieties/idiolects of English that do not exhibit the *that*-trace effect (Chomsky & Lasnik 1977, Sobin 1987, 2002). By tying the *that*-trace effect to phonological exponence and linearisation, such systems are predicted not to exist, unless Spec-to-Spec Anti-locality could be suspended in these systems. Second, it is not at all clear how the *that*-trace effect could be reconciled with the anti-*that*-trace effect of short subject RCs. If movement from SpecTP to SpecCP is required to establish the $\bar{A}$-dependency necessary for RCs, this should be ruled out by Spec-to-Spec Anti-locality. We would thus predict short subject RCs to be impossible, contrary to fact. Alternatively, if short subject RCs could be derived without movement from SpecTP to SpecCP, we would expect them to exhibit the *that*-trace effect, also contrary to fact. This highlights the empirical benefits of adopting a split CP. I thus conclude that the present analysis, although very much in the spirit of Erlewine's analysis, offers considerable empirical and conceptual improvements over it.

However, the analysis proposed here does raise conceptual issues of its own. For example, Chomsky's (2000) reason for saying the phase head's complement undergoes spellout was to ensure the presence of a phase edge (=escape hatch) through which elements from within the phase complement could escape spellout. If the escape hatch is in fact lower than the phase head itself, as I have proposed, the conceptual question re-emerges of why the phase head's complement is the spellout (i.e. linearisation) domain. One possibility is that this ensures symmetry breaking, as in the recent proposal by Boeckx (2014). By targeting a subpart of the phase domain (i.e. the phase head's complement, and perhaps the phase head itself as well (see Ott, 2011), spellout serves to break the symmetry of the phase domain resulting from Merge, thereby allowing the interfaces to interpret it. If this is correct, this may provide a (biolinguistically motivated) rationale for the nature of spellout that is independent of considerations of successive cyclicity.

We may also ask the question of why SpecForceP does not constitute an escape hatch. I propose that heads like Force, Topic and Focus trigger criterial freezing effects (den Dikken 2009, Rizzi & Shlonsky 2007, Rizzi 1997, 2006). More generally, it could be the case that the phase edge is dedicated to criterial-type interpretive properties, e.g. clause-type, topichood and focus. Escape hatches, however, are not part of the phase edge and so do not freeze the elements in them. By trig-

gering the appearance of an escape hatch outside of the phase edge, the system permits successive cyclic movement whilst keeping criterial and non-criterial positions maximally distinct. This suggests that heads like Top(ic) and Foc(us) (Rizzi 1997) are also phase heads, and that Fin may be a phase head in Ø-RCs (recall that SpecFinP in such cases is a final landing site). In other cases, however, Fin is not a phase head (see also Totsuka 1978, who proposes that Force, Topic and a dedicated Rel(ative) head are phase heads whilst Fin is not).

The general schema for phases that emerges can be summarised as follows (H is the phase head):

(42) $[_{HP}$ SpecHP H $[_{XP}$ SpecXP X […

a. $[_{HP}$ **SpecHP H** $[_{XP}$ SpecXP X [… (phase edge)
b. $[_{HP}$ SpecHP H $[_{XP}$ **SpecXP** X [… (escape hatch)
c. $[_{HP}$ **SpecHP** H $[_{XP}$ SpecXP X [… (criterial position)

This raises the question of what happens when there are multiple phase heads, e.g. if Force and Topic are both present. We would have the following schematic structure:

(43) $[_{HP1}$ SpecHP1 H1 $[_{HP2}$ SpecHP2 H2 $[_{XP}$ SpecXP X […

According to the analysis being pursued here, H2 would trigger linearisation of XP. SpecXP would be the escape hatch. H1 would trigger linearisation of HP2. However, SpecHP2 is a criterial position. Consequently, no element can move to SpecHP2 as an intermediate landing site. Furthermore, movement from SpecXP would result in ordering contradictions, hence there can be no successive cyclic movement. This then derives topic island effects (see also Lasnik & Saito 1992).

(44) a. You think that, this book, I'll give to John.
b. *Who do you think that, this book, I'll give to?

I conclude that there is successive cyclic movement through the C domain in the presence of a phase head (cf. den Dikken 2009, Rackowski & Richards 2005), which is more or less the standard assumption (Abels, 2012; Van Urk & Richards, 2015), but that successive cyclic movement takes place through an escape hatch position (i.e. SpecFinP) that is outside the phase edge proper.

4.3 Anti-locality: towards an analysis

We now return to the postulate that movement from SpecTP to SpecFinP is prohibited. This allowed us to straightforwardly capture the anti-*that*-trace effect exhibited by English short subject RCs. Since movement to SpecFinP from a position lower than SpecTP is perfectly acceptable (the anti-*that*-trace effect is not exhibited by anything other than short subject RCs), I suggested this prohibition is a manifestation of anti-locality.

This prohibition is consistent with a more general proposal of Erlewine's 2015, namely Spec-to-Spec Anti-Locality (see also Brillman & Hirsch 2015).

(45) Spec-to-Spec Anti-Locality:

Ā-movement of a phrase from the Specifier of XP must cross a maximal projection other than XP.

(46) Definition: crossing

Movement from position α to position β *crosses* γ if and only if γ dominates α but does not dominate β.

However, this proposal is still quite stipulative. First, it differs from other more principled proposals concerning anti-locality (e.g., Abels 2003, 2012, Boeckx 2008, Bošković 1994, Frampton 1990, Grohmann 2000, Jeong 2007, Pancheva 2009, Saito & Murasugi 1999) and is essentially a geometric condition on syntactic operations. It also specifically relates to Ā-movement, but it is not clear why this should be.

An anti-locality proposal similar in spirit to the present one is made by Ishii (1999, 2004). Ishii proposes that, as a phase head, the C head *that* can be assigned an EPP feature. The *wh*-subject in SpecTP is in the minimal domain of C, and so is able to satisfy the EPP feature of C without moving to SpecCP. Consequently, such movement is ruled out by Economy. Since non-subjects are not in the minimal domain of C, they must move to satisfy C's EPP feature. Now, by economically satisfying the EPP feature of C, a *wh*-subject does not move to the edge of CP. Consequently, when the CP phase is spelled out, the *wh*-subject that remains in SpecTP will be inaccessible to further operations. It will thus be unable to undergo successive cyclic movement and uninterpretable features on the phase heads of the matrix clause will go unvalued, causing the derivation to crash. Following Bošković (1997), Ishii (2004) proposes that when *that* is absent, the C layer is absent. Consequently, a *Ø*-clause is not a CP and hence not a phase. The *wh*-subject in SpecTP is thus accessible to operations in the matrix clause.

The present analysis is built on the same basic intuition as Ishii's. However, Ishii's analysis does not easily account for the anti-*that*-trace effect. If the C head *that* can satisfy its EPP feature and any other feature via an Agree relation with the subject in SpecTP, we would expect any movement of the *wh*-subject into the C domain to be uneconomical and hence ruled out. Indeed, this is precisely Ishii's account of the Vacuous Movement Hypothesis. Short subject RCs thus pose a problem because it seems that, in such cases, subjects must move into the C domain. Why would Agree not be enough in short subject RCs? Some authors do suggest that there is no movement in short subject RCs (e.g. Doherty 1993, 2000), but this means that *that* in short subject RCs is not only different from complementiser *that* (which would leave unexplained the similar behaviour with respect to preposed material), but also different from *that* in all other *that*-RCs. Furthermore, Ishii's analysis relies on Agree underlying all movement, even intermediate movement steps. This is arguably not conceptually desirable (see Bošković 2007).

I will sketch an attempt to motivate at least some instances of Spec-to-Spec anti-local movement, though my proposal will remain tentative. In essence, I will adapt recent ideas on labelling from Chomsky (2013, 2014) and suggest that the ban on movement from SpecTP to SpecFinP might be a kind of derived Comp-to-Spec anti-

local movement (Abels 2003, 2012).

Chomsky (2013, 2014), building on and adapting an intuition of Moro's (2000), raises the issue of how to label a structure formed by merging two non-minimal projections {XP YP}. In the case of a subject in SpecTP, the question is how to label {DP TP}. Chomsky proposes that the label comes from the features shared by both DP and TP, namely φ-features, i.e. {φ {DP_{φ} TP_{φ} }}. For Chomsky, labelling occurs at spellout, i.e. there is a separation of the principles of labelling/projection from the principles of structure building. But suppose we adopt Chomsky's idea about labelling without making such a separation, as has been previously assumed (Chomsky 1995). Fin would merge with {φ {DP TP}}. Now, suppose that the subject DP moves to SpecFinP or, in other words, that the subject DP internally merges with Fin. Since the φ label of Fin's complement and the subject's φ-features are identical, it will appear that Fin has merged with the same φ(-bearing) element twice.

(47) $[_{\text{FinP}}\ DP_{\varphi}\ [_{\text{Fin'}}\ \text{Fin}\ [_{\varphi}\ t_{DP\varphi}\ T\ ]]]$

It thus appears that Fin's complement has moved to its specifier, which can be ruled out as a case of 'derived' Comp-to-Spec anti-locality. As Abels (2003, 2012) argues, Comp-to-Spec anti-locality can be reduced to Economy. Merging a phrase twice with the same head is vacuous and hence violates Economy.

For this account to work, labelling must take place in the syntax. {DP TP} must be labelled φ before becoming the complement of Fin, otherwise moving the subject to SpecFinP could allow Fin's complement to be labelled TP in accordance with Chomsky's assumptions. I do not see labelling in the syntax as a problem, and it remains the standard assumption in the literature. Note that projection of shared features cannot result in freezing either, i.e. there is no Subject Criterion effect as a result of labelling {DP TP} as φ.

Thus far we have said nothing about the *wh*-subject's *wh*-feature. If this and the subject's φ features projected such that Fin's complement is labelled φ,wh, as in (48), our suggestion above remains essentially unchanged.

(48) $[_{\text{FinP}}\ DP_{\varphi,\ \mathbf{wh}}\ [_{\text{Fin'}}\ \text{Fin}\ [_{\varphi,\ \mathbf{wh}}\ t_{DP\varphi,\ \mathbf{wh}}\ T\ ]]]$

Fin appears to have merged with the same φ,wh element twice, in violation of Economy. However, this is problematic in a number of respects. First, there is the question of how Fin's complement is semantically interpreted. Clauses labelled as wh have been argued to be interpreted as free relatives (see Donati & Cecchetto 2011, Donati 2006, Ott 2011), but this does not seem to be plausible in the present context. Second, it is not clear that T should have a *wh*-feature, so the *wh*-feature is not shared and hence should not project. Third, if Fin's complement is labelled as φ,wh, we might expect extraction of the *wh*-subject to involve pied-piping of Fin's complement. This does not seem to be well-motivated for English (though clausal pied-piping is used for *wh*-extraction of embedded subjects in other languages, e.g. Imbabura Quechua (Cole & Hermon 1981, Hermon 1984)).

These problems suggest that the *wh*-feature of the *wh*-subject does not project

with φ (at least in English). However, if that is the case, we might expect this *wh*-feature to allow the system to recognise that the element in SpecFinP is distinct from Fin's complement, as in (49).

(49) $[_{\text{FinP}}\ \text{DP}_{\varphi,\ \textbf{wh}}\ [_{\text{Fin'}}\ \text{Fin}\ [_{\varphi,\ \textbf{wh}}\ t_{\text{DP}\varphi,\ \textbf{wh}}\ \text{T}\ ]]]$

In other words, we might expect φ,wh and φ to be treated as distinct, but apparently they are not. One possibility is that, though distinct, they are not 'distinct enough' (φ being a proper subset of φ,wh). This is reminiscent of issues arising in feature-based approaches to Relativized Minimality (Rizzi 2004, 2013, Starke 2001) and in some accounts of transparency effects in phonology (Calabrese 2005, Nevins 2010). Perhaps inter-speaker variation in the computation of (non-)distinctness is responsible for the differences in absolute judgements concerning *that*-trace examples. I leave this suggestion for future research.

4.4 Summary

We have explored several implications of our analysis for phase theory as applied to more cartographic phrase structures. Our account suggests that Force is a phase head and that Fin is not (except for the Fin in Ø-RCs). SpecForceP cannot serve as an intermediate landing site because the phase edge is reserved for criterial-type interpretations. The phase escape hatch is SpecFinP, lower than the phase head. This was derived from Fox & Pesetsky's (2005) cyclic linearisation combined with the idea that a phase head triggers spellout of the phase head's complement. We have thus completely separated the phase edge and the phase escape hatch, thereby providing a clear distinction between intermediate and final landing sites (for independent arguments that such a distinction is desirable, see Richards 2001 and Sheehan & Hinzen 2011). Finally, we tentatively derived the ban on movement from SpecTP to SpecFinP by suggesting that labelling creates a configuration virtually non-distinct from the one resulting in Comp-to-Spec anti-local movement, which is ruled out by Economy.

5 A view from acquisition

In the last section, we discussed the implications of our analysis for the theory of phases and successive cyclicity within a more fine-grained approach to phrase structure. In this section, we will discuss the analysis in broader terms, namely in terms of how this analysis relates to the question of acquisition and the poverty of the stimulus.

The major contribution of this article has been to provide a unified analysis of the *that*-trace and anti-*that*-trace effects. The ability to account for both effects is already a major advantage of this analysis over the many other analyses of the *that*-trace effect in the literature.

I believe this analysis also provides a significant step towards an account that is more plausible from an acquisition perspective. As should be clear, variation in the *that*-trace and anti-*that*-trace effects reduces to lexical variation, i.e. the effects are

what they are because *that* lexicalises Force and *Ø* lexicalises Fin. Appeals to lexical variation have been made in previous accounts of the *that*-trace effect (see Boeckx 2008, Lohndal 2009), but in these proposals Force and Fin can both be lexicalised by *that* in various circumstances. This complicates the learner's task, and raises serious concerns about the type and availability of relevant evidence that would lead a learner to such a conclusion.

Our analysis avoids this complication and so arguably lessens the severity of the concerns relating to the type and availability of relevant evidence. Nevertheless, these concerns still exist and need to be addressed. Evidence suggests that sparseness of the data and poverty of the stimulus are serious problems for the acquisition of *that*-trace effects. As Phillips (2013) notes, the difference between *that* and *Ø* only makes a difference in one specific environment. The vast majority of a child's input would suggest that there is no real difference. Furthermore, based on Pearl & Sprouse's (2013a, 2013b) frequency counts of a corpus of over 11,000 child-directed *wh*-constructions, Phillips (2013: 143-144) notes that overt *that* is extremely rare in *wh*-interrogatives in the input (2 tokens out of 11, 308). Phillips points out that it is not only unclear how a learner would acquire the *that*-trace effect, but also how a learner would not conclude that overt *that* is prohibited in environments where it is deemed grammatical in the adult language.

If we maintain the idea that lexical variation is involved, the acquisition considerations strongly suggest that a learner is able to determine which complementisers lexicalise which C heads without having direct access to *that*-trace data. I believe the present analysis is a step in the right direction. Although exactly what allows a learner to determine that *that* lexicalises Force and *Ø* lexicalises Fin remains unclear, a tentative suggestion could be as follows.

In the present analysis, there is nothing in principle ruling out a situation of 'mirror-English', i.e. in mirror-English, *that* lexicalises Fin and *Ø* lexicalises Force. This system would exhibit a *that*-trace effect in short subject RCs and an anti-*that*-trace effect in long-distance subject *wh*-extraction (exactly the opposite of what we find in English). But suppose there is a learning bias, i.e. a Third Factor in the sense of Chomsky (2005), such that, when an alternation between an overt and a null form is noticed, the overt form is taken to indicate additional structure. The existence of *Ø*-clauses and *that*-clauses would thus be taken to indicate that *that*-clauses had an extra layer of structure relative to *Ø*-clauses. In the terms of Cardinaletti & Starke (1999), we could think of *that*-clauses as 'strong' clauses and *Ø*-clauses as 'weak' clauses, analogous to the distinction between strong and weak pronouns, for instance. This would yield the English system proposed here. Now consider the reverse. If *that*-clauses had less structure than *Ø*-clauses, the learner would have to posit suppression of the overt form as well as additional structure headed by another covert form. Intuitively, this process of abduction seems less motivated.

Another option, of course, is that the learner simply concludes that *that* and *Ø* are phonological variants of the same head (this is a widespread assumption in syntactic analyses). However, there is independent evidence for a learning bias against maintaining free variation in lexical learning, e.g. the mutual exclusivity bias/assumption (Barrett 1995, Clark 1995, Guasti 2002, Markman, Wasow &

Hansen 2003, Markman 1990). Similarly, competition between minimally different grammars and 'doublets' of various sorts is disfavoured in diachrony (Fuß & Trips 2004, Kroch 1994).

To summarise, I believe the present analysis reduces the learner's task during acquisition by its appeal to learning biases. General learning biases during acquisition combined with accessible data involving complementisers leads the learner to conclude that, in English, *that* lexicalises Force and Ø lexicalises Fin. This conclusion, combined with the general mechanisms of phases and successive cyclicity, then automatically yields the *that*-trace and anti-*that*-trace effects.

6 Conclusion

The major contribution of this paper lies in the development of a unified analysis of both the *that*-trace and anti-*that*-trace effects, a long-standing and recalcitrant problem in the literature. The analysis is based on highly general principles, thereby lending it real explanatory depth. This is firmly within the minimalist spirit of seeking explanations at deeper levels of abstraction than those at which the problem is stated.

The *that*-trace and anti-*that*-trace effects arise from anti-locality combined with a split CP analysis of the differences between clauses with and without *that*. This led to a new perspective of the mechanics of phases and successive cyclicity. Our analysis suggests that Force is a phase head whilst Fin is not (except in Ø-RCs), and that the escape hatch of a phase is actually lower than the phase head, which implies that the phase edge and phase escape hatch are completely distinct, contrary to standard assumptions.

Furthermore, the only thing that has to be learned during acquisition is that *that* lexicalises Force, whilst Ø lexicalises Fin. This is especially noteworthy given the poverty of the stimulus. This strongly suggests that the lexicalisation of the different C heads will determine what sorts of effects will arise in different languages. In this respect, it would be fruitful to extend this analysis to the Scandinavian languages, along the lines of Lohndal (2007, 2009). In addition, given that the generalisations were framed in terms of phase theory, it would be desirable to investigate anti-locality effects in other phasal domains.

Jamie Douglas
jad71@cam.ac.uk

References

Abe, J. 2015. The EPP and subject extraction. *Lingua* 159. 1–17.

Abels, K. 2003. *Successive cyclicity, anti-locality, and adposition stranding*. University of Connecticut: Doctoral dissertation dissertation.

Abels, K. 2012. *Phases: An essay on cyclicity in syntax*. Berlin: De Gruyter.

Arsenijević, B. 2009. Clausal complementation as relativization. *Lingua* 119(1). 39–50.

Barrett, M. 1995. Early lexical development. In P. Fletcher & B. MacWhinney (eds.), *The handbook of child language*, 362–392. Blackwell.

Baunaz, L. 2014. On the various sizes of complementizers. *Probus* 27(2). 193–236.

Baunaz, L. & E. Lander. submitted. Syncretisms with the complementizer: From Dem to the nominal core. *Syntax* .

Berwick, R. C. & N. Chomsky. 2011. The biolinguistic program: The current state of its evolution. In A. M. Di Sciullo & C. Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 19–41. Oxford: Oxford University Press.

Bhatt, R. 2002. The raising analysis of relative clauses: Evidence from adjectival modification. *Natural Language Semantics* 10. 43–90.

Bianchi, V. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin/New York: Mouton de Gruyter.

Bianchi, V. 2000a. The raising analysis of relative clauses: A reply to Borsley. *Linguistic Inquiry* 31(1). 123–140.

Bianchi, V. 2000b. Some issues in the syntax of relative determiners. In A. Alexiadou, P. Law, A. Meinunger & C. Wilder (eds.), *The syntax of relative clauses*, 53–81. Amsterdam/Philadelphia: John Benjamins.

Biberauer, T. & I. G. Roberts. 2015. Rethinking formal hierarchies: A proposed unification. *Cambridge Occasional Papers in Linguistics* 7. 1–31.

Boeckx, C. 2008. *Bare syntax*. Oxford: Oxford University Press.

Boeckx, C. 2014. *Elementary syntactic structures: Prospects of a feature-free syntax*. Cambridge: Cambridge University Press.

Bošković, Ž. 1994. Categorial status of null operator relatives and finite declarative complements. *Language Research* 30(2). 387–417.

Bošković, Ž. 1996. Selection and the categorial status of nonfinite complements. *Natural Language and Linguistic Theory* 14(2). 269–304.

Bošković, Ž. 1997. *The syntax of nonfinite complements*. Cambridge, MA: MIT Press.

Bošković, Ž. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38(4). 589–644.

Bresnan, J. W. 1972. *Theory of complementation in English syntax*. Doctoral dissertation: MIT dissertation.

Bresnan, J. W. 1977. Variation in the theory of transformations. In P. W. Culicover, T. Wasow & A. Akmaijan (eds.), *Formal syntax*, 157–196. New York: Academic Press.

Brillman, R. & A. Hirsch. 2015. An anti-locality account of English subject/non-subject asymmetries. Handout, LSA Annual Meeting, Portland, Oregon.

Browning, M. A. 1996. Cp recursion and that-t effects. *Linguistic Inquiry* 27(2). 237–255.

Calabrese, A. 2005. *Markedness and economy in a derivational model of phonology*. Berlin/New York: Mouton de Gruyter.

Cardinaletti, A. & M. Starke. 1999. The typology of structural deficiency: A case study of three classes of pronouns. In H. Van Riemdsijk (ed.), *Clitics in the languages of Europe*, 145–233. Berlin: Mouton de Gruyter.

Chomsky, N. 1977. On wh-movement. In P. W. Culicover, T. Wasow & A. Akmajian

(eds.), *Formal syntax*, 71–132. New York: Academic Press.
Chomsky, N. 1995. *The minimalist program.* Cambridge, MA: MIT Press.
Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by step*, 89–156. Cambridge, MA: MIT Press.
Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: a life in language*, 1–52. Cambridge: MIT Press.
Chomsky, N. 2004. Beyond explanatory adequacy. In A. Belletti (ed.), *The cartography of syntactic structures*, vol. 3, 104–131. Oxford: Oxford University Press.
Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.
Chomsky, N. 2008. On phases. In R. Freidin, C. P. Otero & M. L. Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge: MIT Press.
Chomsky, N. 2013. Problems of projection. *Lingua* 130. 33–49.
Chomsky, N. 2014. Problems of projection: Extensions. Ms., MIT.
Chomsky, N. & H. Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8(3). 425–504.
Cinque, G. 1995. *Italian syntax and universal grammar.* Cambridge: Cambridge University Press.
Citko, B. 2014. *Phase theory: An introduction.* Cambridge: Cambridge University Press.
Clark, E. V. 1995. Later lexical development and word formation. In P. Fletcher & B. MacWhinney (eds.), *The handbook of child language*, 393–412. Oxford: Blackwell.
Cole, P. & G. Hermon. 1981. Subjecthood and islandhood: evidence from quechua. *Linguistic Inquiry* 12(1). 1–30.
Cowart, W. 1997. *Experimental syntax: applying objective methods to sentence judgments.* Thousand Oaks/London/New Delhi: SAGE publications.
De Vries, M. 2002. *The syntax of relativization.* Amsterdam: University of Amsterdam Doctoral thesis.
Deal, A. R. 2015. TP is a phase in relative clauses. Handout, WCCFL 33, Simon Fraser University.
den Dikken, M. 2009. On the nature and distribution of successive cyclicity: Adjunction, resumption and scope marking as the roads to success in long-distance relation building. Ms., CUNY.
Doherty, C. 1993. *Clauses without that: The case for bare sentential complementation in English*: University of California, Santa Cruz Doctoral dissertation.
Doherty, C. 2000. *Clauses without "that": The case for bare sentential complementation in English.* New York/London: Garland.
Donati, C. 2006. On wh-movement. In L. L.-S. Cheng & N. Corver (eds.), *Wh-movement: Moving on*, 21–46. Cambridge, MA: MIT Press.
Donati, C. & C. Cecchetto. 2011. Relabeling heads: A unified account for relativization structures. *Linguistic Inquiry* 42(4). 519–560.
Erlewine, M. Y. 2014. Why the null complementizer is special in the English that-trace effect. Available on Lingbuzz: http://ling.auf.net/lingbuzz/002029.
Erlewine, M. Y. 2015. Anti-locality and optimality in Kaqchikel agent focus. To appear in *Natural Language and Linguistic Theory*. Available on Lingbuzz:

http://ling.auf.net/lingbuzz/001841.
Fox, D. & D. Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31. 1–45.
Frampton, J. 1990. Parasitic gaps and the theory of wh-chains. *Linguistic Inquiry* 21(1). 49–77.
Fuß, E. & C. Trips. 2004. Introduction. In E. Fuß & C. Trips (eds.), *Diachronic clues to synchronic grammar*, 1–29. Amsterdam/Philadelphia: John Benjamins.
Gallego, Á. 2010. *Phase theory*. Amsterdam/Philadelphia: John Benjamins.
Gallego, Á. 2012. *Phases*. Berlin: De Gruyter.
Grimshaw, J. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28(3). 373–422.
Grohmann, K. K. 2000. *Prolific peripheries: A radical view from the left*. College Park: University of Maryland dissertation.
Guasti, M. T. 2002. *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.
Halle, M. & A. Marantz. 1993. Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (eds.), *The view from building 20*, 111–176. Cambridge, MA: MIT Press.
Halle, M. & A. Marantz. 1994. Some key features of distributed morphology. *MIT Working Papers in Linguistics* 21. 275–288.
Heim, I. & A. Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
Hermon, G. 1984. *Syntactic modularity*. Dordrecht: Foris.
Ishii, T. 1999. Cyclic spell-out and the that-t effects. In S. Bird, A. Carnie, J. D. Haugen & P. Norquest (eds.), *WCCFL 18 proceedings*, Somerville, MA: Cascadilla Press.
Ishii, T. 2004. The phase impenetrability condition, the vacuous movement hypothesis and that t-effects. *Lingua* 114(2). 183–215.
Jeong, Y. 2007. *Applicatives: Structure and interpretation from a minimalist perspective*. Amsterdam/Philadelphia: John Benjamins.
Kandybowicz, J. 2009. Embracing edges: syntactic and phono-syntactic edge sensitivity in Nupe. *Natural Language and Linguistic Theory* 27. 305–344.
Kayne, R. S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
Kayne, R. S. 2014. Why isn't "this" a complementizer. In P. Svenonius (ed.), *Functional structure from top to toe: The cartography of syntactic structures*, vol. 9, 188–231. Oxford: Oxford University Press.
Koopman, H. 2000. *The syntax of specifiers and heads: Collected essays of Hilda J. Koopman*. London: Routledge.
Koopman, H. & D. Sportiche. 2014. The que/qui alternation: New analytical directions. In P. Svenonius (ed.), *Functional structure from top to toe: The cartography of syntactic structures*, vol. 9, 46–96. Oxford: Oxford University Press.
Kroch, A. 1994. Morphosyntactic variation. In K. Beals, J. Denton, B. Knippen, L. Meinar, H. Suzuki & E. Zeinfeld (eds.), *Proceedings of the thirtieth annual meeting of the chicago linguistics society*, vol. II, 180–201. Chicago: Chicago Linguistic Society.
Lasnik, H. & M. Saito. 1992. *Move α: Conditions on its application and output*. Cam-

bridge, MA: MIT Press.
Lohndal, T. 2007. That-t in scandinavian and elsewhere: Variation in the position of c. *Working Papers in Scandinavian Syntax* 79. 47–73.
Lohndal, T. 2009. Comp-t effects: Variation in the position and features of c. *Studia Linguistica* 63(2). 204–232.
Maling, J. & A. Zaenen. 1978. The nonuniversality of a surface filter. *Linguistic Inquiry* 9(3). 475–497.
Manzini, M. R. 2014. The Romance k-complementizers. In P. Svenonius (ed.), *Functional structure from top to toe: The cartography of syntactic structures*, 148–187. Oxford: Oxford University Press.
Markman, E. M. 1990. Constraints children place on word meanings. *Cognitive Science* 14. 57–77.
Markman, E. M., J. L. Wasow & M. B. Hansen. 2003. Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology* 47. 241–275.
McCawley, J. D. 1998. *The syntactic phenomena of English*. Chicago: University of Chicago Press.
Moro, A. 2000. *Dynamic antisymmetry*. Cambridge, MA: MIT Press.
Neeleman, A. & H. Van de Koot. 2006. Syntactic haplology. In M. Everaert & H. Van Riemdsijk (eds.), *The blackwell companion to syntax*, vol. IV, 685–710. Oxford: Blackwell.
Nevins, A. 2010. *Locality in vowel harmony*. Cambridge, MA: MIT Press.
Ott, D. 2011. A note on free relative clauses in the theory of phases. *Linguistic Inquiry* 42. 183–192.
Pancheva, R. 2009. More sudents attended FASL than CONSOLE. In W. Browne, A. Cooper, A. Fisher, E. Kesici, N. Predolac & D. Zec (eds.), *Formal approaches to Slavic linguistics 18: The Cornell meeting*, 383–400. Ann Arbor: Michigan Slavic Publications.
Pearl, L. & J. Sprouse. 2013a. Computational models of acquisition for islands. In J. Sprouse & N. Hornstein (eds.), *Experimental syntax and island effects*, 109–131. Cambridge: Cambridge University Press.
Pearl, L. & J. Sprouse. 2013b. Syntactic islands and learning biases: Combining experimental syntax and computational model to investigate the language acquisition problem. *Language Acquisition* 20. 23–68.
Perlmutter, D. M. 1968. *Deep and surface structure constraints in syntax*. Cambridge, MA: MIT Doctoral dissertation.
Perlmutter, D. M. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.
Pesetsky, D. 1982. Complementizer-trace phenomena and the nominative island condition. *The Linguistic Review* I. 297–343.
Pesetsky, D. 2015. Complementizer-trace effects. Ms., MIT. Available on Lingbuzz: http://ling.auf.net/lingbuzz/002385.
Pesetsky, D. & E. Torrego. 2001. T-to-C movement: Causes and consequences. In M. Kenstowicz (ed.), *Ken Hale: a life in language*, 355–426. Cambridge, MA: MIT Press.
Pesetsky, D. & E. Torrego. 2007. The syntax of valuation and the interpretability of

features. In S. Karimi, V. Samiian & W. K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation*, 262–294. Amsterdam: John Benjamins.

Phillips, C. 2013. On the nature of island constraints II: Language learning and innateness. In J. Sprouse & N. Hornstein (eds.), *Experimental syntax and island effects*, 132–157. Cambridge: Cambridge University Press.

Rackowski, A. & N. Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36(4). 565–599.

Richards, M. D. 2007. On feature inheritance: An argument from the Phase Impenetrability Condition. *Linguistic Inquiry* 38(3). 563–572.

Richards, N. 2001. *Movement in language: Interactions and architectures*. Oxford: Oxford University Press.

Rizzi, L. 1982. *Issues in italian syntax*. Dordrecht: Foris.

Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer Academic Publishers.

Rizzi, L. 2004. Locality and left periphery. In A. Belletti (ed.), *Structures and beyond: The cartography of syntactic structures*, vol. 3, 223–251. New York: Oxford University Press.

Rizzi, L. 2006. On the form of chains: Criterial positions and ECP effects. In L. L.-S. Cheng & N. Corver (eds.), *Wh-movement: Moving on*, 97–133. Cambridge, MA: MIT Press.

Rizzi, L. 2013. Locality. *Lingua* 130. 169–186.

Rizzi, L. 2014. Some consequences of criterial freezing: Asymmetries, anti-adjacency and extraction from cleft sentences. In P. Svenonius (ed.), *Functional structure from top to toe: The cartography of syntactic structures*, vol. 9, 19–45. Oxford: Oxford University Press.

Rizzi, L. & U. Shlonsky. 2007. Strategies of subject extraction. In U. Sauerland & H.-M. Gärtner (eds.), *Interfaces + recursion = language?*, 115–160. Berlin/New York: Mouton de Gruyter.

Roussou, A. 2002. C, t, and the subject: That-t phenomena revisited. *Lingua* 112(1). 13–52.

Roussou, A. 2010. Subjects on the edge. In E. P. Panagiotidis (ed.), *The complementizer phase: Subjects and operators*, 76–116. Oxford: Oxford University Press.

Saito, M. & K. Murasugi. 1999. Subject predication within ip and dp. In K. Johnson & I. G. Roberts (eds.), *Beyond principles and parameters: Essays in memory of Osvaldo Jaeggli*, 167–188. Dordrecht/Boston/London: Kluwer Academic Publishers.

Salzmann, M. 2006. *Resumptive prolepsis: A study in indirect A'-dependencies*. Utrecht: LOT.

Sheehan, M. & W. Hinzen. 2011. Moving towards the edge: the grammar of reference. *Linguistic Analysis* 37. 405–458.

Shlonsky, U. 2010. The cartographic enterprise in syntax. *Linguistics Compass* 4(6). 417–429.

Sobin, N. 1987. The variable status of comp-trace phenomena. *Natural Language and Linguistic Theory* 5(1). 33–60.

Sobin, N. 2002. The comp-trace effect, the adverb effect and minimal CP. *Journal of Linguistics* 38(3). 527–560.

Starke, M. 2001. *Move dissolves into merge: A theory of locality.* University of Geneva: Doctoral dissertation Doctoral dissertation.

Starke, M. 2004. On the inexistence of specifiers and the nature of heads. In A. Belletti (ed.), *Structures and beyond: The cartography of syntactic structures*, 252–268. Oxford: Oxford University Press.

Taraldsen. 1978. On the nominative island condition, vacuous application and the that-trace filter. Ms., Indiana University Linguistics Club.

Totsuka, M. 1978. On syntactic distinctness and phase head in English relative clause. 文化 77(3/4). 74–88.

Weisler, S. 1980. The syntax of that-less relatives. *Linguistic Inquiry* 11(3). 624–631.