Complementizer-trace effects*

David Pesetsky
Massachusetts Institute of Technology
pesetsk@mit.edu

14562 words

* For helpful comments on an earlier version of this article, I am grateful to Željko Bošković, Amy Rose Deal, Jeffrey Lidz, Martina Martinović, Colin Phillips, Gary Thoms, and an anonymous reviewer for *Companion to Syntax*.

Abstract

Perlmutter **(**1968; 1971) observed that English *wh*-movement obeys a puzzling constraint: an asymmetry between subject extraction and non-subject extraction that interacts with the complementizer system. While *wh*-extraction of a non-subject from a finite embedded clause is compatible with the presence or absence of the word *that* introducing the clause, extraction of the subject is possible only when *that* is omitted:

a. ✓Who do you think **that** Sue met __? / ✓Who do you think Sue met __?
b. *Who do you think **that** __ met Sue? / ✓Who do you think __ met Sue?

The effect has come to be known as the *that-trace effect*, a member of a family of possibly broader set of phenomena called *complementizer-trace effects*. A strong Poverty of the Stimulus argument makes it clear that these phenomena are rooted in innate properties of human language. The discovery of these effects in multiple diverse and unrelated languages reinforces this conclusion, especially when coupled with the availability of independently supported explanations for their absence in other specific languages. On the other hand, a number of quite distinct accounts have been offered for complementizer-trace effects, and there is not yet a consensus as to which of these approaches is most likely to prove correct. *Linear accounts* have been proposed that specifically bar extraction from positions right-adjacent to elements of the complementizer system, supported by claimed interactions with the presence and absence of prosodic boundaries between the two. Alternative *structure-based accounts* that have proved influential include structural locality requirements on extraction sites occupying particular syntactic positions — as well as proposals taking an opposite approach, banning movement that is too local. Thus, though the importance of complementizer-trace effects for linguistic theory is clear, the deeper source of these effects remains a matter of controversy.

1. Perlmutter's observation

Perlmutter **(**1968; 1971) observed that English *wh*-movement obeys a puzzling constraint: an asymmetry between subject extraction and non-subject extraction that interacts with the complementizer system. While *wh*-extraction of a non-subject from a finite embedded clause is compatible with the presence or absence of the word *that* introducing the clause, extraction of the subject is possible only when *that* is omitted:

(1) ***that*-trace effect**

a. ✓Who do you think **that** Sue met __? / ✓Who do you think Sue met __?
b. *Who do you think **that** __ met Sue? / ✓Who do you think __ met Sue?

The effect has come to be known as the *that-trace effect*[1], one of a broader set of phenomena called *complementizer-trace effects*. Complementizer-trace effects are not limited to *wh*-movement, but are found with the full range of constructions known as Ā-extractions (Bresnan 1977, 178-182):

(2) ***that*-trace effect with all types of Ā-movement**

a.	This is the person who I thought (*that) met Sue.	(relativization)
a.	Mary we think (*that) __ met Sue.	(topicalization)
b.	It is Mary that we think (*that) __ met Sue.	(cleft)
c.	More people like Mahler than we think (*that) ___ like Bruckner.	(comparative)
c.	?Bill will be easy for us to say (*that) __ met Sue.	(*tough*-movement)

The effect has intrigued researchers for many reasons. Its very existence is of course a puzzle, which is one reason why researchers such as Bresnan (1972, 1977) and Chomsky & Lasnik (1977) quickly took up the challenge of exploring and explaining it. A stronger reason for interest concerns its possible *universality*. Chomsky & Lasnik suggested that complementizer-trace effects arise directly from a principle of *Universal Grammar*, as discussed below. Chomsky & Lasnik's proposal has two immediate consequences relevant to all proposals of this sort. First, the effect should be acquired by children even in the absence of relevant input data — that is, there should be a Poverty of the Stimulus argument for its universality. Second, the effect should be observable cross-linguistically, wherever other factors do not intervene.

2. Poverty of the Stimulus

The existence of a poverty of the stimulus argument for the complementizer-trace effect was suggested informally by Chomsky & Lasnik (1977), and has recently been demonstrated more formally by Philips (2013). Phillips' argument is the following: if the effect does not reflect a principle of UG, it must be learnable on the basis of input data. In that case, we would expect that the input data should contain a significant number of instances of non-subject extraction

[1] This nomenclature is rooted in the 1970s idea that movement leaves a "trace" in the position moved from, and reflects Chomsky & Lasnik's (1977) specific proposal about Perlmutter's effect, discussed below.

from embedded clauses introduced by *that*, contrasting with an absence or near-absence of comparable subject extraction from such clauses. In fact, however, in an 11,308-utterance corpus of child-directed speech examined by Pearl & Sprouse (2013), as Phillips notes, extraction of any kind is vanishingly rare from embedded clauses with *that*. This contrasts sharply with extraction from clauses without *that*, which are reasonably numerous:

(3)

	extraction type	**clause introducer**	**# of occurrences in 11,308 utterances**
a.	object	*that*	2
b.	object	ø	159
c.	subject	*that*	0
d.	subject	ø	13

Philips (2013, 144)

Phillips notes further that though the corpora of adult-directed speech analyzed by Pearl and Sprouse corpora contain a slightly greater percentage of object extractions from clauses introduced with *that*, the overall number is still quite low. As Phillips puts it: "even in the most 'helpful' corpus", the adult-directed speech corpus, we can estimate that the crucial object questions with overt *that* occur with sufficient frequency for a child to hear one roughly once every ten days" (p. 144 note 3). The fact that there does not appear to be support for learning this effect in the input is consistent with Chomsky & Lasnik's conjecture that it arises from principles of UG — and thus does not have to be learned from data.

A distinct but closely related issue concerns age of acquisition. If knowledge of the effect is due to UG, we would not be surprised to detect sensitivity to the effect at the very earliest age at which children command clausal embedding. On this point, however, the evidence is mixed. In an elicitation task, Thornton 1990, found that 21 children aged 2;10-5;5 used *that* 18% of the time with subject extraction, as opposed to 25% use of *that* with object extraction (though only two children (both aged 3;9) were consistent in the use of *that* with subject extraction). On a grammaticality judgment task with 32 English-speaking children who ranged from 2;11 to 5;7, McDaniel, Chiu and Maxfield (1995) found a statistically significant but not overwhelming difference between subject vs. object extraction from clauses introduced by *that*, and a correction study by Gathercole (2002) showed a low rate of correction of subject-extraction errors (6%) by second-graders, contrasted with a higher rate of correction (18%) by 35 fifth-graders. We may conclude that there is some, but not overwhelming evidence for early knowledge of the *that*-trace effect seen in (1) and (2).

3. The effect cross-linguistically: Perlmutter's Generalization and Rizzi's emendation

Perlmutter actually began his discussion of the effect in (1) and (2) with French rather than English. French also shows a contrast between subject and non-subject extraction from an embedded clause introduced by *que*, its counterpart to English *that*:

(4) **French** (Perlmutter 1971, 99ff)

a. Qui a-t-il dit que Marie voulait voir __?
who has-he said that Marie wanted to see
'Who did he say that Marie wanted to see?'

b. *Qui a-t-il dit que __ voulait voir Marie?
who has-he said that wanted to.see Marie
'Who did he say wanted to see Marie?'

Since then, effects of a similar sort have been identified in a number of languages, for example:

(5) **Russian** (Pesetsky 1982)[2]

a. %Kogo ty xočeš', čtoby Maša vstretila __?
who.ACC you.NOM want, C.SJN Maša.NOM meet.SJN.F.SG
'Who do you want Masha to meet?'

b. *Kto ty xočeš', čtoby __ vstretil Mašu?
who.NOM you.NOM want, C.SJN meet.SJN.M.SG Maša.ACC
'Who do you want to meet Masha?'

(6) **Wolof** (Martinović 2014)[3]

a. L-an l-a Aali xam ni l-a xale bi gis
CM-Q *l*-CWH Ali know that *l*-CWH child DEF.SG see
'What did Ali know that the child saw'

b. *K-an l-a Aali xam ni l-a __ gis xale bi
CM-Q *l*-CWH Ali know that *l*-CWH see child DEF.SG
'Who did Ali know saw the child '

(7) **Nupe** (Kandybowicz 2006, 220-221)

a. Ke u: bè [ke Musa du __] na o?
what 3.SG seem COMP Musa cook NA O
'What does it seem that Musa cooked?

b. *Zèé u: bè [ke __ du nakàn] na o?
who 3.SG seem COMP cook meat NA O
'Who does it seem cooked meat?'

---

2 Pesetsky (1982, 298-299) observed that many speakers judge the object extraction in (5a) as ungrammatical (Comrie 1973; Zaliznjak & Padučeva 1979), but still detect a contrast in acceptability with the subject extraction in (5b), hinting at the possibility of a particularly strong Poverty of the Stimulus argument concerning this effect, if neither extraction occurs in normal speech.

3 These examples are quoted from an earlier version of the paper, but have been verified with the author.

(8) **Levantine Arabic** (Kenstowicz 1983; 1989)

a. ʔayy fusṭaan [Fariid kaal (innu) l-bint ištarat _]
which dress Fariid said that the-girl bought
'Which dress did Fariid say that the girl bought?'

b. ʔayy bint Fariid kaal [(*innu) __ ištarat l-fusṭaan]
which girl Fariid said that bought the dress
'Which girl did Fariid say bought the dress?'

The discovery of the same contrast in multiple, often unrelated languages supports the idea that the effect is rooted in principles of UG. At the same time, Perlmutter observed that comparable effects are not as straightforwardly observable in all languages. At first glance, for example, the effect appears to be absent in Italian, as illustrated by the lack of contrast between the object extraction in (9a) and the subject extraction in (9b). The significance of the question mark after the indicated gap in (9b) will be clear shortly:

(9) **Italian**

a. Chi pensi che i linguisti hanno incontrato __?
who you.think C the linguists AUX.3pl met

b. Chi pensi che __ [?] ha incontrato i linguisti?
who you.think C has met the linguists

Perlmutter's own discussion concerned Spanish and Serbo-Croatian, which behave like Italian in this respect. He advanced the conjecture that the absence of the effect was connected to the fact that these languages allow *subject pro-drop* (i.e. unpronounced pronominal subjects), while languages like English and French do not. Perlmutter suggested that English and French disallow null subjects because they obey the constraint in (10), which he suggested was simply inactive in subject *pro*-drop languages:

(10) **Perlmutter's constraint**
Any sentence other than an imperative in which there is an S that does not contain a subject in surface structure is ungrammatical.

Perlmutter assumed that both subject *pro*-drop and subject extraction from *that*-clauses involve structures that violate (10) by virtue of lacking a structural subject position at surface structure. (He presupposed that nothing remains in the position of extraction by the time the filter applies.) When not only the subject but also *that* is missing in English, Perlmutter proposed a node-pruning operation that deletes the S node, leaving a bare verb-headed structure as the sole clausal node.[4] This structure by-passes (10) because of the absence of the S-node explicitly referred to in the statement of the filter. In languages such as Italian, where the filter is turned off entirely, not only *pro*-drop is allowed, but also subject extraction in the presence of an overt complementizer.

---

[4] The node labelled S (for "sentence") roughly corresponds to TP or IP in more recent work.

Perlmutter's proposal not only provided an account of cross-linguistic variation with respect to complementizer-trace phenomena, but also provided a model for future ideas about how superficial properties of languages may co-vary as a result of parametric variation in the applicability of deeper UG principles. The generalization about variation that underlies his proposal came to be known as *Perlmutter's generalization*.

Later research suggested, however, that Perlmutter's generalization was not quite correct as stated, and should be replaced by a distinct, but closely related generalization, explored first by Rizzi (1982). Rizzi pointed out that SVO languages like Italian that appear to lack the effect in (1)-(2) share another property besides pro-drop. Under specific discourse conditions that depend on the nature of the verb, these languages (in contrast to languages like English and French) allow the nominative subject to appear post-verbally (yielding VS or VOS) as well as preverbally (yielding SV or SVO):

(11)a. Alcune pietre sono cadute. (preverbal subject)
some stones are fallen
'Some stones have fallen.'

b. Sono cadute alcune pietre. (postverbal subject)

If Perlmutter's effect concerns extraction of *preverbal* subjects only, languages like Italian, Spanish and Serbo-Croatian (but not English or French) might bypass the effect by using the postverbal rather than preverbal position for subject extraction. The connection with subject pro-drop would then be indirect, insofar as the same factor that allows subject pro-drop might be crucial to the possibility of postverbal subjects.

Rizzi discovered direct evidence that it is postverbal position, rather than pro-drop, that allows apparent exceptions the complementizer-trace effect in languages like Italian. To show this, he used as his experimental probe a phenomenon that can independently be shown to distinguish certain preverbal from postverbal subjects in Italian. In preverbal subject position, when the restrictor of a quantifier like *alcune* 'some' is understood pronominally, it may be null, much as in English — but crucially its pronominal component cannot expressed with the genitive clitic *ne*, as (12a) shows. In the (postverbal) direct object position of a transitive verb, however, the facts are reversed, as (12b) shows: the restrictor may not be null, and *must* be expressed with the genitive clitic *ne*. The crucial example is (12c), which shows that a postverbal subject with an unaccusative verb like *cadere* 'fall' behaves like a direct object, requiring the clitic — and unlike the preverbal subject in (12a):

(12)a. ✓Alcune sono cadute in mare. / *Alcune **ne** sono cadute in mare.
some are fallen in sea / CL.GEN
'Some (of them) fell into the sea.'

b. *Mario ha prese alcune. / ✓Mario **ne** ha prese alcune.
Mario has taken some / CL.GEN
'Mario took some (of them).'

c. *Sono cadute alcune in mare. / ✓**Ne** sono cadute alcune in mare.
are fallen some in sea / CL.GEN
'Some (of them) fell into the sea.'

As Rizzi noted, if *alcune* is replaced by a *wh*-quantifier, the paradigm in (12) provides a test for the position from which *wh*-movement takes place. When *wh*-movement takes place from object position, *ne* is obligatory as expected, as (13a) shows. Crucially, when a subject is extracted, *ne* is also obligatory, as seen in (13b), As this finding parallels (12c) rather than (12a), subject *wh*-movement clearly proceeds from a postverbal rather than preverbal position:

(13)a. *Quante hai detto che hai preso __? / ✓Quante hai detto che **ne** hai preso _?
how.many you.have said C you.have taken / CL.GEN
'How many (of them) did you say that you took?'

b. *Quante hai detto che __ sono cadute? / ✓Quante hai detto che **ne** sono cadute __?
how.many you.have said C are fallen
'How many (of them) did you say fell?'

For reasons specific to the syntax of *ne*, the contrast in (13) can only be demonstrated when the embedded verb is unaccusative, but Brandi & Cordin (1981, 1989) discovered a comparable test in certain Northern Italian dialects that extends to unergative as well as unaccusative verbs. In these dialects, preverbal subjects are obligatorily doubled by a pronominal clitic, while postverbal subjects disallow this doubling entirely. In addition, the finite verb shows normal suffixal number agreement with preverbal subjects, but not with postverbal subjects, masked by morphology in some verb forms. The following examples are from the Trentino dialect:

(14) **Trentino** (Brandi & Cordin 1981, 1989)

a. ✓La Maria **la** parla. / *La Maria parla.
the Maria she speak.3
'Maria speaks.'

b. ✓Le putele **le** parla. / *Le putele parla.
the girls they speak.3

b. ***L'** ha parlà tre putele. / ✓Ha parlà tre putele.
they have.3.SG spoken three girls
'Some girls have spoken.'

Long-distance *wh*-movement of an embedded subject shows the cliticless pattern of postverbal subjects, strongly supporting Rizzi's emendation of Perlmutter's Generalization:

(15) *Quante putele te pensi che __ **le** abia parlà? / ✓Quante putele te pensi che abia parlà _?
how.many girls you think C they have spoken.M.SG

Kenstowicz (1983; 1989) presents a very similar argument from Bani-Hassan Arabic, a language that, unlike Levantine Arabic ((8)), superficially fails to show the complementizer-trace effect. Bani-Hassan also contrasts with Levantine Arabic in permitting postverbal subjects. Conveniently for the linguist, in Bani-Hassan, the word 'who', when it fails to undergo overt *wh*-movement (e.g. in a multiple question) shows different allomorphs when preverbal or postverbal. Crucially, the postverbal allomorph is the only possibility when 'who' undergoes long-distance *wh*-movement from subject position, strongly supporting the hypothesis that the complementizer-trace effect is active after all, and that subject extraction must proceed from the postverbal position, exactly as in the Italian dialects disucssed above.

4. Range of the effect

If we have indeed determined that one or more universal principles conspire to prohibit certain kinds of long-distance extraction of preverbal subjects cross-linguistically, the obvious next task is the formulation of specific hypotheses about nature of these principles. Since it is highly unlikely (though not logically impossible) that UG includes a principle whose sole effect is to exclude preverbal subject extraction from *that*-clauses, considerable effort has been spent exploring a range of "sister effects" that might be attributable to the same principles that exclude examples like (1b) — in the hope that the principles that explain the full range of these effects might turn out to be deep and general. The search for possible "sister effects" is thus both a preliminary and a concomitant to the discussion of specific proposals about the effect. To a great extent, this work continues to this day, since a number of strikingly different possible explanations for the complementizer-trace effect remain under active consideration by the field. We return to this topic in section 5.

One question concerns the range of elements that can play the same role as English *that*, blocking extraction of a preverbal subject. In English, Ross (1967, 445ff.) observed that the complementizer *for* blocks extraction of the subject, an observation that Bresnan (1977, 171) suggested should be unified with comparable effects triggered by *that*:

(16) ***for*-trace effects**
a. Who would you prefer (for) Sue to meet __ at the station?
b. Who would you prefer (*for) __ to meet Sue at the station?

While the unification of *that*-trace effects with *for*-trace effects might seem like an obvious step, Chomsky & Lasnik (1977, 455ff) argued against this unification, claiming that certain North American dialects ("Ozark English") show the former but not the latter (though the empirical basis for this argument remains shaky, for lack of systematic fieldwork). It is probably fair to say that most current work assumes that the two effects have the same cause after all.

Kayne (1979) noted that prepositions with accusative-subject gerund complements ("ACC-ING gerunds") behave much like *that* and *for* in blocking subject extraction, as seen in the contrast between (17b) and (17d) — though the effect is weaker than the *that*-trace effect. Example (17a) and (17c) show that when the gerund is the complement of a verb, no such contrast is found.

(17) "**P-trace effect"** (modeled on examples from Kayne 1979**)**
a. How much headway did he anticipate [Mary making __ on the issue]?
b. How much headway did he talk [about Mary making __ on the issue]?
c. How much headway did he anticipate [__ being made on the issue]?
d. ??How much headway did he talk [about __being made on the issue]?

While contrasts like those in (16) and (17) might show that the range of complementizers triggering the effect extends beyond the finite declarative complementizer in English, in some languages there are words that look like complementizers that actually cause the effect to disappear. As Perlmutter (1971, 102 footnote 2) discovered, for example, French examples like (4b) (repeated as (18a) below) improve if the embedded clause is introduced by *qui* rather than by *que*. The use of this kind of *qui* is limited to clauses from which the subject has been extracted, and is not fully acceptable for all speakers (though the contrast with *que* is usually clear). The celebrated "*que/qui* alternation" discovered by Perlmutter is generally considered one of the core facts that a theory of the effect should explain:

(18) **French *que/qui*** alternation
a. *Qui a-t-il dit **que** __ voulait voir Marie?
who has-he said **QUE** wanted to.see Marie
'Who did he say wanted to see Marie?'

b. Qui a-t-il dit **qui** __ voulait voir Martin?
who has-he said **QUI** wanted to.see Martin
'Who did he say wanted to see Martin?'

A similar alternation between *da* and *die* in West Flemish, observed by Bennis & Haegeman (1984, 35), is almost as celebrated, with *die* (like French *qui*) limited to clauses from which the subject has been extracted. One difference between the French and West Flemish paradigms is the fact that *da* can also be used with subject extraction, i.e. there is no straightforward example excluded as an instance of the *that*-trace effect:

(19) **West Flemish *da/die* alternation**
a. den vent da Pol peinst da/*die Marie __ getrokken heet
the man that Pol thinks DA/DIE Marie made.a.picture.of has
'the man that Pol thinks that Marie has made a picture of'

b. den vent da Pol peinst da/die __ gekommen is
the man that Pol thinks that come is
'the man that Pol thinks has come'
Bennis & Haegeman (1984, 35)

Contrasts that resemble complementizer-trace effects are also found in English and other languages when a preverbal subject is extracted from an embedded interrogative. Extraction from any position within an embedded interrogative, objects included, is often judged unacceptable, since embedded interrogatives are islands, but subject extraction is far less acceptable than extraction of non-subject arguments, despite the absence of a complementizer introducing the embedded interrogative in languages like English:

(20) **"wh-trace effect" (subject extraction blocked from embedded interrogative)**
   a. [??]Remind me which person you were asking whether Sue had invited __.
   b. *Remind me which person you were asking whether __ had invited Sue.

Other possible sister phenomena will be discussed in later sections, in the context of particular accounts of the effect.

5. Accounts of the effect

Even 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 languages commonly discussed in connection with this effect, the offending extraction site linearly follows the complementizer, and occupies a nearby specifier position that is structurally immediately below it. The set of possible proposals thus bifurcates into two general families of hypotheses: (1) those that start with the hunch that *linear order* is a crucial component of the correct account, and (2) those that start with the hunch that the key factor is *phrase-structural position*. In other words, either (1) there is something special about Ā-movement from a position immediately to the right of the complementizer, or else (2) there is something special about Ā-movement from a position structurally right below it (or alternatively both factors could be relevant).

Theories of both types have been proposed ever since Perlmutter discovered the effect, differing in empirical coverage and theoretical consequences. As a consequence, especially in light of the Poverty of the Stimulus argument that the effect is rooted in UG (section 2), the correct account of complementizer-trace effects has become something of a "Hilbert Problem" for the field. In this section, we will discuss linear accounts first, and then turn to structural accounts (which themselves bifurcate into two general families, as we shall see).

5.1 Linear accounts

Two of the earliest influential accounts of the phenomenon suggested that linear adjacency between the complementizer and the extraction site held the key to the effect. In her dissertation, Bresnan (1972, 95 ff.) proposed a condition on variables in the spirit of Ross (1967) that she called the *Fixed Subject Condition* (FSC):

(21) **Fixed Subject Condition (**Bresnan 1972; 1977)
No NP can be crossed over [i.e. moved so as to cross] an adjacent complementizer.[5]

Bresnan formulated the FSC as a constraint on movement. Building on the then-new idea that a moved element retains a presence in its former position in the form of a phonologically silent *trace*, Chomsky & Lasnik (1977) advanced a similar proposal that was crucially not a constraint on movement, but a constraint on the output of movement. Rather than directly block movement from a post-complementizer position, their constraint applied at surface structure, stigmatizing the trace left by movement. Their proposal was thus a surface filter, in the sense pioneered by Perlmutter **(**1968; 1971), which came to be popularly known as the *that-trace filter* — though it actually included embedded interrogative examples like (20) in its purview in addition to *that*-clauses. Omitting an qualification that dealt with relative clauses (see section 6.2), their filter is reproduced in (22):[6]

(22) ***That-trace* filter (**Chomsky & Lasnik 1977)
*[$_{\bar{S}}$ ±WH [$_{NP}$ *e*] ...]

As Chomsky & Lasnik noted, Perlmutter's generalization that subject *pro*-drop languages fail to show complementizer-trace effects is predicted by their filter if subject *pro*-drop arises from a deletion rule that applies to traces as well as pronouns.[7] Applying to a subject trace, this deletion rule bleeds the *that*-trace filter. Though the filter in (22) was largely abandoned by the time Rizzi (1982) argued that it was postverbal position rather than pro-drop that allows languages like Italian to escape the effect, Chomsky & Lasnik's theory could have accommodated Rizzi's emendation as well — so long as no trace-like element was posited in a preverbal subject position when the full subject is post-verbal.

If the key configuration for the complementizer-trace effect involves linear order, i.e. an extraction site immediately following a complementizer, then we expect the effect to disappear if some other element intervenes linearly between the extraction site and the complementizer — even if the structural position of the extraction site and the complementizer remains the same. Bresnan (1977, 194 footnote 6) confirmed this prediction, noting that the placement of an adverbial expression immediately after the complementizer ameliorates the effect noticeably, an observation rediscovered by Culicover (1993a, 1993b), from whom (23a-b) are taken:

---

[5] The Fixed Subject Condition is ultimately reformulated in a more complex manner for reasons mostly irrelevant to the current discussion. See Bresnan (1972, 305 ff. esp. 308) and Bresnan (1977, 174), where it is renamed the *Complementizer Constraint on Variables*, for discussion.

[6] The notation $\bar{S}$ corresponds to modern CP, ±WH corresponds to complementizer *that* and interrogative C, and "[$_{NP}$ *e*]" indicates a trace left by movement. As noted above, Chomsky & Lasnik did not extend their filter to the *for*-trace effect for empirical reasons, but they did explicitly consider the possibility, and could have easily revised the filter to accommodate it.

[7] Bresnan's account could not be directly extended in this way, but she suggested an acquisition scenario that could capture the generalization.

(23) **Adverb intervention effect (English)**

a. Robin met the man who Leslie said that for all intents and purposes __ was the mayor of the city.

b. I asked what Leslie said that in her opinion __ had made Robin give a book to Lee.

Kandybowicz (2006) later observed very similar effects in Nupe. Omission of the boldfaced adverbials eliminates the effect (cf. (7) above):

(24) **Adverb intervention effect (Nupe)**

a. Zèé Musa gàn [gànán **pányi lèé** __ nì enyà] o?
who Musa say C before PST beat drum O
'Who did Musa say that a long time ago beat the drum?'

b. Bagi [na Musa kpe [gànán **pányi lèé** __ ba nakàn]] na
man C Musa know C before PST cut meat NA
'the man that Musa knew cut the meat a long time ago'

If linear adjacency is crucial to the effect, it might also be predicted (depending on the architecture of one's grammatical theory) that the placement of a prosodic boundary between complementizer and extraction should eliminate their adjacency for the purposes of the effect. De Chene (2000) argued that this prediction is confirmed when Right Node Raising of the sister of C induces a prosodic boundary immediately to the left of the shared TP continuation. In his judgment, the complementizer-trace effect disappears, in this situation — though some speakers do not detect the contrast. Example (25a) shows the claimed amelioration, and (25b) shows that no amelioration occurs when the prosodic boundary precedes, rather than follows, the complementizer:

(25) **Prosodic boundary ameliorating the complementizer-trace effect**
(data and judgments from De Chene 2000)

a. That's the meeting I've been wondering if, and Jim's been saying that, __ is going to be canceled.

b. *That's the meeting I've been thinking, and Jim's been saying, that __ is going to be canceled.

In response to observations like (23) and the more controversial (25), de Chene (2000) and Kandybowicz (2006), among others, have suggested that the complementizer-trace effect is not just linear, but makes crucial reference to *prosodic* boundaries. To support this conclusion, Kandybowicz reports that the effect is ameliorated by manipulation of these boundaries — for example, by reducing or destressing the complementizer.[8] Once again, English speakers differ in their degree of assent to this contrast (as in the judgment study of Ritchart et al. 2015):

[8] Kandybowicz also suggests that a contrast in Nupe makes the same point. Specificially, while *gànán*, glossed as a complementizer, triggers the effect, as seen in (7), a reduced version *'án* does not. See Kandybowicz (2006, 224 footnote 3) for more discussion.

(26) **Phonological reduction of C claimed to ameliorate the effect**
(data and judgments from Kandybowicz 2006, 222)
a. Who do you hope *for/✓?fer __ to win?
b. The author that the editor predicts *that/✓?th't __ will be adored

Kandybowicz proposed a constraint quite similar to Chomsky & Lasnik's filter in (22), except that it makes reference to prosodic boundaries (and happens to encompass *for*-trace effects):

(27) **Prosodic filter (**Kandybowicz 2006, 223)
*<C°, t> iff:
(i) C° & t are adjacent within a prosodic phrase, and
(ii) C° is aligned with a prosodic phrase boundary

In (25a), the prosodic boundary orthographically marked by de Chene's comma before the extraction site allows it to escape the filter by not meeting condition (i). Phonological reduction of the complementizer in (26) eliminates an otherwise present prosodic phrase boundary to the left of C, according to Kandybowicz, bleeding condition (ii) of the filter.).

The three proposals considered in this section (Bresnan's Fixed Subject Condition, Chomsky & Lasnik's *that*-trace filter and Kandybowicz's prosodic filter) all share the assumption that what triggers the effect is linear adjacency between the complementizer and the extraction site. Given that the extraction site of *wh*-movement is phonologically null in languages like English, one can also imagine a linear account that stigmatizes linear adjacency not between the complementizer and the extraction site per se — but rather between the complementizer and what *follows* the extraction side, for example the finite verb. In an experimental study conducted by Salzmann et. al. (2013), German-speaking subjects performing a magnitude estimation task were asked to judge German sentences that contain the sequence C-V for reasons other than subject extraction (rightward verb projection raising that includes the subject) — as well as traditional subject-extraction examples of the complementizer-trace effect. Subjects judged both structures as degraded compared to a variety of reference sentences. Salzmann et al. took this result as support for an account of the complementizer-trace effect that bars the sequence C-V, in contrast to accounts that focus on extraction as a crucial factor.[9]

Finally it is worth noting that though the various linear proposals discussed in this section were formulated so as to apply to somewhat different sets of phenomena, they can all be adjusted so as to include or exclude particular elements from the roster of "complementizers" that trigger the filter. For example, Chomsky & Lasnik could have easily extended their filter to include *for*-trace effects, had they wished. Likewise, as they themselves noted, the amelioration of the effect in French when *que* is replaced by *qui* can be explained if *qui* "differs in at least one feature" from *que* (p. 452) so that it does not satisfy the conditions of the filter. Similar variations on the proposals by Bresnan, Kandybowicz or Salzmann et al. can easily be envisaged that include or exclude various clause-initial elements from their purview. The disappearance of the effect

[9] The account favored by Salzmann et al. is not a straightforward "C-V filter", but has a strong structural component: a condition requiring overt material in some specifier position between V and C.

when overt *that* is absent in English could also be attributed to a variety of factors: for example, a C-deletion rule that removes the complementizer from the purview of the relevant constraint, or a feature matrix for null C with similar effect. Such flexibility is not necessarily a virtue, however, since it reveals the degree to which all these accounts fall short of *predicting* the scope of the effect, as opposed to merely stipulating its existence.

5.2 Structure-based accounts

The linear proposals that have been advanced as explanations for the complementizer-trace effect have a related disadvantage. The are all *sui generis*, essentially singling out the string that corresponds to judgments of unacceptability and prefixing it with an asterisk. This fact might not be a deep demerit of these proposals, but might be an accident of our ignorance about related phenomena. Since we know far less about prosody-syntax interactions, for example, than about other aspects of syntactic structure, it could easily turn out that a filter such as Kandybowicz's might be a special case of a more general class of phenomena that have not yet been discovered and explored.

On the other hand, there is a different family of prominent proposals that are more directly *structural,* situating the explanation for complementizer-trace effects in the syntax proper, exploring possible connections and extensions to other syntactic phenomena. Some of these proposals are not *sui generis*, but attempt to draw links between complementizer-trace effects and other syntactic phenomena, which is probably why they have attracted the greatest interest from the beginning (including Perlmutter's original proposal in (10), discussed in section 3) At the same time, as we shall see, each of these proposals stumbles at some point in accounting for the range of phenomena thought to be central to the discussion. This is one reason why there is still no consensus concerning the right approach.

Most of the structure-based accounts build on the independent observation that successful Ā-movement appears to proceed successive-cyclically through the edges of domains such as CP. Complementizer-trace effects are attributed by these proposals to some consequence of movement from the subject position to a position in the complementizer system that is affected by the presence, absence or featural content of the complementizer that such movement crosses. The most prominent accounts from the late 1970s through the end of the twentieth century attributed phenomena such as the unacceptability of subject extraction over *that* to locality principles requiring structural closeness between the subject and its CP-peripheral intermediate landing site. More recently, a variety of different proposals have been advanced, including several discussed below, that rely (in one way or another) on the requirement of successive-cyclicity. The following subsections survey some of the more prominent or promising approaches.

5.2.1 Nominative Island Condition accounts

In the late 1970s, several of the earliest structure-based accounts suggested a strong link between complementizer-trace effects and the ban on nominative reflexives seen in (28), observable in many (though not all) languages. Chomsky (1980, 26) had accounted for this ban with the constraint in (29): where for our purposes we can understand the term "anaphor" as singling out reflexives:

(28) *Mary said that herself saw Bill.

(29) **Nominative Island Condition (NIC)**
A nominative anaphor cannot be free in $\bar{S}$ [= modern CP].

Chomsky himself speculated (1980, 14) that the *wh*-trace effect in (20a-b) might follow from (29). In an embedded question, Chomsky reasoned, the sole position that might have been available for overt successive-cyclic extraction through the edge of the embedded clause is filled by a distinct element. If a nominative trace in subject position has the status of an anaphor for (29), it will inevitably violate the NIC when the landing site for successive-cyclic movement within its own clause is unavailable.

Chomsky's proposal immediately suggested to other researchers the possibility that the *that*-trace effect might also be a consequence of the NIC. The main obstacle to this unification concerned the role of the complementizer in regulating the effect. Since successive-cyclic movement of a non-subject is possible through the periphery of an embedded *that*-clause (in contrast to embedded interrogatives), it was not easy to see why overt *that* in English or *que* in French should block binding of a nominative trace when a subject is extracted. A series of papers written in the immediate wake of Chomsky's (1980) suggestion by Kayne (1980), Taraldsen (1978) and Pesetsky (1979,1982) advanced several different suggestions for overcoming this obstacle.

Kayne (1980) proposed that the binder for a nominative trace must itself be assigned case, and suggested that the presence of *that* in C blocks case-assignment to a trace of successive-cyclic movement in the C-domain. In support of this claim, Kayne noted that extraction of a nominative element is degraded even when *that* is absent, when the clause from which extraction takes place occupies a non-case position such as the complement position of an adjective. Though Kayne's judgments, reproduced in (30), probably overstate the contrast (and (30b) becomes even less acceptable when *that* is added to the embedded clause), the contrast seems real, and remains an interesting puzzle for all accounts:

(30) ***that*-less subject/non-subject contrast in a non-case environment** (Kayne 1980, 77)
a. The only person who it's not essential she talk to __ is Bill.
b. *The only person who it's not essential __ talk to her is Bill.
(judgments from Kayne 1980)

On the other hand, even granting Kayne's idea that the intermediate trace must itself be case-marked, it remained unclear why the presence of *that* should block the required case-marking. (Kayne advanced a rather complex stipulation on this point.) Consequently, Kayne's proposal was not much developed in subsequent work:

Taraldsen (1978) proposed an interaction between linearization and the NIC to explain the interaction of *that*-trace effects with the overtness of *that* (thus straddling the divide between linear and structural accounts). His proposal, however, crucially presupposed a structure for the

complementizer system that differs from accounts supported by subsequent research, so his proposal will not be discussed in detail here.

Inspired by Taraldsen's paper, Pesetsky (1979, 1982) proposed an account of the interaction of the NIC with English *that* that had the virtue of relying on an independently motivated property of the complementizer system: the (still poorly understood) restriction against the simultaneous occurrence of an overt *wh*-phrase and an overt element in C commonly dubbed the *Doubly Filled Comp Filter* (Keyser 1975; Chomsky & Lasnik 1977). This filter interacts with the ability of certain *wh*-expressions and C-elements to undergo free deletion, successfully predicting the existence of three — but crucially not four — varieties of finite relative clauses in English. If neither the *wh*-phrase nor *that* is null, the Doubly Filled comp Filter is violated:

(31)a.	[The person who ~~that~~ Mary met] is my friend.	*that* deleted
b.	[The person ~~who~~ that Mary met] is my friend.	*who* deleted
c.	[The person ~~who~~ ~~that~~ Mary met] is my friend	*who* and *that* deleted
d.	*[The person who that Mary met] is my friend.	violates Doubly Filled Comp Filter

Pesetsky proposed that a trace of intermediate Ā-movement might also be forced to delete in order to avoid violating the Double Filled Comp Filter, when it cooccurs with an undeleted *that*. Deletion of an intermediate trace that is crucial to the satisfaction of the NIC, he argued, might explain the *that*-trace effect, as seen in (32). Crucially, if *that* is not deleted, either the intermediate trace is deleted, violating the NIC in (32b), or the Doubly Filled Comp Filter is violated, as in (32d):

(32) **Pesetsky's (1979, 1982) proposal**

a.	Who$_i$ do you think [t_i ~~that~~ —$_i$ met Sue]?	*that* deleted, NIC satisfied
b.	*Who do you think [~~t_i~~ that —$_i$ met Sue]?	*t* deleted, NIC violated
c.	*Who do you think [~~t_i~~ ~~that~~ —$_i$ met Sue]?	*t* and *that* deleted, NIC violated
d.	*Who do you think [t_i that —$_i$ met Sue]?	violates Doubly Filled Comp Filter

Pesetsky extended this proposal to the French *que/qui* alternation by suggesting that *qui* is a variant of *que* that acquires an index from a local trace of successive-cyclic movement by a special agreement rule. This index allows the complementizer to save a nominative trace from the effects of the NIC even after the trace of successive-cyclic movement that originally bore it is deleted. (The West Flemish *da/die* alternation was not known at the time, but could have been analyzed similarly.) In the spirit of this analysis of French *qui*, Rizzi (1990, 52-53 and 56) later suggested an alternative to Pesetsky's analysis of the English *that*-trace effect in which the absence of overt *that* arises not from the deletion of the overt complementizer, but from the choice of a null allomorph of *that* with the same kind of index (acquired from an intermediate trace) as French *qui*.

5.2.2 The Empty Category Principle

The general idea that *that*-trace effects might be attributed to the NIC was attractive precisely because it grounded these effects in a broader generalization. Unfortunately, a number of problems were noted that seemed fatal to the idea of unifying the two families of phenomena.

One of the most serious problems concerned the claim that the trace of Ā-movement behaves binding-theoretically like an reflexive, subject in effect to what later became known as Principle A of the Binding Theory (which subsumed the NIC). As Friedin & Lasnik (1981) noted, if anything, such traces behave like full non-pronominal noun phrases (i.e. r-expressions), obeying what later became known as Principle C — in the form of so-called Strong Crossover effects. This is as true of nominative Ā-bound traces as it is of any other Ā-bound trace, as can be seen in (33) — which crucially cannot be understood as "for which *x*, *x* says *x* saw Mary":

(33) ***wh*-trace behaves as an r-expression, not a reflexive**
*Who$_i$ did he$_i$ say [— $_i$ saw Mary]?

Though this objection was generally taken to be devastating for NIC explanations of complementizer-trace phenomena, Aoun (1981, 1983) noted a possible way out. He observed that in examples like (33), the antecedent *he* with respect to which the subject trace behaves like an r-expression occupies an A-position. By contrast, in the *that*-trace configuration, the intermediate trace with respect to which the subject trace behaves like a reflexive anaphor occupies an Ā-position. Aoun proposed that this bifurcation might be real, and that the traces of *wh*-movement might quite generally behave like r-expressions for binders in A-positions, and like reflexive anaphors (obeying NIC) for binders in Ā-positions — embedding this suggestion in a more general theory of binding relations.

The NIC theory of *that*-trace effects faced other problems, however. Though it extended nicely to *wh*-trace effects, as Chomsky (1980) had noted, it did not extend to infinitival *for*-trace effects, which look very much like the same effect in an infinitival context (Chomsky & Lasnik's 1977 skepticism notwithstanding). Furthermore, it was not clear even for finite clauses that the effect singles out nominative expressions, as opposed to whatever happens to occupy preverbal subject position, nominative or not. Bresnan (1977, 186), for example, had observed that fronted locative expressions in the Locative Inversion construction generate *that*-trace effects, despite the fact that that the nominative noun phrase with which the finite verb agrees is located elsewhere, as the plural verb in (33) makes clear:

(34) ***that*-trace effect with Locative Inversion**
In which villages do you believe (*that) __ are found the best examples of this cuisine.
(adapted from Bresnan 1977, 186, ex. 41)

Bresnan's observation strongly suggests that it is movement from a particular position that triggers the effect, not movement of an element bearing a particular case. A similar point arises from Rizzi's and Brandi & Cordin's observations concerning Italian and North Italian preverbal vs. postverbal subjects discussed in section 3. Just as (33) shows a preverbal non-nominative element triggering the effect, the Italian and Trentino contrasts in (13) and (14) show postverbal

nominative elements failing to trigger the effect. Once again, it looks like it is preverbal subject position, rather than the nominative case that is sometimes assigned to this position, that is the true locus of the phenomenon. If the preverbal vs. postverbal distinction is itself structural, corresponding, for example, to a position outside the (extended) verb phrase vs. positions inside it, a more accurate, but still locality-based account of the phenomenon might be (35), modeled on the NIC, but with no reference to case — which could still be coupled to any of the accounts discussed above for the relevance of complementizer alternations to the effect:

(35) **Preverbal trace condition**[10]
A trace in VP-external subject position must be bound within $\bar{S}$ [= CP].

The proposal in (35) as stated, however, would have represented a retreat to a *sui generis* condition specific to the complementizer-trace effect, had it been adopted in response to the failures of the NIC approach. For this reason, Chomsky (1981) proposed that (35) is a special case of a more general licensing requirement on silent syntactic elements (notated [*e*]) that included not only traces of $\bar{A}$-movement, but also traces of A-movement and silent pronouns as well:

(36) **Empty Category Principle (ECP)** (Chomsky 1981)
[$_{\alpha}$ *e*] must be properly governed.

In Chomsky's original formulation, there were two ways in which an empty element could be properly governed, only one of which was available to VP-external (preverbal) subjects in languages like English and Italian.

First, any empty element c-commanded very locally by a lexical category such as V counted as properly governed by that lexical category. This possibility came to be known as *head-government*. A trace of movement in direct object position or a postverbal VP-internal subject trace could satisfy the ECP by virtue of head-government by the verb. The "very local" c-command relation relevant to head-government was the relation called simply *government*.

An empty element that is not head-governed, however, such as a preverbal subject trace, could only satisfy the ECP in a different way: by virtue of coindexation with a very local c-commanding element, i.e. by coindexation with an antecedent that *governs* it the way a head might otherwise govern it. This possibility came to be known as *antecedent-government.* Antecedent-government was the only form of government available to a VP-external (preverbal) subject trace, since this subject position is not governed by a lexical category such as V. On this approach, (35) became just a special case of the ECP. Once again, any of the accounts available to NIC theories that explained the unacceptability of *that* but not its null counterpart when the subject is extracted could be straightforwardly adapted to the antecedent-government requirement imposed on preverbal subject traces by the ECP. The effect of the French *que/qui*

---

[10] No condition in precisely this form was proposed in the literature, but it was an implicit intermediate formulation on the way to the development of the ECP, discussed below. A version of (35) was discussed in the "underground literature" of the time under the name "Residue of the NIC", abbreviated RES(NIC).

and West Flemish *da/die* alternations could also be handled under the ECP much as they had been handled by the NIC theories that preceded it.

Chomsky's formulations of proper government and government are given for convenience in (37):

(37)a. **Proper Government:** α properly governs β if and only if α governs β [and α is a lexical category].

b. **Government***:* Consider the structure [$_{\beta}$ ... γ... α...γ...], where

(i) α = X° [*head government*] or is coindexed with γ [*antecedent government*],
(ii) where ϕ is a [barrier to government], if ϕ dominates γ, then ϕ dominates α, and
(iii) α c-commands γ.

In this case, α *governs* γ.

Much as NIC accounts of the complementizer-trace effect attempted to link them to independent puzzles about the distribution of reflexives, ECP accounts attempted to link the effect to other independent syntactic puzzles. Indeed, an enormous quantity of syntactic research for about a decade was devoted to possible extensions of the ECP (and modifications of the ECP) to ever wider ranges of phenomena. Chomsky himself suggested, for example, that the ECP holds of traces of A-movement as well as Ā-movement, regulating raising to subject from embedded clauses (among other things). In structures like (38a), he suggested that the trace of raising satisfies the ECP thanks to head-government by *likely* across a bare TP complement that was said to not block government. In the unacceptable but superficially parallel (38b), by contrast, the full CP posited as the complement of *probable* was said to block government of the subject trace by the higher adjective — violating ECP.

(38)a. Mary is likely [$_{TP}$ __ to study syntax].
(cf. *It is likely that Mary studies syntax*.)

b. *Mary is probable [$_{CP}$ [$_{TP}$__ to study syntax]].
(cf. *It is probable that Mary studies syntax*.)

Silent elements other than traces were also viewed as subject to the ECP, restricting them to positions of head government, antecedent government being an impossible rescue strategy for silent elements not resulting from movement. Stowell (1981), for example, noted that a CP whose complementizer head is silent is excluded from preverbal subject position, as shown in (39) — an effect that he attributed to an ECP-imposed need for head-government of the null-headed CP:

(39)a. Mary believes [ø$_{C}$ the world is round].

b. *[ø$_{C}$ the world is round] is obvious.

The absence of subject *pro*-drop in languages like English and French was similarly explained by the ECP (due to lack of head government for the null subject), with a lively literature devoted to the question of why languages like Italian contrast in this respect.

If the ECP is correct as stated, extraction from any position that is not head-governed should be subject to an antecedent-government requirement, not just extraction from subject position. This possibility was famously explored by Huang (1982, 23), who noted that Ā-extraction of words like *why* and *how* appear to produce a *wh*-trace effect, just like extracted preverbal subjects. Though the strings of words in (40a-b), for example, have an acceptable parse in which *why* belongs to the higher clause (and is associated with *ask*), they cannot be parsed as indicated by the brackets and underscore. On the unacceptable parses, (40a) would presuppose that I had asked about a possible reason for Sue inviting Mary, and (40b) would be an inquiry about the identity of a method for solving a problem (specifically, the method *x* such that John asked who solved the problem using *x*). The unacceptability of examples like (40-b) appears to be cross-linguistically robust, perhaps exceptionless across the languages of the world:

(40)a. *This is the reason why I asked [whether Sue invited Mary __].
b. *How did John ask [who solved the problem __]?

Huang suggested that these effects had the same explanation as *wh*-trace effects involving subject extraction such as (20-b), and were explained by the ECP, on the assumption that *how* and *why* are syntactic adjuncts of some sort (and not head-governed).

One immediate difficulty concerns the absence of any sensitivity to the presence or absence of overt *that* with such elements, i.e. the absence of any *that*-trace effect. Most speakers appear to find no difference in availability of the indicated parse with and without overt *that* in examples like (41a-b):

(41)a. This is the reason why I heard [(that) Sue invited Mary __].
b. How did John believe [(that) we should solve the problem __]?

To solve this problem without entirely giving up on Huang's attempt at unification, Lasnik & Saito (1984), in a celebrated paper, proposed an elegant but complex architecture for ECP satisfaction and trace-deletion, developed further by Chomsky (1986). Its elegance notwithstanding, many of its details were motivated only by the puzzle of the missing *that*-trace effect for adjunct extraction, leading some to suspect that Huang's unification might be spurious (eliminating the motivation for Lasnik & Saito's ECP architecture).

Other worries raised similar suspicions. As Huang's own work showed in part, the extraction of adjuncts is extremely sensitive to all types of islands, not just embedded questions, contrasting with both subject and non-subject extraction of nominal arguments in this respect. In later work, Rizzi (1990), building on Ross (1984), showed that adjunct extraction is also blocked by a larger range of other elements such as intervening negation and downward entailing quantifiers — elements that normally fail to affect non-adjunct extraction (though see Beck 1996, Pesetsky 2000, and Kotek 2014 for possible exceptions). These observations do not of course prove that the principle at stake in (40) is distinct from the principles that account for complementizer-trace phenomena, but they did reinforce the suspicion that Huang's unification might be spurious, given the wide range of environments from which adjuncts may not be extracted.

In theory, such worries should not have affected the independent investigation of complementizer-trace effects as possible consequences of the ECP, but in actual fact they did. Both complementizer-trace effects and constraints on adjunct extraction posed urgent and exciting questions on their own, and the fact that both sets of phenomena have their roots in UG should have only added to the fascination of both topics. Sadly, disenchantment with the idea that the two sets of phenomena have a common origin led to a rapid and marked decline in research on both topics — perhaps for no reason other than general discouragement. Whereas a sizable fraction of the syntax papers at any general conference in the mid-1980s would have been devoted to subject/non-subject asymmetries and extensions of the ECP, a decade later papers on these topics had almost disappeared — and not because the outstanding problems had been solved, or the vein of new discoveries exhausted. Instead, there seemed to be an unspoken consensus that the research had taken a wrong turn somewhere, and fresh topics beckoned. The abandonment of the notion "government" by Chomsky in his early papers developing the new "Minimalist Program" helped deliver a *coup de grâce* to enthusiasm for the ECP as an engine of syntactic explanation, as many linguists found themselves convinced that the true theory of the human language faculty could not encompass complex syntax-internal relations such as government.

5.2.3 C and the complementizer-trace effect in the wake of the Minimalist Program

The "ECP boom" of the 1980s, which stimulated widespread research on complementizer-trace effects and sister phenomena, predated several significant ideas about the workings of natural-language syntax that came to prominence only in the 1990s. Among these was the idea that syntactic movement was not optional and free as had been argued by Chomsky & Lasnik (1977), Chomsky (1981), and many others (under the slogan "Move α") — but instead was a triggered and obligatory response to the presence of an unvalued feature F on a higher head H, acting as a *probe*. The probing feature F, it was argued by Chomsky (1995) and others, seeks the closest bearer of a comparable feature in its c-command domain (its *goal*), and (if F bears the so-called EPP property) copies the goal as a sister to the phrase headed by H. The result was a characterization of movement as "internal Merge", licensed only when a probing feature on a syntactic head requires it. This "last resort" property of internal Merge, subject to a strong locality requirement (only the closest probe counts), was one of several economy conditions on derivations proposed as part of the development of the Minimalist Program in the 1990s.

In light of this reorganization of the theory of syntactic movement, Pesetsky & Torrego (2001) suggested that it might be time to reopen the search for the right account of complementizer-trace phenomena and its sister effects. More specifically, they noted that the probe-goal/last-resort theory of movement offered new possibilities for understanding effects of this sort. Their particular proposal also rested on a new analysis of words like *that*, traditionally viewed as complementizers.

Their starting point was a much older proposal by Koopman (1983), who had suggested that an asymmetry in English T-to-C movement (observable in matrix *wh*-questions) might constitute a sister effect in the "complementizer-trace" family. In matrix questions that lack a

modal or aspectual auxiliary verb, whenever a non-subject is questioned, T-to-C movement can be observed in the form of the auxiliary *do* appearing to the left of the subject. Crucially, this raised *do*, though obligatory in matrix questions when a non-subject is *wh*-moved, is obligatorily absent when it is a subject that is questioned. We might call this a *do*-trace effect:

(42) ***do-trace* effect in matrix questions** (Koopman 1983)
a. What did Mary buy __?
b. *What Mary bought __?
c. *Who did __ buy the book?[11] [*unless *did* is emphatic]
d. Who __ bought the book?

Presupposing an ECP account of complementizer-trace effects, Koopman had suggested that the fronted auxiliary in (42c) might block antecedent government of the subject trace, though she did not develop this proposal in detail. Pesetsky & Torrego (2001) added the new observation that the same *do*-trace effect can be observed in an embedded *declarative* clause in the Belfast English dialect studied by Henry (1995). In this dialect, when *that* is absent from a subordinate clause successive-cyclic interrogative *wh*-movement may trigger T-to-C movement instead:

(43) **Belfast English: T-to-C triggered by successive-cyclic *wh*-movement**
a. What did Mary claim [did they steal __]?
b. Who did John say [did Mary claim [had Sue feared [would Bill attack __]?

In an embedded declarative clause introduced by *that*, Belfast English shows a *that*-trace effect when a subject is extracted, as (44a) shows. Crucially, in an embedded declarative clause without *that*, subject extraction may not trigger T-to-C movement like that seen in (43a-b), as (44b) shows:

(44) **Belfast English: *that*-trace and *do*-trace effect**
a. Who do you think [(***that)** __ left].
(Henry 1995, 128)

b. *Who did John say [**(*did)** __ go to school]?
(Alison Henry, p.c to Pesetsky & Torrego; *unless *do* is emphatic)

Pesetsky & Torrego suggested that the parallel between (44a) and (44b) might hold the key to the complementizer-trace effect. Since some principle evidently prevents otherwise obligatory T-to-C movement from applying when a subject is extracted as in (44a), they asked whether the same principle might not govern the comings and goings of *that* in C as well. Controversially, they proposed that the word *that*, contrary to popular belief, is not really a complementizer, but is instead a tense or aspect element raised to C — an invariant allomorph of

---

[11] *Do* is possible with contrastive focus on the truth value, but this use of *do* has a different source, as first argued by Chomsky (1957), as can be seen from the fact that it is available when there is no *wh*-movement, e.g. *You're wrong! Mary did buy the book*.

the auxiliary verbs found in matrix questions like (42a-d) and embedded declaratives like (43a-b). On this view, English declarative C is actually a null morpheme, and elements previously thought to instantiate C such as *that* and *for* have actually moved to C from a lower position such as T.

The interaction of a feature-based theory of movement with economy conditions on movement, Pesetsky & Torrego argued, explains why movement of the subject to the specifier of CP suppresses the possibility of T-to-C movement, deriving both complementizer-trace and *do*-trace effects as instances of the same phenomenon. They noted that C has at least two distinct probing features: one that triggers Ā-movement (e.g. a *wh*-probe) and one that triggers T-to-C movement (a *Tense*-probe). Each probe must find the closest possible goal, in keeping with locality requirements. Furthermore the needs of these probes should be satisfied by the fewest overall number of operations, in keeping with economy requirements.

Crucially, Pesetsky & Torrego argued that a preverbal subject with relevant Ā-features (for example, a *wh*-phrase) is capable of satisfying the needs of both of C's probes at once. Equally crucially, they argued that it counts as a maximally local goal for both of them. In this situation, Economy dictates that it is the subject that moves to C, rather than T, since movement of T is superfluous. The result is the obligatory absence of *that* and fronted *do* when the local subject is Ā-extracted — i.e. the complementizer-trace and *do*-trace effects. In contrast, when only a non-subject phrase satisfies C's Ā-feature requirements, T is the closest bearer of Tense features to C. In such a situation, both T and the non-subject move separately, to avoid violating locality, resulting in auxiliary fronting, *that* or *for* in C (depending on finiteness and the dialect-specific rules governing the realization of the moved element).

The reason nominals were said to bear T-features relates to another important aspect of Pesetsky & Torrego's proposal: the idea that nominal case is the uninterpretable counterpart of tense in the verbal system (much as φ-featural agreement in the verbal system is the uninterpretable counterpart of meaningful φ-features in the nominal system). Though this idea forms an essential part of their analysis, space considerations prevent a fuller presentation here.

A key aspect of Pesetsky & Torrego's proposal, as noted above, was the claim that English C is phonologically null, and that the comings and goings of *that* and *for* actually represent the obligatoriness vs. impossibility of T-to-C movement, not the appearance and disappearance of the complementizer itself. A gap in their discussion concerned the obvious expectation that in some languages C might not be a null morpheme, in contrast to English. In such languages, the *that*-trace effect should surface not as an alternation between overt material in C vs. its absence under subject extraction — but rather as an alternation between a bimorphemic element in C (the complementizer plus the raised T) when a non-subject is extracted vs. a monomorphemic C (containing only the complementizer itself) when a preverbal subject is extracted.

This prediction appears to be supported by Wolof. As was shown in (6a), long-distance Ā-extraction of a non-subject triggers a bimorphemic complementizer cluster ***l-a*** in the subordinate clause. As (6b) demonstrated, subject extraction is impossible with this cluster. However, as (45) demonstrates, subject extraction *is* possible with a monomorphemic

complementizer cluster containing only ***a*** without its prefix. As Martinović notes, if *a* is overt C, then the comings and goings of *l-* can be understood as mirroring the interaction of English *that* with Ā-extraction, supporting Pesetsky & Torrego's argument that English C is null and *that* is an instance of T moved to C:

(45) **Wolof long-distance subject extraction (Martinović** 2014**)**
K-an l-a Aali xam ni mu (***l-**)a (>moo) __ gis xale bi[12]
CM-Q *l*-CWH Ali know that 3SG CWH see child DEF.SG
'Who did Ali know saw the child?'

The comings and goings of *l-* can also be observed in short-distance extraction, reinforcing the parallel with the *do*-trace effect highlighted by Koopman and by Pesetsky & Torrego:

(46) **Wolof short-distance subject vs. non-subject extraction (Martinović** 2014**)**
a. K-an **a** jox Musaa téere bi?
CM-*an* CWH hand Musa book DEF.SG
'Who handed the book to Musa?'

b. K-an **l-a** Musaa gis?
CM-*an l*-CWH Musa see

The next subsection, which introduces the final account of the complementizer-trace effect surveyed here, discusses short-distance extraction effects in greater detail.

5.2.4 *Short-distance movement and anti-locality accounts*

Though most the discussion so far has concerned long-distance extraction, the observation that something like a complementizer-trace effect also holds of short-distance extraction is not necessarily surprising, so long as the relevant configurations show an interaction of some sort with the contents of C — as is the case with Koopman's *do*-trace and Wolof contrasts like (46a-b). Another example is provided by French, where some dialects and registers fail to show the effects of the Doubly Filled Comp filter. The *que/qui* alternation surfaces in these dialects even in short-distance questions:

---

[12] The morpheme *mu* that precedes C in (45) is a so-called "subject marker", variously analyzed as agreement or as a doubling clitic, and occurs (as discussed by Martinović, citing Dunigan 1994 and Russell 2006) to the left or right of C in certain declarative clauses, as well when a non-subject is extracted:

film bi mu wax-oon ni l-a-ñu bëgg
movie DEF.SG 3SG.SBJ say-PAST that *l*-CWH-1PL.SBJ like
'the movie that s/he said we liked' Martinović (2013, 311 ex. (9b))

The element glossed as 'that' is argued by Martinović to be a second instance of C. See Martinović (2014) for further discussion of both subject marker and multiple C in Wolof.

(47)a. Quel garçon **que** tu as vu?
which boy *que* you have seen
'Which boy have you seen?'

b. Quel garçon **qui** a vu Marie?
which boy *qui* has seen Mary
'Which boy saw Mary?'

Significantly, however, there is cross-linguistic evidence that short-distance extraction of a preverbal subject is stigmatized even in environments where there is no obvious interaction with the contents of C. In an often-neglected concluding section of his famous paper on subject extraction, Rizzi (1982, 152 ff) noted that the prohibition on Ā-movement of a preverbal subject seen for long-distance extraction in (12)-(13) holds for short-distance movement as well. Recall that the genitive clitic *ne* is obligatory with a postverbal bare quantifier in complement position, a class that includes the postverbal subject of an unaccusative as well as direct objects of transitive verbs. When a bare-quantifier subject of an unaccusative undergoes short-distance *wh*-movement, Rizzi noted, *ne* is obligatory. This strongly suggests that short-distance extraction of the subject must proceed from the postverbal position, just like long-distance movement:

(48) **Italian short-distance subject extraction** (compare (13))
*Quante __ sono cadute? / ✓Quante **ne** sono cadute __?
how.many are fallen **CL.GEN**
'How many (of them) have fallen?'

Northern Italian dialects such as Trentino are, as before, even more informative, showing by the obligatory absence of both subject clitic and subject-verb agreement that extraction of the subject of every type of verb proceeds from the postverbal rather than preverbal position:

(49) **Trentino short-distance subject extraction** (compare (15))

a. Quante putele è vegnú __ con te (no other variant possible)
how.many girls is come-M.SG with you *unaccusative*
'How many girls came with you?'

b. Quante putele ha parlá con te __? (no other variant possible)
how.many girls has spoken with you *unergative*
'How many girls spoke with you?'

The absence of any obvious alternation in the contents of C relevant to these contrasts seems to put the phenomenon entirely out range of accounts that rely on linear or structural proximity to C to rule out preverbal subject — including the Fixed Subject Constraint, the *that*-trace filter and its prosody-sensitive variants, and Pesetsky & Torrego's account in terms of T-to-C movement. Furthermore, the fact that extraction of a preverbal subject would be maximally local seems to put these facts out of range of NIC and ECP accounts as well, since these accounts mandate *locality* between extraction site and landing site for subjects. It is, of course possible, as Jaeggli (1984) suggested, that extraction from preverbal position in languages like Italian and Trentino is ruled out by factors distinct from those that produce complementizer-trace effects in

languages like English. But if one wishes to maintain the conjecture that these effects have the same source, a different account is needed.

A recent proposal by Erlewine (2015) has attempted to meet this challenge, responding to the existence of contrasts parallel to (48)-(49) in Kaqchikel, a Mayan langyage of Guatemala. As is the case in many Mayan languages, the Kaqchikel verb shows a special morphological pattern when its subject is Ā-extracted. The so-called class A (ergative) morpheme that typically agrees with the subject of transitive clauses is replaced by an invariant morpheme whose traditional name is *Agent Focus* (AF) — while class B (absolutive) agreement remains intact. The extraction in question can involve focus, quantification, or a variety of other Ā-processes. Here we show examples with interrogative *wh*-movement. AF does not appear when any element other than the subject is extracted:

(50) **Agent Focus morphology replaces A-agreement iff subject is extracted**

a. Achike *x-ø-u-tëj /✓ x-ø-tj-ö ri wäy?
who COM-B.3sg-A.3sg-eat / COM-B3sg-eat-AF the tortilla
'Who ate the tortilla?'

b. Achike ✓x-ø-u-tëj / *x-ø-tj-ö ri a Juan?
what COM-B3sg-A3sg-eat / COM-B3sg-eat-AF Juan
'What did Juan eat?'

The clearest sign that the distribution of AF is related to the English *that*-trace effect is Erlewine's discovery of Adverb Intervention effects like those seen in English and Nupe in (23) and (24). When an adverb belonging to a certain class intervenes, AF is both unnecessary and impossible:

(51) **Intervening adverb makes AF unnecessary and impossible**

a. Achike kanqtzij x-ø-u-tëj ri wäy?
who actually COM-B3sg-A3sg-eat the tortilla
'Who actually ate the tortilla?'

b. *Achike kanqtzij x-ø-tj-ö ri wäy?
who actually COM-B3sg-eat-AF the tortilla

Erlewine demonstrates that other intervening elements, including "tucked in" *wh*-phrases in multiple questions and raised quantifiers, trigger the same difference in AF.

To account for these facts, Erlewine argues for a constraint on Ā-movement that is ironically the near-opposite of the locality requirement imposed in ECP and NIC accounts. Erlewine's proposal belongs to a family of suggestions distinct from those discussed so far that have been occasionally advanced as a possible source for the *that*-trace and *wh*-trace effects. In the spirit of earlier proposals by Saito & Murasugi (1998, 182), Bošković (1994, 261), Ishii (1999, 2004) (as well Grohmann 2003, in other domains), Erlewine suggests that what prevents local subject movement in Kaqchikel is an *anti-locality* constraint:

(52) **Spec-to-Spec Anti-Locality** (Erlewine 2015**)**
Ā-movement of a phrase from the Specifier of XP must cross a maximal projection other than XP.

What goes wrong when Ā-movement from the specifier of TP targets the specifier of CP in simple sentences, according to (52), is the fact that only TP is crossed by this movement, violating Spec-to-Spec Anti-Locality. If adverbs like *kanqtzij* 'actually' are hosted by a maximal projection that lies between CP and TP, the Anti-Locality condition does not rule out the extraction. Erlewine argues further that the AF morpheme that appears in examples like (50a) indicates that the subject has moved from a position that is lower than its normal position as the specifier of TP — much like absence of subject-verb agreement observed in comparable circumstances in Trentino. By moving from a lower position, Anti-Locality is once again by-passed.

A very similar account can clearly be offered for the Italian and Trentino facts in (48) and (49), and in fact can be extended to long-distance extraction. If long-distance subject extraction in an English clause introduced by *that* must stop at the specifier of CP for reasons of locality, but is prevented from doing so by Anti-Locality, the movement in question can be excluded. The ameliorating effects of intervening adverbs, which seemed largely beyond the reach of NIC and ECP accounts,[13] is a central fact supporting Erlewine's approach.

On the other hand, the fact that complementizer alternations are relevant to the effect when the subject is extracted long-distance (central to most previous accounts) is now problematic. It is not clear on an Anti-Locality approach why the presence or absence of particular material in C should affect the acceptability of subject extraction. Both Erlewine (2014) and Brillman & Hirsch (to appear) suggest that long-distance extraction might sometimes be permitted to skip landing in specifier of CP, perhaps because some embedded clauses are smaller than CP in the first place. Simple English *wh*-questions such as *who left* are also problematic, raising the possibility that the subject might be able to remain in situ in such examples.[14]

6. Other unifications and puzzles

As should be clear by now, the puzzle of the complementizer-trace effect has not been conclusively solved. Certain aspects of the phenomenon interact with C, displaying an interaction between the contents of C (both externally merged or moved) and the locality of movement into its specifier (subject vs. non-subject movement). Other aspects of the puzzle appear to have an Anti-Locality character, as just discussed. Whether the phenomenon will

---

[13] See Browning (1996) for more discussion. Pesetsky & Torrego's approach does offer an account of Adverb Intervention not unlike Erlewine's in spirit (Pesetsky & Torrego 2001, 376).

[14] In essence, Erlewine's proposal freezes a subject in place, when it occupies the specifier of TP, so extraction of a subject is only possible if it launches from a lower position or somehow moves in a non-standard manner. In this respect, the proposal resembles another recent account by Rizzi & Shlonsky (2007), which directly stipulates the immovability of subjects, with exceptions deriving from the distribution of features in the clausal left-periphery that impose the freezing effect.

finally yield to a new synthesis of these ideas, or whether the correct account requires an entirely new approach, is a topic for the future.

At the same time, a number of additional facts and problems have played a role in the discussion and should be briefly noted here, since they too may form part of the puzzle that leads to a comprehensive solution:

6.1 Covert movement

Kayne (1979) observed that for some speakers of French, the particle *ne* that accompanies negative phrases (and some other phrase types) in high registers may be separated by a clause boundary from that negative phrase — and acts as a scope marker. For many of the relevant speakers, when wide scope for the negative element is forced by the placement of *ne* in a higher clause, a subject negative phrase is degraded in acceptability:

(53) The "*ne-personne*" facts

a. J'ai exigé qu'ils **n**'arrêtent personne. (narrow scope object)
I have required that they *ne* arrest nobody
'I have required that there be nobody *x* such that they arrest *x*.'

b. Je **n**'ai exigé qu'ils arrêtent personne. (wide scope object)
I *ne* have required that they arrest nobody
'There is nobody *x* such that I have required that they arrest *x*.'

c. J'ai exigé que personne **ne** soit arrêté. (narrow scope subject)
I have required that nobody *ne* be arrested
'I have required that there be nobody *x* such that *x* is arrested.'

d. *Je **n**'ai exigé que personne soit arrêté. (*wide scope subject)
I *ne* have required that nobody be arrested
'There is nobody *x* such that I have required that *x* be arrested.'

If scope is assigned to a quantifier as a consequence of covert movement, these facts might be another case of a sister phenomenon to the complementizer-trace effect.

Kayne observed similar contrasts in English. It has been widely argued, for example, that in a multiple *wh*-question, an in situ *wh*-phrase undergoes covert movement to a scope position near the overtly moved *wh*-phrase (Aoun, Hornstein & Sportiche 1981; Huang 1981; May 1985; Pesetsky 1987, 2000). Kayne observed that when a clause boundary separates the two, a subject/non-subject asymmetry is observed:

(54)a. ?I know perfectly well who thinks that he's in love with who.
b. *I know perfectly well who thinks that who is in love with him.

Kayne suggested an explanation for these and other contrasts in terms of the NIC, replaced by the ECP in later accounts. Rizzi (1982) in turn noted that the French paradigm in

(53) can be reproduced in Italian (once certain confounds are controlled for), and Kenstowicz (1983; 1989) demonstrated a paradigm similar to (54) in Bani-Hassan Arabic. Preverbal subjects show the effect and postverbal subjects lack it, strongly supporting a unification of these facts with the complementizer-trace effect.

The intensive discussion of subject/non-subject asymmetries in the 1980s took these observations as central. Whatever accounted for the complementizer-trace effect clearly held of covert movement as well as overt. Consequently, these observations were taken as strong arguments against accounts that crucially rely on a surface gap, such as Perlmutter's proposal, the *that*-trace filter or any purely linear account.

Problems remained, however. Aoun & Hornstein (1985) noted that the syntax of French *ne* and its interaction with quantifiers was more complex than had been assumed, with multiple dialects presenting slightly different pictures. Picallo (1985) observed that the effect in (53) is sensitive to verbal mood: detectable when the embedded clause is subjunctive, but weak or absent when the embedded clause is indicative. For English, a constant puzzle was the fact that the presence vs., absence of overt *that* seemed to make little difference in (54b). As with so many of the puzzles connected with the complementizer-trace effect, the study scope contrasts that distinguish subjects from non-subjects, a core topic of research in the 1980s, faded into near-obscurity in succeeding decades, without an obvious conclusion. Among prominent post-ECP approaches, it is far from obvious how the complementizer-centric proposal of Pesetsky & Torrego could explain facts like (53) and (54), and there is no obvious extension of Anti-Locality that might accomplish this task either.

6.2 English relative clauses

In example (31), repeated below, we examined a traditional analysis of English restrictive relative clause, according to which *wh*-movement, *that* in C and a rule of free deletion for either element interact with the Doubly Filled Comp Filter to predict the three (but crucially not four) variants that are allowed:

(55)a.	[The person who ~~that~~ Mary met] is my friend.	*that* deleted
b.	[The person ~~who~~ that Mary met] is my friend.	*who* deleted
c.	[The person ~~who~~ ~~that~~ Mary met] is my friend	*who* and *that* deleted
d.	*[The person who that Mary met] is my friend.	violates Doubly Filled Comp Filter

This picture holds only when the relativized position is not the subject of the relative clause. When the subject is relativized, possibility (c) disappears:

(56)a.	[The person who ~~that~~ met Mary] is my friend.	*that* deleted
b.	[The person ~~who~~ that met Mary] is my friend.	*who* deleted
c.	*[The person ~~who~~ ~~that~~ met Mary] is my friend	*who* and *that* deleted
d.	*[The person who that met Mary] is my friend.	violates Doubly Filled Comp Filter

For many accounts of complementizer-trace phenomena, this paradigm is a double surprise. Linear-order accounts and any structural account that blocks movement from a

preverbal subject position across *that* seem to predict that (56b) should be unacceptable, contrary to fact. Similarly, these accounts seem to predict that (56c) should be acceptable, again contrary to fact. Because the judgments are the exact opposite of what most accounts of complementizer-trace phenomena predict, this puzzle is often called the "anti-*that*-trace effect", and to this day has no obvious solution. Investigation of extant proposals would require a deeper look at the analysis of English relative clauses and their counterparts cross-linguistically than is possible here.

6.3 Variation puzzles and open questions

Finally, though Poverty of the Stimulus considerations and cross-linguistic ubiquity strongly suggest that the complementizer-trace effect has its roots in UG, compelling explanations have not been found for every case in which the effect fails to appear as expected. Within Germanic, for example, there are open questions concerning the existence and nature of the effect in German (see section 5.1; also Featherston 2005) and Dutch (Reuland 1983; Bennis 1986; Den Dikken et al. 2007), as well as significant variation within Norwegian, as discussed by Lohndal (2007). It is not clear that elegant resolutions of the issues raised by these languages have yet been found, like those provided by Perlmutter, Rizzi and their successors for other languages.

Even for English, it has often been informally observed that not all speakers report the standard judgments, and the possibility of dialect variation has been raised. A small judgment study by Sobin (1987) at the University of Iowa found that a majority of (undergraduate) subjects failed to detect the standard *that*-trace effect contrast, though an equal majority was sensitive to the *wh*-trace effect. Sobin suggested the possibility of a dialectal difference, and it is often informally claimed that the absence of the *that*-trace effect might be a characteristic of the American Midwest region. A later, more robust study by Cowart (1997; 2003), however, which collected data from more than 1,100 participants at five distinct U.S. locations, found no evidence of any correlation between *that*-trace judgments and geography, while continuing to find considerable variation in judgments. Recent unpublished work by Chacón, Fetters, Kandel, Pelzl & Phillips (2014) confirms these findings, stressing how surprising it would have been to find dialect variation in a subtle effect whose existence in the first place cannot be deduced from the data available to the child (cf. section 2).

***

Hopefully, it is clear by now that the problem of complementizer-trace effects is significant for linguistic theory and fascinating in its own right. Of this there should be no doubt. Great progress has been made on discovering some dots that need to be connected by the correct account of these effects, and numerous attempts have been made to connect them. But what picture these dots will present when all the right connections are properly made is still, perhaps, anyone's guess.

SEE ALSO: Complementizer Agreement; Complementizer Deletion; Left Periphery of the Clause; Multiple-Wh-Questions; Overtly Marked Wh-Paths; Strong vs. Weak Islands; Subject Auxiliary Inversion; Wh-in-Situ

REFERENCES

Aoun, Joseph, and Norbert Hornstein. 1985. Quantifier types. *Linguistic Inquiry.* 16:623-637. [http://www.jstor.org/stable/4178459]

Aoun, Joseph, Norbert Hornstein, and Dominique Sportiche. 1981. Aspects of wide scope quantification. *Journal of Linguistic Research Journal of Linguistic Research.* 1:67-95.

Aoun, Joseph. 1981. The formal nature of anaphoric relations.

Aoun, Joseph. 1983. The formal nature of anaphoric relations.

Beck, Sigrid. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics.* 4:1-56.

Bennis, Hans. 1986. *Gaps and dummies.* Dordrecht: Foris Publications.

Bennis, Hans, and Liliane Haegeman. 1984. On the status of agreement and relative clauses in west-Flemish. In Wim de Geest, and Yvan Putseys, eds. *Sentential complementation.* 33-54. Dordrecht: Foris. [https://pure.knaw.nl/portal/files/489847/14935_292_bennis.pdf]

Bošković, ☐Željko. 1994. D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis.* 24:247-286.

Brandi, Luciana, and Patrizia Cordin. 1981. Dialetti I italiano: Un confronto sul parametro del soggetto nullo. *Rivista di Grammatica Generativa.* 6:33-87. [http://arca.unive.it/bitstream/10278/2019/1/2_brandi_cordin.pdf]

Brandi, Luciana, and Patrizia Cordin. 1989. Two Italian dialects and the null subject parameter. In Osvaldo Jaeggli, and Kenneth J. Safir, eds. *The null subject parameter.* 111-142. Dordrecht: Kluwer.

Bresnan, Joan. 1972. Theory of complementation in English syntax. Doctoral dissertation, Massachusetts Institute of Technology.

Bresnan, Joan. 1977. Variables in the theory of transformations. In Peter W. Culicover and Thomas Wasow and Adrian Akmajian, ed. *Formal syntax.* 157-196. New York: Academic Press.

Brillman, R. J., and Aron Hirsch. to appear. An anti-locality account of English subject/non-subject asymmetries. eds. Proceedings of CLS 50. Chicago: Chicago Linguistic Society. [https://dl.dropboxusercontent.com/u/30985229/Website/Papers/BrillmanHirschCLS50Proceedings.pdf]

Browning, M. A. 1996. CP recursion and that-t effects. *Linguistic Inquiry.* 27:237-255. [http://www.jstor.org/stable/4178935]

Chacón, Dustin, Michael Fetters, Margaret Kandel, Eric Pelzl, and Colin Phillips. 2014. Indirect learning and language variation: Reassessing the *that*-trace effect. Unpublished manuscript. University of Maryland & Yale.

de Chene, Brent. 2000. Prosody and subject traces. Unpublished manuscript. Waseda University. Tokyo. [http://www.f.waseda.jp/dechene/PST.pdf]

Chomsky, Noam. 1957. *Syntactic structures.* Mouton.

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

Chomsky, Noam. 1986. *Barriers.* Cambridge, Mass: MIT Press.

Chomsky, Noam. 1995. Categories and transformations. In *The minimalist program.* 219-394. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1980. On binding. *Linguistic Inquiry.* 11:1-46. [http://www.jstor.org/stable/4178149]

Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry.* 8:425-504. [http://www.jstor.org/stable/4177996]

Comrie, Bernard. 1973. Clause structure and movement constraints in Russian. In Claudia Corum, T. Cedric Smith-Stark, and Ann Weiser, eds. *You take the high node and I'll take the low node: Papers from the comparative syntax festival.* 291-304. Chicago: Chicago Linguistic Society.

Cowart, Wayne. 1997. Experimental syntax : Applying objective methods to sentence judgments. Thousand Oaks, Calif: Sage Publications.

Cowart, Wayne. 2003. Detecting syntactic dialects: The that-trace phenomenon. Talk delivered at 39th Regional Meeting of the Chicago Linguistic Society.

Culicover, Peter W. 1993a. The adverb effect: Evidence against ECP accounts of the *that*-t effect. In Amy J. Schafer. ed. Proceedings of the North East Linguistic Society 23 pp. 97-111. University of Ottawa: Graduate Linguistic Student Association.

Culicover, Peter W. 1993b. Evidence against ECP accounts of the *that*-t effect. *Linguistic Inquiry.* 24:557-561. [http://www.jstor.org/stable/4178827]

Den Dikken, Marcel, Judy Bernstein, Christina Tortora, and Rafaella Zanuttini. 2007. Data and grammar: Means and individuals. *Theoretical Linguistics.* 33:335-352.

Dunigan, Melynda B. 1994. On the clausal structure of Wolof. Doctoral dissertation, University of North Carolina at Chapel Hill.

Erlewine, Michael Yoshitaka. 2014. Why the null complementizer is special in the English that-trace effect. Unpublished manuscript. MIT. [http://ling.auf.net/lingbuzz/002029]

Erlewine, Michael Yoshitaka. 2015. Anti-locality and optimality in kaqchikel agent focus. *Natural Language & Linguistic Theory.* to appear [unpublished ms., MIT; http://ling.auf.net/lingbuzz/001841]

Featherstone, Sam. 2005. *That*-trace in German. *Lingua.* 115:1277-1302.

Freidin, Robert, and Howard Lasnik. 1981. Disjoint reference and wh-trace. *Linguistic Inquiry.* 12:39-53. [http://www.jstor.org/stable/4178204]

Gathercole, Virginia C. Mueller. 2002. Monolingual and bilingual acquisition: Learning different treatments of *that*-trace phenomena in English and Spanish. In D. Kimbrough. Oller, and Rebecca E. Eilers, eds. *Language and literacy in bilingual children.* Vol. Child language and child development ; 2. 220-254. Clevedon, England ; Buffalo: Multilingual Matters.

Grohmann, Kleanthes K. 2003. *Prolific domains: On the anti-locality of movement dependencies.* Amsterdam: John Benjamins.

Henry, Alison. 1995. *Belfast English and standard English.* Oxford: Oxford University Press.

Huang, C. T. James. 1981. Move wh in A language without wh movement. *The Linguistic Review.* 1:369-416.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, Massachusetts Institute of Technology. [http://dspace.mit.edu/handle/1721.1/15215]

Ishii, Toru. 1999. Cyclic spell-out and the that-t effects. In Sonya Bird, Andrew Carnie, Jason Haugen, and Peter Norquest, eds. *Proceedings of WCCFL 18(220).* 220-231.

Ishii, Toru. 2004. The phase impenetrability condition, the vacuous movement hypothesis, and that-t effects. *Lingua.* 114:183-215. [http://www.sciencedirect.com/science/article/pii/S0024384103000457]

Jaeggli, Osvaldo. 1984. Subject extraction and the null subject parameter. In Charles Jones, and Peter Sells, eds. *Proceedings of NELS 14.* 132-153. Amherst, MA: UMass Graduate Student Linguistic Association.

Kandybowicz, Jason. 2006. *Comp-trace* effects explained away. In Donald Baumer, David Montero, and Michael Scanlon, eds. *Proceedings of the 25th West Coast Conference on Formal Linguistics.* 220-228. Somerville, MA: Cascadilla Proceedings Project. [http://www.lingref.com/cpp/wccfl/25/paper1452.pdf]

Kayne, Richard S. 1979. Two notes on the NIC. In Adriana Belletti, Luciana Brandi, and Luigi Rizzi, eds. *Theory of markedness in generative grammar.* 317-346. Pisa: Scuola Normale Superiore di Pisa.

Kayne, Richard S. 1980. Extensions of binding and case-marking. *Linguistic Inquiry.* 11:75-96. [http://www.jstor.org/stable/4178151]

Kenstowicz, Michael. 1983. The null-subject parameter in modern Arabic dialects. In Charles Jones, and Peter Sells, eds. *Proceedings of NELS 14.* 207-219. Amherst, MA: Graduate Linguistics Student Association.

Kenstowicz, Michael. 1989. The null subject parameter in modern Arabic dialects. In Osvaldo Jaeggli, and Ken Safir, eds. *The null subject parameter.* 263-275. Dordrecht: Kluwer.

Keyser, Samuel Jay. 1975. A partial history of the relative clause in English. In Jane Grimshaw, ed. *Papers in the history and structure of English.* Amherst, MA: Massachusetts Occasional Papers in Linguistics.

Kotek, Hadas. 2014. Composing questions. Doctoral dissertation, MIT. [http://ling.auf.net/lingbuzz/002231]

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry.* 15:235-289. [http://www.jstor.org/stable/4178382]

Lohndal, Terje. 2007. *That*-t in scandinavian and elsewhere: Variation in the position of C. *Working Papers in Scandinavian Syntax.* 79:47-73. [http://ling.auf.net/lingbuzz/000474]

Martinović, Martina. 2013. The subject/non-subject asymmetry in Wolof. In Stefan Keine, and Shayne Sloggett, eds. *That-trace effects without traces: An experimental investgiation.* Vol. Proceedings of NELS 42. 309-322. Amherst, MA: GLSA, UMass Amherst. [http://home.uchicago.edu/martinam/Selected_papers_&_presentations_files/Martinovic_1.pdf]

Martinović, Martina. 2014. *Wh*-morphology and cyclicity in Wolof. Unpublished manuscript. University of Chicao. [http://ling.auf.net/lingbuzz/001874]

May, Robert. 1985. *Logical form: Its structure and derivation.* Cambridge, Mass.: MIT Press.

McDaniel, Dana, Bonnie Chiu, and Thomas Maxfield. 1995. Parameters for wh-movement types: Evidence from child language. *Natural Language and Linguistic Theory.* 13:709-753.

Pearl, Lisa, and Jon Sprouse. 2013. Computational models of acquisition for islands. In Jon Sprouse, and Norbert Hornstein, eds. *Experimental syntax and island effects.* 109-131. Cambridge, UK: Cambridge University Press.

Perlmutter, David M. 1968. Deep and surface structure constraints in syntax. Doctoral dissertation, Massachusetts Institute of Technology. [http://hdl.handle.net/1721.1/13003]

Perlmutter, David M. 1971. *Deep and surface structure constraints in syntax.* New York: Holt, Rinehart and Winston.

Pesetsky, David. 1979. Complementizer-trace phenomena and the nominative island condition. In Edward Battistella, ed. *Proceedings of the ninth annual meeting of the North Eastern Linguistic Society (CUNY forum 7-8).* Vol. 2. New York: Queens College Press.

Pesetsky, David. 1982. Complementizer-trace phenomena and the nominative islands condition. *The Linguistic Review.* 1:297-344.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric J. Reuland, and Alice G. B. ter Meulen, eds. *The representation of (in)definiteness.* Cambridge, Mass.: MIT Press.

Pesetsky, David. 2000. *Phrasal movement and its kin.* Cambridge: MIT Press.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In Michael Kenstowicz, ed. *Ken Hale: A life in language.* 355-426. Cambridge, Mass.: MIT Press.

Phillips, Colin. 2013. On the nature of island constraints. Ii: Language learning and innateness. In Jon Sprouse, and Norbert Hornstein, eds. *Experimental syntax and island effects.* 132-157. Cambridge, UK: Cambridge University Press.

Picallo, Carme. 1985. Opaque domains. Doctoral dissertation, CUNY.

Reuland, Eric. 1983. Movement vs. Merger: Relations between inflection and verb. In *13th meeting of the northeast linguistics society.* Montreal: Université du Québec à Montréal.

Ritchart, Amanda, Grant Goodall, and Marc Garellek. 2015. Prosody and the *that*-trace effect: An experimental study. Talk delivered at 33rd West Coast Conference on Formal Linguistics. Vancouver, Canada.

Rizzi, Luigi. 1982. Negation, wh-movement and the null subject parameter. In *Issues in Italian syntax.* 117-185. Dordrecht - Holland: Foris.

Rizzi, Luigi. 1990. *Relativized minimality.* Cambridge, Mass.: MIT Press.

Rizzi, Luigi, and Ur Shlonsky. 2007. Strategies of subject extraction. In Uli Sauerland, and Hans-Martin Gärtner, eds. *Interfaces + recursion = language? : Chomsky's minimalism and the view from syntax-semantics.* Vol. Studies in generative grammar, 89. 115-160. Berlin ; New York: Mouton de Gruyter. [http://www.loc.gov/catdir/toc/ecip0714/2007013045.html]

Ross, John R. (writing as Haj Ross). 1984. Inner islands. In Claudia Brugman, Monica Maccaulay, Amy Dahlstrom, Michele Emanatian, Birch Moonwoman, and Catherine O'Connor, eds. *Control and command in tzotzil purpose clauses.* 258-265. Berkeley: Berkeley Ling. Soc. of California, 1984. xvii, 698. [http://elanguage.net/journals/bls/article/view/2372]

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT. [dspace.mit.edu/handle/1721.1/15166]

Russell, Margaret A. 2006. The syntax and placement of Wolof clitics. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Saito, Mamoru, and Keiko Murasugi. 1998. Subject predication within IP and DP. In Kyle Johnson, and Ian Roberts, eds. *Beyond principles and parameters: Essays in memory of osvaldo jaeggli.* 159-182. Dordrecht: Kluwer Academic Publishers.

Salzmann, Martin, Jana Häussler, Markus Bader, and Josef Bayer. 2013. *That*-trace effects without traces: An experimental investgiation. In Stefan Keine, and Shayne Sloggett, eds. *Proceedings of NELS 42.* Vol. 2. 149-162. Amherst, MA: GLSA, UMass Amherst. [http://www.academia.edu/3445743/2013f_with_Jana_Häussler_Josef_Bayer_and_Markus_Bader_That-trace_effects_without_traces._An_experimental_investigation]

Sobin, Nicholas. 1987. The variable status of comp-trace phenomena. *Natural Language and Linguistic Theory.* 5:33-60.

Stowell, Tim. 1981. Complementizers and the empty category principle. In Victoria Burke, and James Pustejovsky. eds. Proceedings of the 11th Annual Meeting of the North Eastern Linguistic Society pp. 345-363. University of Massachusetts/Amhert: Graduate Linguistic Student Association.

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

Thornton, Rosalind. 1990. Adventures in long-distance moving: the acquisition of complex *wh*-questions. Doctoral dissertation, University of Connecticut. [http://search.proquest.com/pqdtft/docview/303848860/fulltextPDF/D087A73A1588482FPQ/1]

Zaliznjak, Andrey Anatolyevich, and Elena Viktorovna Padučeva. 1979. Sintaksičeskie svojstva mestoimenija *kotoryj* [syntactic properties of the (relative) pronoun *kotoryj*]. In T.M. Nikolaeva, ed. *Kategorija opredelennosti-neopredolennosti v slavjanskix I balkanskix jazykax [the category definiteness/indefiniteness in the Slavic and balkan languages]*. 289-329. Moscow: Nauka.