

CHAPTER SIX

FOR IN ENGLISH INFINITIVES HAS NOTHING TO DO WITH CASE

THOMAS MCFADDEN

1. Introduction

Consider some well-known facts about *for* and overt subjects in English infinitives. Unlike finite clauses, infinitives frequently have no overt subject, as in 1a, and adding an overt subject leads to ungrammaticality as in 1b.

- (1) a. [To drink the hot sauce] would be difficult.
- b. * [Hank to drink the hot sauce] would be difficult.

An overt subject becomes possible in many infinitives if *for* appears in front of it, as in 2a. Crucially, this *for* is only allowed in an infinitive with such an overt subject, being strictly ruled out in infinitives without one, as 2b.

- (2) a. [For Hank to drink the hot sauce] would be difficult.
- b. * [For to drink the hot sauce] would be difficult.

An overt subject is also possible if the infinitival clause is the complement of certain transitive verbs like *believe*, as in 3.

- (3) Frank believed [Hank to have bought the beer].

Following other verbs, infinitives again obligatorily lack an overt subject, as the contrast between 4a and 4b with *campaign* shows. Again, an overt subject is possible in the complements of at least some of these verbs if *for* is inserted before it, as in 4c.

- (4) a. Frank campaigned [to drink the hot sauce].
 b. *Frank campaigned [Hank to drink the hot sauce].
 c. Frank campaigned [for Hank to buy the beer].

This is one of the central sets of data that inspired the theory of abstract syntactic Case formulated in the late 1970s and early 1980s (Vergnaud, 1977; Chomsky, 1980, and following). The basic intuition behind abstract Case is that DPs are somehow defective and can only appear in configurations where some other element can license them by assigning them Case. Finite tense or agreement licenses subjects in finite clauses by assigning them structural nominative Case. The crucial element for nominative assignment is lacking in infinitives, and so by default these are not allowed to have overt subjects. However, the defectiveness of infinitives in this respect can be overcome if something else is present to assign Case to the subject DP — a transitive ECM verb like *believe* or the special prepositional complementizer *for*. The notion that what is involved here is some abstract version of case comes from the fact that, in languages with full-fledged systems of overt case-marking, verbs and prepositions play an important role in determining the case in which their complement DPs appear. What is happening with subjects in English infinitives can then be seen as driven by the abstract features involved in case-marking, even though the relevant distinctions are not always realized in the overt morphology of the language.

An account of the distribution of *for* in terms of the Case needs of the following subject is initially attractive when we concentrate on the data presented above. However, I will argue in this article that it clearly fails when we broaden our empirical perspective. The difficulties begin when we consider infinitival complements of other classes of verbs in English, where the distribution of *for* and overt subjects is rather different, and Case theory offers no insight into why this should be. Matters get worse when we bring in diachronic data on *for-to* infinitives, which are strikingly inconsistent with the predictions of Case theory and cast serious doubt on accounts devised for the synchronic data. The trouble continues when we bring in cross-linguistic evidence for comparison. The conclusion we will come to is that English *for* is actually rather odd, and the details of its distribution are more likely to be due to language-particular quirks than something deep and universal like abstract Case.

I will argue that the distribution of *for* can be more accurately described in terms of restrictions on overt and covert complementizers, which apply to that as well, which has a suspiciously similar distribution. I will present evidence that these restrictions must be characterized at least

partly in prosodic terms, and will make a first attempt at precisifying and motivating these. I will then propose that the differences of detail in the restrictions on the two complementizers can be accounted for if we adopt central aspects of Rizzi (1997)'s structures for the English left periphery, but will argue against certain peripheral aspects of his account which depend on Case theory. While I cannot present a fully principled explanation of the distribution of *for*, I will lay out as much as possible the properties that such an explanation must and must not have.

2. Problems for a Case-based Approach

2.1. *want*-Class Verbs

The attractiveness of the Case-theoretic approach to the distribution of overt subjects and *for* in infinitives is based on the behavior of three classes of verbs that take non-finite complements. First, verbs like *campaign* can either have an overt subject, obligatorily preceded by *for*, or a controlled PRO, obligatorily without *for*:

- (5) a. Frank campaigned [(*)for] Hank to drink the hot sauce].
 b. Frank campaigned [(*)for] PRO to drink the hot sauce].

Second, verbs like *believe* require an overt subject in their complement, without *for*, and do not allow controlled PRO:

- (6) a. Hank believed [(*)for] Bill to have drunk the hot sauce].
 b. *Hank believed [PRO to have drunk the hot sauce].

Third, verbs like *try* simply don't allow an overt subject in their complement, either with or without *for*, taking only a controlled PRO subject:

- (7) a. *Hank tried [(for) Bill to drink the hot sauce].
 b. Hank tried [PRO Bill to drink the hot sauce].

Case theory can make sense of these three patterns as follows. Overt DPs require Case, while PRO cannot have it.¹ Infinitival clauses by default

¹ Alternatively, following Chomsky & Lasnik (1993); Martin (2001), we could say that PRO does require Case, but a specific type—null Case—which is barred on overt DPs. For present purposes, this comes out to the same thing as the version of

have no source of Case for DPs in the subject position, and thus typically allow only PRO subjects. However, one type of infinitive has a special complementizer *for* which does assign Case to the subject position and thus appears with overt DP subjects. Verbs like *campaign* can take either type of infinitival clause as their complement, and thus allow both options. Verbs like *believe* can only take infinitival clauses without *for* as their complements, but themselves assign Case to the embedded subject position. They thus require an overt subject without *for* and do not allow PRO. Finally, verbs like *try* only take infinitival clauses without *for* as their complements and do not themselves assign Case, so they only allow PRO subjects in the embedded clause.

However, English has an additional class of verbs like e.g. *want* which don't fit into any of these three patterns, and which cannot so easily be accommodated within Case theory. These verbs allow variation between an overt subject and PRO in the complement clause:

- (8) a. I want [PRO to join the circus].
 b. I want [Jeff to join the circus].

With some members of the class, *for* is preferred before an overt subject (9a), with others *for* is optional (9b), and with still others—including *want* itself—*for* is clearly dispreferred (9c):²

- (9) a. I would hate [?(for) Jeff to join the circus].
 b. I would like [(for) Jeff to join the circus].
 c. I want [(?*for) Jeff to join the circus].

Consider why this is problematic. The basis of the Case-theoretic analysis of the distribution of subject types is the assumption that PRO and overt DPs have mutually inconsistent Case needs. However, these verbs seem capable of satisfying both. Similarly, the basis of the Case-based story of the distribution of *for* is that it shows up in all and only the non-

the theory where PRO doesn't get Case at all. Either way, overt DPs and PRO have distinct and complementary Case requirements which can be used to model the differences in their distribution.

² The strength of the preference for or against *for* with particular verbs is subject to extensive dialectal and inter-speaker variation. What is crucial is that there is broad agreement on the existence of the various sub-classes, even if there is disagreement on which verbs belong in which. In fact, verbs like *campaign* mentioned above essentially belong in the *want*-class broadly defined, at the end with strongest preference for *for* along with an overt subject.

finite clauses where it is required for Case purposes. But after at least some of these verbs, it can (optionally) go missing. To put it a different way, if *for* shows up where it does in order to deal with something so important as Case licensing, how can it be optional in a sentence like 9b?

How could we make sense of the behavior of this class of verbs and their complements within Case theory? I can think of two main possibilities. The first possibility is that (some) *want*-class verbs can in fact license the embedded subject via ECM, i.e. they are like *believe* in this respect. The initial problem with this idea is that we would have to say that, unlike the *believe*-class verbs, the *want*-class verbs only optionally assign Case to the subject of their complement. The non-Case-assigning option would have to be what underlies both the examples with *for* before an overt subject and those with a controlled PRO. In other words, we would have no principled explanation of why these verbs differ from the *believe*-class verbs in allowing control complements. A further problem for this approach is that the embedded subject below *want*-class verbs simple does not show the close association with the matrix clause that is found with verbs like *believe*. In a tradition of work going back at least to Postal (1974), it has been shown that the embedded subject behaves as though it were the matrix direct object of these verbs according to a number of syntactic and semantic diagnostics. The simplest one to demonstrate is that, under passivization, the embedded subject can become the matrix subject, as in 10a. Crucially, subjects embedded in infinitives below *want*-class verbs do not show these kinds of behavior, including not being able to raise to matrix subject position in passives as in 10b (for additional evidence and discussion, see Lasnik & Saito, 1991; Bošković, 1997; Martin, 2001):

- (10) a. John_i was believed/proven/made out t_i to be sick.
 b. *John_i was wanted/preferred/liked t_i to be sick.

The patterns of object-like behavior are the main justification for the ECM analysis of *believe*-class verbs — the embedded subject behaves like the object of the matrix verb because, like a normal object, it receives structural accusative case from that verb. Since this object-like behavior is missing in the complements of *want*-class verbs, an analysis of their subject licensing properties in terms of ECM seems seriously misguided.

The second possibility for making sense of *want*-class complements within Case theory is to suppose that there's a non-overt counterpart of *for* that is involved in Case-licensing in examples like 9b and 9c, where we have an overt subject but no overt *for* (see e.g. Bošković, 1997; Martin,

2001, for different versions of this idea, which goes back at least to Bresnan, 1972). In other words, the difference between the two variants in 9b would not be whether *for* is present, but whether it is pronounced. This would provide a uniform account for the Case-licensing of the overt subjects that appear here and would correctly distinguish between these structures and those with ECM verbs like *believe*.

While I ultimately agree that the correct analysis of the relevant sentences does involve a null version of *for*, and will adopt this position in the analysis I propose below, I submit that doing so undermines any explanatory power that a Case-based story about the distribution of *for* had. Such a theory of *for* was attractive in that it provided a way to understand the apparent simple correlation between the presence of overt *for* and overt subjects. Once we begin positing null instances of *for* to plug gaps in that correlation, we can no longer say that we can explain where *for* must and must not appear. Furthermore, if there is a null version of *for* which can appear in sentences like 9b and is sufficient to Case-license the overt embedded subject, we are at a loss to explain why it is not sufficient in sentences like 9a and must give way to the overt version. There is nothing in principle to stop of from saying that both null and overt *for* are able to assign Case to a following subject, but once we do this, we are admitting that Case doesn't explain the distribution of overt *for*. Case theory could be used to account for the occurrence of the relevant syntactic head, but something else would be needed to model the conditions on its pronunciation.

2.2. Middle English

Additional questions are raised by data on the origins of the *for-to* construction. The first infinitives with *for-to* show up in the transitional period between Old English and Middle English (henceforth ME). However, these early example are a bit funny from a modern English perspective. First of all, throughout ME, the vast majority of *for-to* infinitives don't have an overt subject.³

- (11) a. I ne come not in-til erþe [for to do mi wille]
 'I didn't come to the earth to do my will.'
 (BENRUL,10.333)

³ According to my searches of the PPCME2 (Kroch & Taylor, 1999), 1 of 391 *for-to* infinitives in the period 1150-1250 has an overt subject. For 1250-1350, it is 3 of 315, 1350-1420 16 of 1292, and for 1420-1500, 18 of 352. See also Pak (2006).

- b. ... and wente wythout the wal for to walke
 ‘... and went outside the wall in order to walk.’
 (REYNAR, 11.179)

The few that do have an overt subject place it before *for*, not between it and to:

- (12) For it es a velany, a **man for to** be curiously arrayed apou his
 heuede with perré and precyous stanes
 ‘For it is a disgrace for a man to be strangely adorned on his
 head with jewels and precious stones.’
 (ROLLTR, 29.609, from Pak, 2006)

So we don’t yet see the familiar correlation between overt subjects and *for* in infinitive clauses. This is interesting, but doesn’t necessarily have much significance for the nature of Case theory or the analysis of modern English—it just seems to indicate that the syntax of *for* was different in ME, perhaps that it was merged somewhere in the articulated Infl domain rather than the C domain. Indeed, so far ME looks quite similar to Belfast English, which has been analyzed in Case-theoretic terms by Henry (1992).

The ME data which are truly revealing and problematic (and crucially different from those in Belfast English) are those like 13 below. These are sentences containing infinitival clauses which are not complements of transitive verbs, have no *for*, and yet do have an overt subject:⁴

- (13) The thridde grevance is [a man to have harm in his body].
 ‘The third grievance is for a man to be bodily injured.’
 (CTPARS, 310.C1.941)

Infinitivals of this type appear as predicate clauses following a copula (13), as surface clausal subjects (14a) and as extraposed subjects of predicate adjectives (14b) and nouns (14c):

- (14) a. [a man to pride hym in the goodes of grace] is eek an
 outrageous folie
 ‘For a man to pride himself in the gifts he has received
 by grace is also an outrageous folly.’
 (CTPARS, 302.C2.562)

⁴ See Garrett (2011) for discussion of the development of this clause type.

- b. And so was hit syttyng [Crist to teche this mekenesse].
‘And so it was fitting for Christ to teach this meakness.’
(WYCSER, 350.2203)
- c. it is a woodnesse [a man to stryve with a strengre or a
moore myghty man than he is hymself]
‘It is madness for a man to quarrel with a stranger or a
man who is stronger than himself.’
(CTMELI, 231.C1.537)

They also appear as adjunct and appositional clauses:

- (15) a. ‘That were shame unto the,’ seyde sir Launcelot, ‘[thou
an armed knyght to sle a nakyd man by treson].’
““That would be a disgrace on you,” said sir Lancelot,
“for you, an armed knight, to slay a naked man by
treason”.’
(MALORY, 206.3373)
- b. hym thought no worship to have a knyght at such avayle,
[he to be on horsebacke and hys adversary on foote]
‘It seemed to him no honor to have a knight at such a
disadvantage, for himself to be on horseback and his
adversary to be on foot.’
(MALORY, 40.1297)

The construction is thus securely attested and reasonably diverse in its distribution. Still, all of the examples I have found have two related things in common. First, they would most naturally be translated into modern English as *for-to* infinitives, and second, the subject has no obvious source for abstract Case.

Why does this matter for present concerns? Consider that parallel modern English sentences (including the translations of the examples above) absolutely require an overt *for*—they are ungrammatical without it. It is presumed that this is because *for* is there to handle Case on the embedded subject. Now, if something as deep and universal as Case-licensing is what is at stake here, we don’t expect much cross-linguistic variation. No language should allow an overt subject in such a clause without *for*, unless some other element is present to assign it Case. The fact that ME does show precisely this kind of variation, but without any other indication that something else is different about these clauses in a way that should be relevant for Case, suggests that the Case-theoretic understanding of modern English *for* is incorrect. Instead, something more

parochial and shallow that's at least partly specific to modern English must be involved here.

2.3. Language Comparison

Comparative data from languages somewhat further afield confirm that the overt distribution of *for* in modern English is more a reflection of language-specific factors than an instantiation of a universal principle. The obligatory co-occurrence of a 'prepositional complementizer' with overt subjects in certain non-finite clauses is not a pattern that consistently recurs across languages. Of course, 'prepositional complementizers' are themselves rather common—e.g. Italian *di* 'of', to which we will return below—but their appearance generally does not co-vary with that of an overt subject. Indeed, they frequently only appear in clause types that only (or at least predominantly) take a controlled PRO subject.

On the other hand, a wide array of languages allow overt subjects in various kinds of non-ECM infinitives, again reminiscent of modern English *for-to* infinitives, but with nothing like *for* or other plausible Case-licensors present. e.g., Irish allows overt subjects in extraposed infinitives like 16 (McCloskey, 1985):

- (16) Níl iontas [é mac mí-nádúrtha a thógáil.]
 is-not wonder him son un-natural raise:INF
 'It is no wonder that he should raise an unnatural son.'

Note that there can be no question of ECM here, since the matrix clause doesn't have a transitive verb, and an extraposed clause wouldn't be in the right structural position. Crucially, there is no preposition like *for* or anything else in the embedded clause that we might expect to be a Case-assigner. The embedded verb is not marked for tense, and unlike in the famous European Portuguese construction discussed by Raposo (1987), it also doesn't bear any agreement morphology. Similar patterns can be found under the traditional label Acl (accusative-with-infinitive) in Latin and Ancient Greek (see e.g. vanden Wyngaerd, 1994).

An even more striking example is the Tamil adjunct infinitive in 17 discussed by Sundaresan & McFadden (2009):

- (17) [vasu poori porikk-a] raman maavu
 vasu.NOM poori.ACC fry-INF raman.NOM flour.ACC
 vaangi-n-aan
 buy-PST-M.3SG 'Raman bought flour for Vasu to fry pooris'

The matrix clause already has a direct object *poori*, so there can't be any structural accusative from the matrix verb left over to license the embedded subject. Furthermore, the embedded clause is an adjunct, so the structural conditions for ECM aren't even met. Nonetheless, an overt subject is perfectly grammatical in the clause, bearing morphological nominative case no less. Again, patterns like this, while not ubiquitous, are also not rare cross-linguistically (see Sundaresan & McFadden, 2009, for more examples and discussion).

What is relevant for our analysis of the distribution of *for* is that the Irish and Tamil examples here are essentially parallel in terms of their semantics and distribution within the main clause as the modern English *for-to* construction. Crucially, as was the case in ME, nothing like *for* is found in the relevant examples in these other languages, yet overt subjects are licensed. Again, this makes it difficult to see how the overt distribution of *for* in modern English could be explained in terms of the Case-licensing needs of a following subject.

3. Parallels with *that*

Now that we've rejected accounts in terms of Case, we need to find an alternative—and superior—treatment of the distribution of overt *for*. In the remainder of this article, I will work towards developing such an alternative, which still leaves a number of theoretical questions open, but seems to be on the right track empirically. I propose that the right place to start is with the set of (partly English-specific) restrictions on overt and covert complementizers. Such restrictions have been independently assumed, in order to deal with the distribution of *that*. The complementizer *that* is not usually thought to be implicated in Case and subject licensing, and thus requires an account in different terms even in theories which adopt a Case-based account of *for*. As discussed by Pesetsky & Torrego (2001) and others, however, *that* and *for* have a strikingly similar distribution, so what we say about the one should actually apply to the other. In other words, if Case is not the right way to think about *that*, it isn't the right way to think about *for* either.

Here are the basic facts supporting the claim that the two complementizers have a parallel distribution. First, they both optional in

verbal complements, with variation in how easily they can be left off according to the specific verb involved.⁵

- (18) a. I would like (for) him to buy the book.
 b. I believe (that) he bought the book.
- (19) a. I would prefer ^{?*}(for) Jeff to join the circus.
 b. I dispute ^{?*}(that) Jeff joined the circus.

Second, there is a clear preference for both to be overt when an adverbial intervenes between the matrix verb and the clausal complement that they head:

- (20) a. I would like very badly ^{**?}(for) Frances to buy the book.
 b. I believe quite strongly ^{**?}(that) Frances bought the book.

Third, both complementizers are obligatory overt when heading subject clause. Both finite and non-finite clauses in subject position without an overt complementizer are sharply ungrammatical:

- (21) a. [**(For)* him to buy the book] was unexpected.
 b. [**(That)* he bought the book] was unexpected.

Fourth, both overt complementizers are impossible when a subject *wh*-trace follows. This is what has been called the *that*-trace or COMP-trace effect.

- (22) a. Who_i do you think (*that) t_i bought the book?
 b. What_i do you think (that) he bought t_i?
 c. Who_i would you like (*for) t_i to buy the book?
 d. What_i would you like (for) him to buy?

Descriptively, we can distinguish two restrictions here. On the one hand, there is a restriction *against* overt complementizers in the context preceding an extracted subject. On the other hand, there is a restriction *against covert* complementizers in a variety of contexts, largely boiling down to anywhere that isn't a complement position highly local to one of a

⁵ See Dayal & Grimshaw (2009) for intriguing discussion of the semantic and pragmatic factors that help to determine whether the complementizer actually will be overt or covert in these contexts.

particular set of verbs. I will discuss the details and possible implementations of these two restrictions in turn.

4. On the COMP-Trace Effect

The restriction against overt complementizers in particular bears some elaboration. As a result of its effects, in certain contexts overt *for* can go from being strongly preferred to being disallowed, simply because of *wh*-movement of the embedded subject:

- (23) a. I would like very badly ^{*(?)}(for) Frances to buy the book.
 b. Who_i would you like very badly (^{*}for) t_i to buy the book.

Note that this furnishes one final argument against explaining the distribution of *for* in terms of Case. As a kind of \bar{A} -movement, *wh*-movement shouldn't have any effect one way or the other on the Case-licensing needs of a DP. The Case status of *who* in 23b thus shouldn't be any different than that of Frances in 23a. Since *who* is licensed in 23b without an overt *for*, the fact that there is a strong preference for *for* to be overt in 23a simply cannot have anything to do with Case (see also Landau, 2006, for similar discussion and argumentation).

We must ask then what the nature of this effect is. How is it properly characterized, and can we find any motivation for it in terms of deeper principles? This topic has been the subject of extensive research and theorizing since the 1970s, but no fully satisfactory account has ever been developed. Chomsky & Lasnik (1977) proposed a filter operating at surface structure, which ruled out structures where a complementizer locally preceded a trace. This achieved good empirical coverage, but was a pure stipulation, which was furthermore particularly suspicious in that it seems to require reference to phonological factors within the syntax.

In the course of the 1980s, a series of attempts were made to account for the effect in terms of the Empty Category Principle (ECP), a principle motivated independently to cover an array of other facts (e.g. Lasnik & Saito, 1984; Chomsky, 1986, among many others). Details varied from one version to another, but the basic idea that all of them had in common was that the *wh*-trace of the subject, being an empty category, needed to be appropriately licensed via government. For one reason or another, the overt complementizer intervened and blocked government, whereas its covert counterpart did not. ECP based accounts avoided the modularity problem of the filter approach, but had additional problems of their own. Empirically, they were ill-equipped to deal with certain exceptions to the

effect (which we will discuss below) which seem to involve purely phonological factors rather than structural ones which would be expected to be relevant for government. Theoretically, they required certain ad hoc assumptions and very particular definitions of different kinds of government in order to work properly (in part because the ECP was also relevant for the licensing of the null complementizers themselves). In no small part because of problems with the ECP, definitions of government came to be excessively complicated in ways that seemed arbitrary and difficult to motivate in independent terms.

As a result, government was abandoned in the general house-cleaning that went along with the advent of the minimalist program. The ECP was no longer formulable, and has had to be replaced by other principles in minimalist accounts for the relevant phenomena. Newer proposals for the COMP-trace effect have taken a number of different approaches, e.g. attributing it to interactions between T-to-C movement and wh-movement of the subject (Pesetsky & Torrego, 2001) or to the Phase Impenetrability Condition (Ishii, 2004). While perhaps more theoretically palatable—at least from a current perspective—these accounts however still suffer from the empirical failings inherent to any purely syntactic approach to the effect, which will be made more clear as our discussion proceeds.

There is a strand of more recent research, however, which has produced evidence for a rather different approach to the COMP-trace effect (see e.g. Postal, 2004; Kandybowicz, 2006; Bruening, 2009, though much of the relevant facts were already known in the 1970s), which promises to avoid the problems of the other approaches discussed above. This work has shown that the COMP-trace effect has nothing specifically to do with extraction of the subject. Rather, what it actually seems to care about is whether the overt complementizer is adjacent to an empty subject position. In fact, there is reason to believe that the effect is actually phonological. I present here a selection of the relevant facts that have been noted in this connection, and direct the reader to the references cited for further evidence and discussion.

First, ellipsis eliminates the effect, both with *that* and *for* (Merchant, 2001):

- (24) a. John said that someone would write a new textbook, but I can't remember who, ~~John said that t_i would write a new textbook.~~
 b. John said it was important for someone to come, but I can't remember who, ~~John said it was important for t_i to come.~~

Second, the intonational break created by Right Node Raising, while not completely removing the effect, greatly improves things, again with both complementizers (de Chene, 1995):

- (25) a. Who_i does John doubt whether and Bill suspect that t_i cheated?
 b. That's the president I've been hoping for and you've been petitioning for t_i to be impeached.

Third, at least for some speakers, using heavily reduced pronunciations of the complementizers can lessen the effect (Kandybowicz, 2006):

- (26) a. Who_i do you hope *for/?fer t_i to win.
 b. The author_i that the editor predicts *that/?th't t_i will be adored.

Fourth, the effect can be substantially ameliorated by having an adverbial intervene between the overt complementizer and the presumed position of the trace before the verb adverbs (Bresnan, 1977, and others):

- (27) Who_i do you think [that, against doctor's orders, t_i drank the hot sauce]?

This adverb effect can only be observed with that, since, for independent reasons which I will discuss in Section 6 below, adverbs cannot intervene between for and the subject position.

In the face of these and other facts, Kandybowicz (2006) proposes to characterize the COMP-trace effect in terms of the following prosodic constraint:

- (28) *<C°, t> iff:
 i. C° & t are adjacent within a prosodic phrase, AND
 ii. C° is aligned with a prosodic phrase boundary

As Kandybowicz himself notes, this is not, as it stands, an explanation of the effect, but rather a descriptive generalization which will need to be motivated in terms of more general principles. However, it represents an improvement over previous approaches in that it states the generalization in the *correct* terms—at least partly prosodic rather than completely

syntactic—and thereby achieves better empirical coverage.⁶ It thus has a better chance of being the correct descriptive generalization and serving as the basis for an ultimate explanation.

If we assume that 28 is on the right track in this sense, an interesting possibility arises. Since 28 is supposed to be a prosodic restriction, it would be odd for it to apply specifically to traces—traces are syntactic, so the phonology should not be able to directly refer to them or distinguish them. What the prosody could be sensitive to is whether something is pronounced or not. Thus in fact, a truly prosodic version of 28 should apply to phonologically empty subject positions in general.

- (29) *The COMP-null-subject effect (CNSE)*
C_{overt} Subj_{null}

Note that this is entirely consistent with the conclusion that the effect doesn't really have anything to do with extraction. Indeed, Bruening (2009) concludes, largely on the basis of evidence and argumentation from Postal (2004), that the effect applies to other kinds of null subjects, in particular null expletives. The crucial consequence for our concerns is that this generalized CNSE will straightforwardly rule out appearances of overt *for* before PRO. In other words, we can reduce the ban on *for*-PRO-*to* in standard modern English to a generalized version of the COMP-trace effect. Note that this is a significant move towards an analysis of the distribution of overt *for* that is independent of Case. The fact that *for* never appears before PRO had a simple account based on the idea that PRO had incompatible Case requirements from those satisfied by *for*. But now an equally simple account is available, if we can say that overt complementizers are never possible before null subjects in English.

⁶ As it stands, Kandybowicz's constraint faces problems with modularity, since it involves both prosodic notions and purely syntactic ones like the category C and a subject trace. If we want to maintain that this is a prosodic constraint, the syntactic notions will have to be restated in prosodic terms—presumably in terms of how syntactic structure is fed into the PF component. There is some reason to think that by doing this we can actually move towards an explanatory account of the phenomenon. I will go part of the way towards this goal here, but much will have to be left for future research. This approach to null complementizers has a lot in common with theories which have attempted to characterize their distribution in terms of a requirement that they cliticize to the matrix predicate at some level of the grammar (see e.g. Pesetsky, 1992; Bošković & Lasnik, 2003). There are, however, important differences, e.g. in that the cliticization accounts are still implemented syntactically, whereas that of An is implemented in terms of the prosodic domains created on the basis of the output of the narrow syntax.

Now, the CNSE as it stands in 29 still lacks independent motivation, and it remains problematic from a modularity standpoint. We have reason to think that it is a prosodic constraint, but we still have reference to the category C and to subjects. If we can eliminate that syntactic residue, we will hopefully move closer to a constraint that can actually be motivated in prosodic terms. Precisely how this should be done must remain a topic for future research, but some initial speculation is possible at this point. Note first that it is commonly assumed that prosody—while phonological in nature—is clearly related to syntactic structure. The relationship between the two is presumably mediated by an operation at the interface, by which as a first approximation certain syntactic units are (indirectly) mapped onto prosodic units. In particular, much current research has argued from the perspective of (phase-based) multiple spell-out that spell-out domains as transferred from the narrow syntax to the PF branch form the basis for the construction of prosodic domains (see e.g. Adger, 2003; Kratzer & Selkirk, 2007; Kahnemuyipour, 2009). It must be the case then that the configuration described by 29, rather than directly involving C and subjects, is more properly defined in terms phase-based spell-out domains. Note in this direction that C is normally taken to be a phase-defining head, meaning that the complement TP would constitute a spell-out domain. The CNSE configuration would then involve an overt, unstressable element (the C head) adjacent to a prosodic domain with an empty left edge. If we can show that this configuration generally leads to unacceptability in English, then we will have achieved a deeper generalization, and will be closer to a true explanation.

5. The Ban on Null Cs is Prosodic Too

Turning now to the restrictions against *covert* C, there is evidence that these, too, are essentially prosodic, and here more of the details have already been worked out. In particular, An (2007) proposes that they result from the following prosodic restriction, which he argues holds quite generally:⁷

⁷ This approach to null complementizers has a lot in common with theories which have attempted to characterize their distribution in terms of a requirement that they cliticize to the matrix predicate at some level of the grammar (see e.g. Pesetsky, 1992; Bošković & Lasnik, 2003). There are, however, important differences, e.g. in that the cliticization accounts are still implemented syntactically, whereas that of An is implemented in terms of the prosodic domains created on the basis of the output of the narrow syntax.

- (30) *Intonational Phrase Edge Generalization (IPEG)*
The edge of an Intonational phrase (Int-P) cannot be empty.

According to his analysis, the contexts where *that* and *for* must be overt in English correspond to those where CP is obligatorily parsed as the edge of an Int-P. Subject clauses, being at the left edge of the sentence, are always at the left edge of an Int-P, so the IPEG forces C to be overt.⁸

- (31) a. $\boxed{\text{Int-P } [\text{CP For him to buy the book}] \text{ was unexpected}}$
b. $\boxed{\text{Int-P } [\text{CP That he bought the book}] \text{ was unexpected}}$

Complement clauses can—at least under certain circumstances—be parsed as part of the Int-P with the matrix predicate, in which case C can be null without violating the IPEG:

- (32) a. $\boxed{\text{Int-P I would like } [\text{CP } \emptyset_{\text{for}} \text{ him to buy the book}]}$
b. $\boxed{\text{Int-P I believe } [\text{CP } \emptyset_{\text{that}} \text{ he bought the book}]}$

However, if an adverbial follows the matrix verb, this breaks up the Int-P, and the following complement clause is parsed as an independent Int-P. The IPEG thus requires the complement C to be overt:

- (33) a. $\boxed{\text{Int-P I would like with all my heart}}$
 $\boxed{\text{Int-P } [\text{CP for Frances to buy the book}]}$
b. $\boxed{\text{Int-P I believe quite strongly}}$
 $\boxed{\text{Int-P } [\text{CP that Frances bought the book}]}$

6. Cartography and the Differences between *that* and *for*

In spite of their generally parallel distribution, there are some contexts in which *that* and *for* clearly differ in their behavior. When a finite clause is

⁸ I will use rounded boxes to delineate relevant prosodic domains, distinct from the standard square brackets used to delineate syntactic constituents.

the complement of an adjective, as in 34a, *that* can be pronounced or not, essentially like in the complement of a verb. In contrast, *for* cannot go unpronounced in parallel non-finite complements, as shown in 34b:

- (34) a. I'm happy [(that) you came].
 b. I'd be happy [**(for)* you to come].

I propose that we can account for this difference if we adopt Rizzi (1997)'s structure for the left periphery, along with his proposals about *that* and *for*. Rizzi posits the following hierarchy of functional material in the C domain:

- (35) [Force [Top* [Foc [Top* [Fin . . .]]]]]

He further proposes that in English *that* is in Force^o, while *for* is in Fin^o, on the basis of word order facts like the following:

- (36) a. I believe that, tomorrow, John will leave.
 b. *I would hate for, tomorrow, John to leave.

If we assume that fronted adverbials like *tomorrow* appear in a Top projection, then we expect them to be able to appear between an element in Force and the subject (which is presumably in Spec-TP, below Fin). On the other hand, they will not be able to intervene between an element in Fin and the subject. Placing *that* in Force and *for* in Fin thus accounts for the facts.

Now we can suggest the following. Adjectives that take clausal complements select ForceP. Some verbs, however, can select a reduced complement, consisting only of a FinP. Recall then that, following An (2007), null complementizers are only possible when the relevant embedded clause has been parsed into a single Int-P with the matrix predicate. Let us assume that this only happens when the complementizer is highly local to the embedding predicate. The finite complementizer *that* will be maximally local to any matrix predicate, being in Force, but the non-finite complementizer *for*, being in Fin, will be separated from the matrix predicate by an intervening Force head. Thus when Force is present, as in the complement of a matrix adjective, the infinitival complement will always form its own Int-P and *for* will have to be overt. In the complement of the right set of verbs, however, no Force will be present, meaning that Fin will be sufficiently local to the verb and parsed into the same Int-P with it, allowing *for* to be null.

The variation in clause size I'm proposing here is not exactly standard for the clause-types in question, but note that it just extends to them a certain kind of standard reasoning used for other cases. It is normally assumed that ECM/raising infinitives are TPs, while control infinitives are CPs, and this distinction is used to explain why we never get *for* in ECM and raising. If we allow the traditional variation in clause size between TP and CP, and supplement this with more recent cartographic proposals for an articulated C layer, then we should also expect variation in clause size involving newly distinguished levels like Force and Fin.

7. Against the Case Adjacency Generalization

One potential problem with adopting Rizzi (1997)'s proposals is that by so doing we may be letting crucial reference to Case theory in through the back door. This concern arises because Rizzi explicitly attempts to derive the positional difference between *that* and *for* from their involvement in Case licensing. Specifically, he retains the assumption that *for* assigns Case to the subject below it, while *that* does not. He then proposes the following generalization, which depends upon this assumption:

- (37) *Case adjacency*
 “[W]henver the case properties of the subject depend on an element of the complementizer system, no preposed phrase of any sort can intervene between this element and the subject.”
 (Rizzi, 1997, p. 300)

Rizzi offers this as an explanation for the fact that *that* can appear in Force, while *for* must be in Fin. He adopts the traditional assumption that *for* is relevant for the Case properties of the subject, which for him requires maximum locality (specifically head government). If *for* appeared in Force^o, it would be too far away, and the subject wouldn't be able to get Case. Instead, it must appear in Fin^o, sufficiently local to the following subject. On the other hand, *that*, is not involved in the Case of the subject, so it is allowed to be in Force^o. To bolster these claims, Rizzi provides an example with a different preposition behaving the complementizer of what looks like a small clause:

- (38) a. With John away on the weekend. . .
 b. * With, on the weekend, John away. . .
 c. *? John away on the weekend, we'll have to meet at another time.

We see that it is not possible to prepose an adverbial between *with* and the subject, indicating that in Rizzi's system, *with* must be in Fin° , not Force° . Again, *with* is ostensibly involved in the Case needs of the subject, as suggested by the relative unacceptability of the version in 38c without it. Rizzi thus argues that we have here further support for the connection between involvement in Case assignment to the embedded subject and restriction to Fin° .

However, I think the connection Rizzi draws here is incorrect. English has other clause types with prepositional complementizers, e.g. the minimally different gerundival in 39:

- (39) With John being away on the weekend. . .

Here again, fronting of the adverbial is impossible, so we surmise that *with* can only be in Fin° :

- (40) *With on the weekend John being away...

However, *with* is transparently not involved in Case with gerundivals, since it can be left off in contexts with an overt subject and no possibility of anything like ECM:

- (41) John being away, we had to get Peter to do it.

Indeed, it is clear that whatever handles licensing of overt subjects in gerundival clauses, it is internal to the clauses themselves and requires no support from prepositional, verbal or other external elements (see Pires, 2007, for discussion). In any case, the Case adjacency generalization about complementizer position clearly breaks down here. The requirement for particular C elements to be in a position adjacent to the subject thus does not derive from locality restrictions on Case-assignment.

Another fact that Rizzi discusses from Italian adds to the trouble for the Case adjacency generalization. We find an overt complementizer *di* in nonfinite clauses without a subject. Adverb ordering facts make it fairly clear that while the *che* of finite clauses is in Force° , this *di* is in Fin° , again like English *for*.

- (42) a. *Penso, a Gianni, che gli dovrei parlare.
'I think, to Gianni, that I should speak to him.'
b. Penso che, a Gianni, gli dovrei parlare.
'I think that, to Gianni, I should speak to him.'

- (43) a. Penso, a Gianni, di dovergli parlare.
 ‘I think, to Gianni, ‘of’ to have to speak to him
 b. *Penso di, a Gianni, dovergli parlare.
 ‘I think ‘of’, to Gianni, to have to speak to him.’

However, *di* isn’t needed in that position to assign Case to an overt subject, so why can’t it be higher? Of course, we could imagine that *di* has to assign null Case to PRO, and thus still needs to be local—this is essentially what Rizzi proposes. But the use of null Case to account for the distribution of PRO has been repeatedly and thoroughly called into question in recent years (see Manzini & Roussou, 2000; Landau, 2006; Sigurðsson, 2008; Sundaresan & McFadden, 2009, among many others), and we should be highly suspicious of its use in this context.

What all of this means is that we cannot use Case to explain the hierarchical positions available to specific complementizers. Again, there’s nothing here to support the idea that *for* in particular has anything to do with case. However, this doesn’t undermine the rest of Rizzi’s story. Accounting for differences in possible adverbial positions in terms of distinct height for different complementizers is still a viable option, and is fully consistent with accounts offered for other word-order facts in cartographic approaches. There is sufficient evidence that *for* and *that*, like *di* and *che*, appear at different heights in the structure of the left periphery. The only consequence is that we will need to find a different way to motivate this difference, i.e. to explain why *for* should be in one position while *that* is in another.

In fact, it seems plausible that this could be done in terms of the functions of the actual complementizers. English *for*, e.g. plausibly has to do with finiteness, identifying non-finite clauses with a certain (modal) character. On the other hand, *that* does not indicate finiteness or any particular modality.⁹ Rather, it marks declarative Force, as can be seen by the fact that it alternates with interrogative *if* in finite indicative and subjunctive contexts. In any case, that non-finite complementizers should consistently tend to spell out lower heads than finite ones is precisely what we should expect. Oversimplifying matters a bit, both within and across languages, non-finite clauses tend to be a bit smaller than finite ones, e.g. expressing fewer distinctions and allowing fewer positions for arguments and modifiers.

⁹ Note e.g. that *that* is found both in typical indicatives and in subjunctives, such as *I demand that he leave or I thought that he would leave.*

8. Putting the Restrictions Together

A final question we can ask is how the two types of restrictions on complementizers interact with each other. The answer will ultimately depend on how exactly they are implemented, and thus must wait for future research. However, at this stage we can at least consider some of the relevant data to establish the basic empirical situation. In particular, what happens in contexts where the two constraints would contradict each other? We've already seen one such example, which I repeat here:

- (44) a. I would like very badly ^{*}(for) Frances to buy the book.
 b. Who_i would you like very badly (^{*}for) t_i to buy the book.

Adverbial material between the matrix verb and an embedded non-finite clause strongly favors an overt complementizer in 44a, presumably due to the IPEG, but if we extract the subject as in 44b, we create a context for the CNSE, making overt *for* impossible, and the lack of *for* significantly better than in 44a. This initially suggests that the CNSE can overrule the IPEG.

However, note what happens in a slightly different context:

- (45) a. It would be instructive ^{*}(for) her to emulate the teacher.
 b. ^{*}Who_i would it be instructive (for) t_i to emulate the teacher?
 c. Who_i would it be instructive for her to emulate t_i?

Here, the embedded clause is an extraposed subject clause, and overt *for* is absolutely required. In this case, extracting the subject, as in 45b, doesn't improve the version without *for*, and in fact the result here is ungrammatical no matter what we do. Note that the problem here is not that extraction out of such clauses is banned in general, as 45c shows. This implies that the CNSE cannot actually overrule the IPEG.

Presumably also related is the following contrast:

- (46) a. ^{*}[_{CP} \emptyset_{for} John to drink the hot sauce] would be difficult.
 b. ^{*}[_{CP} For PRO to drink the hot sauce] would be difficult.
 c. [_{CP} \emptyset_{for} PRO to drink the hot sauce] would be difficult.

The fact that 46a is bad is accounted for by the IPEG — the embedded CP is at the beginning of the utterance, so it will clearly be the left edge of an Int-P, and thus the null C at its edge violates the IPEG. We can also

straightforwardly rule out 46b on the basis of the CNSE. What is unexpected is the fact that 46c is grammatical. Here as well, we seem to have a null C at the edge of an Int-P, so the IPEG should rule it out, contrary to fact. Note that a retreat to a Case story (according to which an overt *for* would have been required in 46a to license Case on *John*, but not for PRO in 46c) could solve this problem, but only by also creating others.

For example, it would cost us our unified account of parallel examples with finite clauses, where an null *that* is also ruled out:¹⁰

(47) * [_{CP} \emptyset_{that} John drank the hot sauce] was surprising.

Descriptively, it would seem that the violation of the IPEG can somehow be voided here because the following subject is silent as well as the C. One possibility that we could consider is that the fact that both C and the subject are null allows them to be left out of the construction of prosodic domains entirely, so that the leftmost Int-P actually has to at its edge. It remains to be seen how this could be made precise, and in particular whether it can be formulated in such a way that this escape hatch will not be able to incorrectly rule in examples like 45b.

References

- Adger, D. (2003). Stress and phasal syntax. *Linguistic Analysis* 33: 238–266.
- An, D.H. (2007). Clauses in noncanonical positions at the syntax-phonology interface. *Syntax* 10: 38–79.
- Bošković, Ž. (1997). *The Syntax of Nonfinite Complementation*. Cambridge, MA: MIT Press.
- Bošković, Ž. & H. Lasnik (2003). On the distribution of null complementizers. *Linguistic Inquiry* 34: 527–546.
- Bresnan, J. (1972). Theory of complementation in English syntax. Ph.D. thesis, MIT.
- Bresnan, J. (1977). Variables in the theory of transformations. In P. Culicover, T. Wasow & A. Akmajian (eds.), *Formal Syntax*. New York: Academic Press, pp. 157–196.
- Bruening, B. (2009). The *that*-trace effect is not about extraction. Slides from presentation at University of Delaware Prosody Meeting.
- Chomsky, N. (1980). On binding. *Linguistic Inquiry* 11: 1–46.

¹⁰ Of course, there is no example parallel to 46c with a finite embedded clause, because English finite clauses always have overt subjects.

- Chomsky, N. (1986). *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, N. & H. Lasnik (1977). Filters and control. *Linguistic Inquiry* 8: 425–504.
- Chomsky, N. & H. Lasnik (1993). The theory of principles and parameters. In J. Jacobs et al. (eds.), *Syntax: An International Handbook of Contemporary Research*. Berlin: Walter de Gruyter.
- Dayal, V. & J. Grimshaw (2009). Subordination at the interface. Ms., Rutgers University.
- de Chene, B. (1995). Complementizer-trace effects and the ECP. *Geneva Generative Papers* 3: 1–4.
- Garrett, A. (2011). The historical syntax problem: Reanalysis and directionality. In D. Jonas, J. Whitman & A. Garrett (eds.), *Grammatical Change: Origins, Nature, Outcomes*. Oxford: Oxford University Press.
- Henry, A. (1992). Infinitives in a *for-to* dialect. *Natural Language and Linguistic Theory* 10: 279–301.
- Ishii, T. (2004). The phase impenetrability condition, the vacuous movement hypothesis, and *that-t* effects. *Lingua* 114: 183–215.
- Kahnemuyipour, A. (2009). *The Syntax of Sentential Stress*. Oxford University Press.
- Kandybowicz, J. (2006). *Comp-trace* effects explained away. In D. Baumer, D. Montero & M. Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, pp. 220–228.
- Kratzer, A. & E. Selkirk (2007). Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24: 93–135.
- Kroch, A. & A. Taylor (1999). *Penn-Helsinki Parsed Corpus of Middle English*, 2nd edn. Philadelphia, PA: University of Pennsylvania.
- Landau, I. (2006). Severing the distribution of PRO from Case. *Syntax* 9: 32–66.
- Lasnik, H. & M. Saito (1984). On the nature of proper government. *Linguistic Inquiry* 15: 235–289.
- Lasnik, H. & M. Saito (1991). On the subjects of infinitives. In L.M. Dobrin, L. Nichols & R.M. Rodriguez (eds.), *CLS 27: The General Session*, pp. 324–343.
- Manzini, M.R. & A. Roussou (2000). A minimalist theory of A-movement and control. *Lingua* 110: 409–447.
- Martin, R. (2001). Null case and the distribution of PRO. *Linguistic Inquiry* 32: 141–166.
- McCloskey, J. (1985). Case, movement and raising in Modern Irish. In *Proceedings of WCCFL 4*, pp. 190–205.

- Merchant, J. (2001). *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford: Oxford University Press.
- Pak, M. (2006). Infinitive marking with for: A diachronic account. *Proceedings of the 29th Annual Penn Linguistics Colloquium*, 293–306.
- Pesetsky, D. (1992). *Zero Syntax*, vol. 2. Ms., MIT.
- Pesetsky, D. & E. Torrego (2001). T-to-C movement: Causes and consequences. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press.
- Pires, A. (2007). The derivation of clausal gerunds. *Syntax* 10: 165–203.
- Postal, P. (1974). *On Raising*. Cambridge, MA: MIT Press.
- Postal, P. (2004). *Skeptical Linguistic Essays*. Oxford: Oxford University Press.
- Raposo, E. (1987). Case theory and Infl-to-Comp: The inflected infinitive in European Portuguese. *Linguistic Inquiry* 18: 85–109.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*. Dordrecht: Kluwer Academic Publishers, pp. 281–337.
- Sigurðsson, H.Á. (2008). The case of PRO. *Natural Language and Linguistic Theory* 26: 403–450.
- Sundaresan, S. & T. McFadden (2009). DP distribution and finiteness in Tamil and other languages: Selection vs. Case. *Journal of South Asian Linguistics* 2: 5–34.
- van den Wyngaerd, G. (1994). *PRO-legomena: Distribution and Reference of Infinitival Subjects*. New York: Mouton de Gruyter.
- Vergnaud, J.R. (1977). Letter to Noam Chomsky and Howard Lasnik on ‘Filters and Control’. Published 2008 in R. Freidin, C.P. Otero & M.L. Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*. Cambridge, MA: MIT Press.