DE GRUYTER MOUTON
DOI 10.1515/tlr-2013-0021 — The Linguistic Review 2014; 31(1): 41–71

Ángel J. Gallego

Deriving Feature Inheritance from the Copy Theory of Movement

Abstract: This paper recasts Chomsky's (2007, 2008) *Feature Inheritance*, arguing that the syntactic relation between phase heads and non-phase heads must be regarded as that of *identity*: Non-phase heads are (copies of) phase heads. From this perspective, e.g., C and T are one and the very same unit in the lexicon, a unit that can (and sometimes must) undergo a process of Internal Merge (IM) during the derivation, creating a discontinuous object (i.e., a non-trivial chain) that gives rise to the – apparent – non-phase head / phase head distinction. The analysis has several consequences. Firstly, it provides an explanation for the fact that the features of phases heads are inherited by non-phase heads (under a copy-based approach, this is simply automatic). Secondly, it accounts for the functioning of phase and non-phase heads 'as a unit' in the Case-agreement systems (as argued by Chomsky 2004, Epstein et al. 2012). Thirdly, since non-phase heads are copies of phase heads, it also follows that they need not be seen as feature-less elements in the lexicon (as mere feature receptacles, cf. Richards 2007). Fourthly, this approach allows us to dispense with the idea that either C and T trigger simultaneous operations (Chomsky 2008), or that the operations triggered by T are actually driven by C (derivatively); this is welcome, since both alternatives involve a restricted (phase-based) margin of countercyclicity. Finally, the present analysis makes it possible to dispense with *Feature Inheritance*, understood as a brand-new, UG enriching, feature-depriving mechanism. Rejecting this commonly accepted view of *Feature Inheritance* takes into account the problems attributed to feature-movement in the syntax (Chomsky's 1995 *Move F*).

Keywords: copy theory of movement, discontinuous categories, feature inheritance, non-trivial chain, phase theory

Ángel J. Gallego: Universitat Autònoma de Barcelona. E-mail: angel.gallego@uab.cat

1 Introduction

This paper puts forward an alternative conception of Chomsky's (2008) Feature Inheritance (henceforth, FI), a process whereby some features are transmitted from phase heads (C and v*) to non-phase heads (T and V), a narrow syntactic

process that poses a series of theoretical and empirical questions (cf. Chomsky 2007, 2013, Epstein et al. 2012, Ouali 2008, Obata 2010, Obata & Epstein 2011, Richards 2007). In particular, this paper argues that FI should be seen as a consequence of phase heads and non-phase heads being one and the same element in the lexicon. The main idea behind this proposal can be stated as (1):

(1) *Feature Inheritance as Copying Thesis (FIACT)*
Non-phase heads are (copies of) phase heads

FI is considered as obligatory in the case of uninterpretable φ-features (the source of structural Case), but not much has been said about other features (but cf. Miyagawa 2009, Chomsky 2007, 2013). In principle, one could then assume that the distribution of features of a phase head is as in (2a,b), whith interpretable and uninterpretable features (uFF and iFF) occupying the complement and edge domains of a phase respectively (where α and β stand for phase head and non-phase head):

(2) a. $[_{\alpha P}\ \alpha_{[iFF]\,[uFF]} \ldots [_{\beta P}\ \beta \ldots]\ ]$ BEFORE FI (all features in α)
b. $[_{\alpha P}\ \alpha_{[iFF]} \ldots [_{\beta P}\ \beta_{[uFF]} \ldots]\ ]$ AFTER FI (uFF are inherited by β)

Though consistent with Chomsky's (2007, 2008) view, (2) seems empirically insufficient. On the one hand, we know that [iFF] (tense, mood, etc.) typically appear below C (cross-linguistically, the unmarked option); on the other hand, we also know that [uFF] can appear in C (in different West Germanic languages).

To make things more intricate, the very existence of FI raises lexical integrity worries: If lexical items (henceforth, LIs) are complexes of features in the lexicon, it is not immediately obvious why those features must split in the syntax. This is also an apparent violation of both inclusiveness and the no-tamperng conditions (cf. Chomsky 1995, 2001), which bar manipulation of LIs during the derivation, especially so if FI involves adding or removing features. Viewed this way, FI revolves around the possibility, entertained but quickly discarded by Chomsky (1995:261 and ff.), that features of LIs are accessed by syntactic operations (i.e., *Attract* / *Move F*).

This paper focuses on the status of FI in the theory of phases (cf. Gallego 2010, 2012), claiming that the FI process of Chomsky (2008) can be dispensed with if the conception of feature valuation is adjusted. This allows us to eliminate an extra operation (FI itself), an unnecessary enrichment of UG, and stick to fairly standard methodological considerations. In support of this idea, I consider empirical evidence showing that uFF and iFF can appear in C in different languages. To address this matter, I endorse the FIACT, whereby C and T are not

two independent elements in the lexicon, but actually one and the same. Consequently, I assume that the (overt) presence of features in C or in T is largely determined in the morpho-phonological component of specific languages (being thus part of parametric variation).

Note that the FIACT suggests that the C-T distinction arises once the C head is internally merged with the so far assembled structure, a somewhat unusual type of movement (head-to-phrase movement) that has occasionally been explored (cf. Rosenbaum 1967, Bresnan 1970, Chomsky 1995, Epstein 1998, Donati 2006, Shimada 2007, Saito 2012, and Kitahara 2013). Interestingly enough, there are more antecedents of this idea in the literature. Perhaps the most relevant one is Stowell's (1981) proposal that C and T are to be regarded as a 'discontinuous object.' Here I push Stowell's (1981) idea and recast it in chain-based terms, suggesting that C and T are two occurrences of the same LI (angle brackets below indicate copies).

(3) a. $[_{CP}\ C \ldots [_{v^*P} \ldots]]$
b. $[_{CP}\ C \ldots [_{CP}\ \langle C \rangle \ldots [_{v^*P} \ldots]]]$

From this conception it follows that T shares inflectional features with C simply because they are two occurrences of the same lexical item. Apart from dispensing with an unnecessary computational operation (FI), the FIACT has further welcompe consequences, making it possible to reject three FI-rooted assumptions: the existence of parallel operations (Bošković 2007, Chomsky 2008, Stepanov 2001), the feature-less status of T, and the parasitic nature of this head with respect to the activation of operations.

Discussion is divided as follows: sections 2 and 3 review the conceptual and empirical arguments for FI and the empirical challenges it must deal with. Section 4 introduces the FIACT as an alternative to FI and discusses its main advantages. In section 5, I focus on the (in)dependence of C and T in light of the FIACT. Section 6 summarizes the main conclusions.

2 Feature Inheritance: Motivations

This section reviews the motivations offered to substantiate FI. To this end, it is necessary to step back a bit in order to understand why φ-features, which are typically visible in T, are supposed to be in C in the lexicon.

Although in the system outlined by Chomsky in 2000 and 2001 CP and v*P are the phases, it is T and v* that encode the φ-features responsible for structural Case. Such asymmetry posed a tension in the way feature valuation worked.

Let us consider why. Chomsky (2001) assumes that linguistic expressions are assembled through the application of the operation Merge, which takes two objects (LIs or phrases) and constructs binary sets, as indicated in (4):

(4) a. Merge (the, book) = {the, book} (= K)
b. Merge (read, K) = {read, K} (= M)
c. Merge (v*, M) = {v*, M} (= N)
d. . . .

As the derivation unfolds, different LIs are accessed by the computation, some of which are endowed with uFF. Chomsky (2000, 2001) takes T and v* to be such LIs. Importantly, the uFF on these heads (which are unvalued in the lexicon) must be valued as soon as they enter the derivation and, according to Chomsky (2001), they are deleted "shortly after," when the operation Transfer (previously, "Spell-Out") applies, handing over the syntactic derivation to the semantic and phonological components at the phase level. The relevant timing is depicted in (5), where I indicate the effect of Transfer with outline letters.

(5) a. $[_{CP}\ T_{[uFF]} \ldots [_{v*P}\ DP \ldots]\ ]$ VALUATION (uFF → iFF)

b. $[_{CP}\ C \ldots [\![_{CP}\ T_{[iFF]} \ldots [\![_{v*P}\ DP \ldots]\!]\ ]\!]$ TRANSFER (to SEM and PHON)

A key effect of valuation is that it renders uFF and iFF identical: after it applies, both uFF and iFF have a value. However, as Epstein & Seely (2002) quickly noted, this is problematic: "if the [interpretable/uninterpretable] distinction is lost after valuation, then the distinction is *necessarily* lost "shortly after" valuation as well." (p. 74). Indeed, if φ-features are valued by the time deletion takes place, the system will not be able to distinguish interpretable features and uninterpretable ones (Chomsky 2001:5).

Chomsky (2004) concurs, and adds a slight but crucial twist to the way φ-features are valued. In particular, he assumes that φ-features are generated in C – restoring at the same time the asymmetry between v*P and CP for structural Case assignment purposes. This move offers a solution to Epstein & Seely's (2002) observation, but it has non-trivial consequences for the status of T, which becomes somewhat of a 'second hand' category, its role being entirely parasitic on C's presence. Already at this point, Chomsky suggests that C and T are a species of unit when he points out that "T functions in the Case-agreement system only if selected by C, in which case is also complete. Furthermore, in just this case T has the semantic properties of true Tense [. . .] That makes sense if C-T are really functioning as a unit in inducing agreement" (pp. 115–116).

As we can see, taking φ-features to be encoded in C is motivated on empirical grounds (Epstein & Seely's 2002 observation), but it also has a (welcome) conceptual consequence: The fact that C is the locus of φ-features is consistent with its phase head status. Once this is assumed, we have to find a rationale for φ-features to move from phase heads to non-phase heads. The next sections are devoted to that.

2.1 Chomsky (2008): Feature Inheritance and the A / A-bar distinction

In order to motivate FI, Chomsky (2008) relates this mechanism to the A / A-bar distinction, which qualifies as a third factor condition, since it has interpretive (interface) effects:[1]

> The crucial role [that the A-A′ distinction] plays at the C-I interface suggests the usual direction to determine whether [FI] is consistent with the SMT though violating NTC. If the C-I interface requires this distinction, then SMT will be satisfied by an optimal device to establish it that violates NTC, and inheritance of features of C by the LI selected by C (namely T) may meet that condition. If so, the violation of NTC still satisfies SMT.
> (Chomsky 2008:144)

Chomsky (2008) thus assumes that φ-features descend from C to T to give rise to the distinction between A (i.e., argumental) and A-bar (i.e., operator) positions.

2.2 Richards (2007): Feature Inheritance and the Phase Impenetrability Condition

Soon after Chomsky proposed FI, Richards (2007) provided a motivation for this operation based on the *Phase Impenetrability Condition* (henceforth, PIC). In Chomsky's system, the PIC captures the idea that phases reduce computational burden by cashing out specific chunks of structure in a cyclic fashion (see (5b) above). The original formulation of the PIC is shown in (6):

1 The acronyms SMT and NTC in the following quote stand for Strong Minimalist Thesis and No Tampering Condition (cf. Chomsky 2000, 2001). The SMT states that language is an optimal solution to legibility conditions (so that the traits of the language faculty are motivated in terms of interface needs). As for the NTC, it is a computational efficiency condition according to which objects created during the derivation (and their dependencies) cannot be modified.

(6) Phase Impenetrability Condition
In phase α with head H, the domain of H is not accessible to operations outside α; only H and its edge are accessible to such operations
(Chomsky 2000:108)

The PIC states that the complement domain of a phase is transferred to the interfaces at the end of the phase. This draws a line between the *complement* and the *edge* of a phase: The former is transferred to the external systems (and becomes inaccessible for further computation), whereas the latter is accessible for additional operations.

Richards (2007) provides two arguments to dispense with Chomsky's (2008) view on FI. The first argument concerns the A / A-bar distinction itself, which – Richards argues – is absent in the case of v*-V relation (because A / A-bar asymmetries are weaker in the v*P). The second and strongest argument comes from uF valuation. As Richards (2007:565 and ff.) notes, the need for FI straightforwardly follows from the combination of two independently necessary assumptions, listed in (7):

(7) a. Valuation and deletion of uFF are part of Transfer
b. Only the complement domain of phases is transferred

(7a) is motivated by Epstein & Seely's (2002) point: If valuation of uFF takes place prior to deletion, the semantics will be unable to tell whether a given feature is iF or uF. Likewise, if deletion takes place prior to valuation, the features will never be able to feed the phonological component. Therefore, these three operations (valuation, deletion, Transfer) must take place at the same time. Now it is actually (7b) that is key for Richards' (2007) reasoning to go through: since valuation and deletion must be part of Transfer, and Transfer only affects the complement domain of phases, then φ-features must move to the complement domain (that is, to the non-phase head).

Though appealing, Richards' (2007) proposal rests on two ideas that are not obviously correct: (i) the absence of the A / A-bar distinction within the v*P, and (ii) the restriction of the PIC to the complement domain of phases. The first assumption is threatened by the empirical findings by Fox (2000) and Legate (2003), and Johnson (1991) and Lasnik (1999), who show that the A / A-bar distinction is also present in the v*P. As for the limited effect of Transfer, different proposals in the literature have convincingly argued that, under certain circumstances, Transfer can also affect either the complement plus the phase head (Ott 2011 makes this claim for free relative clauses) or the entire phase (Chomsky 2004 suggests this for root clauses; cf. Obata 2010 for further discussion).

Having seen the (largely conceptual) arguments that support FI, the next section reviews some of the (largely empirical) arguments offered against FI. I will restrict myself to consider data related to the Case-agreement systems, putting aside FI in the case of iFF.

3 Feature Inheritance: Problems

Chomsky's (2008) FI makes a clear empirical prediction: uFF must appear in the complement domain at the end of the phase. Consequently, if some uF remained in the edge when Transfer applies, then a crash at the semantic interface would emerge:

(8) a. *$[_{\alpha P}\ \alpha_{[uF]} \ldots [_{\beta P}\ \beta \ldots]\ ]$
b. ✓$[_{\alpha P}\ \alpha \ldots [_{\beta P}\ \beta_{[uF]} \ldots]\ ]$

In the recent literature, different empirical phenomena have been argued to illustrate the unwanted scenario of (8a). Perhaps the most popular case concerns "agreeing complementizers" (cf. Bayer 1984, Haegeman 1992, Zwart 1993, Carstens 2003, Watanabe 2000), but other data also indicate that uFF can appear in the edge of a phase.

There are of course technical strategies available to get around the problem in (8a). One could for instance follow Miyagawa (2009), Rizzi (1997), or Uriagereka (1988), and postulate an additional functional head between C and T (call it γP), and argue that the φ-features move there. The two options are depicted in (9) (the example is modified from Carstens 2003:393), where only (9b) is in accord with the Chomsky-Richards approach to valuation:

(9) a. Kpeinzen $[_{CP}$ dan-**k** $[_{TP}$ (ik) morgen goan]] (West Flemish)
think-1.SG that-1.SG I tomorrow go-1.SG
b. Kpeinzen $[_{CP}$ C $[_{\gamma P}$ dan-**k** $[_{TP}$ (ik) morgen goan]]] (West Flemish)
think-1.SG that-1.SG I tomorrow go-1.SG
'I think that I'll go tomorrow'

In the next section I discuss these data in more detail, but I would like to note that FI is also conceptually problematic. It is because FI is a new operation (different from Merge, Agree, Transfer, etc.), which automatically entails an enrichment of UG. This is not a problem in and of itself, but the reductionist (i.e. 'from below') perspective envisaged in Chomsky (2007) encourages us to reduce the substantive content of UG.

3.1 Agreements on C: Haegeman & van Koppen (2010)

In Haegeman & van Koppen (2010) two pieces of data are provided in order to argue against FI. It is important to note that these authors do not question that C and T have φ-features, but rather that the features in C and T are *the same*.[2]

The first argument comes from coordinated subjects in Limburgian. As (10) shows, complementizers agree with the first conjunct alone (the pronoun *toow*), and not the entire coordinated subject, whereas the embedded verb manifests full (plural) agreement.

(10) Ich denk [$_{CP}$ de-**s** [$_{TP}$ [**toow** en Marie] **kump**]] (Limburgian)
I think-1.SG that-2.SG you and Marie come-2.PL
I think that you and Marie will come
(Haegeman & van Koppen 2010:3)

As Haegeman and van Koppen argue, the fact that C and T show distinct agreement markers is *prima facie* evidence that the sets of φ-features of C and T are distinct.[3]

The second empirical argument comes from External Possessor Agreement (EPA) in West Flemish. EPA emerges when two DPs (i.e., *die venten* and *underen computer* in (11) below), interpreted as possessor and possessed, agree with T and C in singular and plural respectively. Here, yet again, Haegeman and van Koppen take this to indicate that the φ-features of C and T are two different bundles.

(11) . . . [$_{CP}$ onda -**n** [$_{TP}$ **die venten** tun juste **underen computer** (WF)
because-PL those guys then just their computer
kapot **was**]
broken be-SG
. . . because André and Valère's computer broke down just then
(Haegeman & van Koppen 2010:4)

2 Another argument in favor of the idea that phase heads do not lose their features may come from clitics, if they target C and v* (cf. Roberts 2010).

3 Although the argument is well taken, it is known that agreement is typically associated with the first conjunct alone (so-called 'first conjunct' agreement), which c-commands the second one (cf. Camacho 2003, van Koppen 2005). Also relevant is the fact that the scenario Haegeman and van Koppen discuss involves agreement in a long-distance / head-complement configuration (which is weaker than agreement in a spec-head configuration, as has been shown in the literature; cf. Guasti & Rizzi 2002).

To sum up, the data just reviewed are meant to show that "C agreement" and "T agreement" do not result from the same φ-feature bundle.

Notice that data like these pose a problem for both FI and the FIACT, for strict identity between two elements entails having the same features. I return to this in section 4.

3.2 Epstein, Kitahara & Seely (2012)

Epstein, Kitahara & Seely (2012) (henceforth, EKS) point out two more scenarios, shown in (12), that appear to violate (8).

(12) a. They like him
b. Whom do they like?

EKS argue that the V *like* in (12a) moves to v* before Transfer, which predicts a derivational crash under (8a). (12b) is very similar, but this time it is the Case feature of *whom* that appears out of the deletion zone, under the assumption that *whom* moves to [Spec, v*P] before the VP is transferred.[4] Obata (2010:43 and ff.) offers a solution to the latter conflict. According to this author, the only features that appear in the edge zone correspond to the phonological features of *whom*, while he abstract Case feature moves to [Spec, VP], as part of the raising process associated to φ-feature inheritance (Chomsky 2007). This is shown in (13):

(13) a. $[_{v*P}$ they v* $[_{VP}$ like who$_{[iQ][i\varphi][Phon/Acc][uCase]}$]]
b. $[_{v*P}$ **who$_{[iQ][Phon/Acc]}$** $[_{v*P}$ they v* $[_{VP}$ **who$_{[i\varphi][uCase]}$** [like who$_{[iQ][i\varphi][Phon/Acc][uCase]}$]]]]

(Obata 2010:45)

The LI *who* in (13) is regarded as a bundle of features, some interpretable ([iQ] and [iφ] at the semantic component, and [Phon/Acc] at the phonological component), some uninterpretable ([uCase]). What Obata (2010) proposes is that the components of *whom* can be split in two chunks: (i) the phonological features associated to [uCase] move to [Spec, v*P] (along with [iQ]), and (ii) the rest

4 As EKS point out, Chomsky (2007:18) refers to situations like that (12b) suggesting that abstract accusative on *whom* is valued (and deleted) within the VP before A-bar movement takes place.

of the features move to [Spec, VP]. This strategy has been explored in the recent literature with interesting resuts (cf. Epstein & Obata 2011, Obata 2010) and does tackle the facts in (12). Nonetheless, for this analysis to be tenable one must entertain the idea that the different features of an LI can split during the derivation, a possibility that raises questions about lexical integrity and the nature of chains (uniformity, the A vs. A-bar status, etc.; cf. Chomsky 1995:261 and ff. for discussion).

3.3 Obata (2010)

Building on findings by Carstens (2005), Obata (2010) reports that Kilega (a Bantu language) displays two patterns of wh-movement. In (14a), the wh-phrase *bíkí* (Eng. what) remains *in situ* and the verb agrees with the subject *bábo bíkulu* (Eng. those women); in (14b), the wh-phrase overtly moves to [Spec, CP] and the verb moves to C, which then agrees with the fronted wh-phrase. As a result of the latter agreement process, two things happen: first, the verb gets the complementizer agreement morpheme; second, the fronted wh-phrase and C share the same noun class number (namely "8").

(14) a. [$_{TP}$ Bábo bíkulu b -á -kás -íl -é mwámí **bíkí** (Kilega)
2that 2women 2SA-A-give-PERF-FV 1chief 8what
mu-mwílo]?
18 -3village
What did those women give the chief in the village?

b. [$_{CP}$ **Bíkí** bi -á -kás -íl -é [$_{TP}$ bábo bíkulu mwámí (Kilega)
8what 8CA-A-give-PERF-FV 2that 2woman 1chef
mu-mwílo]]?
18 -3village
What did those women give the chief in the village?
(Carstens 2005:220, [taken from Obata 2010:64])

According to Carsten's (2005) analysis of (14b), T attracts the wh-phrase while C agrees with it, apparently retaining its φ-features (cf. Carstens 2005 for relevant discussion). In order to get around this FI-challenging situation, Obata (2010:71) argues that there are two types of edge features (EF): (i) pure EFs and (ii) φ-related EFs. The first type of EF, available in English, triggers movement without agreement. The second type of EF is assumed to "contain [uPhi] inside," which results in the relevant Probe triggering movement of a φ-agreeing Goal. What does this buy us? According to Obata (2010), who further assumes that EFs are uninterpre-

table, EFs must disappear to allow convergence after Transfer.[5] Now, since the φ-features are part of the EFs, these are deleted too. Although deletion of EFs may be necessary for convergence (if these features are regarded as uFF, which is not immediately obvious, cf. fn. 5), it is not clear to me how Obata's (2010) proposal solves the problem raised by the Kilega examples. This is so because, under (7)–(8), deletion requires for features to be in the complement domain of a phase. Given that the EFs of C are not inherited, deletion cannot expunge them.

3.4 Interim conclusion

The previous sections have shown that (8), a direct consequence of FI, is threatened by different situations. There are of course technical ways to overcome these shortcomings (incorporate additional projections, invoke feature-splitting, etc.), but the question arises whether the alternatives have a solid grounding.

The next section explores a system that dispenses with FI, not only because of the empirical arguments we have just considered, but also because of its dubious conceptual status. I argue that the alternative, the FIACT, provides a reasonable explanation as for why uFF can appear in the edge of phases and to other questions that concern the interaction of phase and non-phase heads.

4 Feature Inheritance through the Copy Theory of Movement

This section develops the idea that the exceptional status of non-phase heads is a consequence of the FIACT, repeated here for convenience:

(15) *Feature Inheritance as Copying Thesis (FIACT)*
Non-phase heads are (copies of) phase heads

5 This take on EFs is unclear, since EFs are unlikely to align with other *bona fide* features. Unlike other features, EFs do not involve Match, EFs do not involve Valuation, EFs do not delete, and EFs do not provide any interface (PF or LF) instruction (as argued for by Chomsky 2007:11). An anonymous reviewer points out that EFs should indeed delete, for they have no interface interpretation. This observation is well-taken, but given that deletion requires valuation, and EFs have no values (unlike features like number, gender, Case, etc.), it would have to operate in a different fashion. This may all well indicate the theory internal nature of EFs, which is precisely Chomsky's point: EFs are a way of capturing the unbounded nature of Merge, not an actual feature that can be valued and deleted.

Below I review some antecedents of the technical aspects behind the FIACT, as well as some conceptual and empirical arguments in favor of the idea that C and T are the same LI.

4.1 The FIACT and head-to-phrase movement: Some antecedents

To see what the FIACT ultimately advocates for, consider the matrix clause in (16), which is analyzed as in (17) (for expository reasons, I concentrate on the CP phase).

(16) John saw Susan in the cafeteria

(17) a. $[_{CP}\ C_{\varphi}\ [_{v^*P}$ John v* $[_{VP}$ saw Susan in the cafeteria]]]
b. $[_{CP}$ John $C_{\varphi}\ [_{v^*P}$ <John> v* $[_{VP}$ saw Susan in the cafeteria]]]

As can be seen, (17) assumes that T is not involved in the derivation of (16). Instead, C is taken from the lexicon as a unit that contains all the relevant features (agreement, tense, mood, and others) and drives the relevant operations. There is no process of FI, there are no parallel operations, and the subject moves to the subject position (as customary) – the only difference being that it moves to [Spec, CP], not [Spec, TP].[6]

The analysis of (17) assumes no TP projection, which resembles the GB idea that root declarative clauses lack a CP layer.[7] Two independent projections (typically labeled "CP" and "TP") are necessary in non-matrix scenarios, though, for reasons I return to. There are two cases to consider: defective (raising, ECM) and non-defective subordinate clauses. The standard analyses of these clauses are shown in (18):

(18) a. I believe $[_{CP}$ that John likes Susan] φ-complete clause
b. I believe $[_{TP}$ John to like Susan] φ-defective (raising / ECM) clause

6 Noam Chomsky (p.c.) suggests the possibility that C always raises, and that this is precisely that yields a phase (cf. fn. 9). More precisely, he suggests that IM of C is required for φ-feature valuation, which in turn makes it possible for Transfer to take place. Chomsky further observes that the movement process may be related to the C-I requirement to create an A / A-bar distinction, which can only be established in a structural / configurational fashion, namely if C undergoes IM.

7 This idea is reminiscent of Chosmky's (1986) "Vacuous Movement Hypothesis," whereby only TP was projected. Here I am assuming that only CP is projected – under the idea that T(P) = C(P).

At first glance, the example in (18a) is difficult to accommodate under the FIACT, since *that* seems to be the spell-out of C, while *John* moves to [Spec, TP]. From the perspective of the FIACT, however, (18a) obtains as follows: once the v*P has been assembled, *that* is selected from the lexicon, and then undergoes IM leaving a copy that plays the role of what we usually call "T:"

(19) a. EM (C, v*P) = [$_{CP}$ that [$_{v*P}$ John likes Susan]]
b. IM (C, CP) = [$_{CP}$ that [$_{CP}$ John <that> [$_{v*P}$ <John> likes Susan]]]

An interesting question raised by this derivation is why IM of *that* is necessary in embedded contexts, but not in (17). I would like to suggest that the reason is related to economy considerations and the nature of Transfer. In matrix environments, IM is not necessary because one projection suffices to represent all the information encoded by a clause (its type, mood, etc.). Therefore, at the end of the CP phase, the φ-features of C are valued and Transfer targets the whole structure. In embedded contexts, however, Transfer cannot affect the entire syntactic object, and additional structure must be created for successive cyclic movement to take place. This alone, which suffices to capture the distinction between root and embedded clauses, motivates IM of C.[8, 9]

Let us now consider embedded defective environments, where I assume the presence of a φ-defective version of C, following Epstein & Seely (2006) and Ormazabal (1995).[10] The main advantage of this approach is that there is always a

8 There may be additional interpretive reasons to motivate IM of C. Clause typing may require some specific morpheme in non-root contexts, creating a more fine-grained representation (Rizzi 1997).

9 Alternatively, one could argue that IM of the phase head is always necessary for φ-feature valuation purposes. In particular, it could be assumed that the external argument raises due to the unlabeled status of (EA, v*P) (Chomsky 2008), leaving a copy. Since this copy is invisible (only the highest occurrence of a chain is; Chomsky 2000, 2001), C needs to undergo IM again to get the external argument in its search space and assign Case, as in (i)–(ii):

(i) [$_{CP}$ John C [$_{v*P}$ <John> likes Susan]]
(the highest occurrence of *John* is not in the search space of C)

(ii) [$_{CP}$ C [$_{CP}$ John <C> [$_{v*P}$ <John> likes Susan]]]

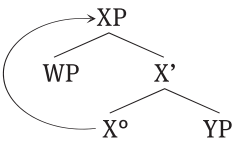
I leave the possibility that IM of the phase head is always mandatory for future research (cf. Epstein et al. 2012 for an alternative view whereby Transfer can only occur if the phase head raises).

10 Arguments against defective clauses to contain a CP layer come from raising-to-subject/object positions and the absence of wh-movement (e.g., *I believe Peter to love Mary – *I believe who Peter to love*). These arguments can be accommodated under the present account if we adopt the φ-defective / φ-complete distinction, which seems to be independently needed. First, if embedded C is φ-defective, then Case would have to be assigned by some matrix Probe. Second, the

C head in embedded scenarios, but there cannot be T without C (*pace* Chomsky 2007).

In sum, I assume that IM of the phase head creates a discontinuous category (i.e., "C" . . . "<C>") that is forced only in embedded clauses. In matrix clauses, this process is optional, unless independent factors demand it (e.g., A-bar movement; cf. fn. 6, 8). It is also important to note that the FIACT is arguing for a rather infrequent movement scenario, whereby a head is internally merged with a phrase, as in (20), where I resort to X-bar notation for presentation purposes:

(20) Internal Merge {X, XP}



Though generally precluded (due to structure-preservation reasons), the configuration in (20) has some precedents in the literature, as an anonymous reviewer points out. In the context of the first formulation of bare phrase structure, Chomsky (1995:320–321) considered the possibilty that v can attach to itself, an option that was dispensed with because it either involved projection of an adjunct (the v head, once moved) or else it created all the problems already present in situations where the target does not project (Chomsky 1995:256 and ff.) – what Hornstein & Uriagereka (2002) call "reprojection."[11] A similar idea is developed in Donati (2006), where heads are argued to project once they reach a would-be specifier position. Donati (2006) makes the interesting observation that this inherent asymmetry of projection can explain the different outcomes of free relatives and embedded interrogatives (cf. (21a) vs. (21b)):

(21) a. I liked [$_{DP}$ what [the student read <what>]] X° movement
b. I know [$_{CP}$ what books [the student read <what books>]] XP movement

lack of wh-movement would follow from the wh-phrase intervening between the relevant upstairs Probe and the external argument (as argued by Chomsky 2004).

11 The proposal considered by Chomsky (1995) is very close to Larson's (1988) analysis of vP shells as involving v movement to empty head positions (a case of substitution). For more general discussion, see Guimarães (2000) and Kayne (2011). The posibility that a head moves to a specifier has also been explored in the Distributed Morphology framework, where a subsequent process of morphological-merger (burying the head within the dominating XP) is assumed to take place.

Donati's (2006) proposal is couched in a projection-based version of Merge, but this should not cloud the fact that her proposal is reminiscent of the FIACT, which reaches the same conclusion without resorting to projections or bar levels. Instead, IM operates displacing the head to the topmost position (as in (22)), where a labeling algorithm (Chomsky 2008, 2013) determines that *what* is the label of the structure:[12]

(22) . . . {what, {C, {the student, {T, {<the student>, {v, {read, <what>}}}}}}}
X° movement

Under this scenario, an LI is internally merged with a phrase, which is not precluded in a theory that dispenses with X-bar theoretic apparatus (labels, bars, spec-head agreement, etc.).[13] The main contribution of the proposal, however, is not that a head moves to the root of the structure (this has already been defended, as we have seen), but that this possibility can be used to capture the close connection between C and T.

To sum up so far, the FIACT takes C and T to be the very same LI expressed in a discontinuous fashion through IM. This has two main consequences. First, it allows us to rationalize the odd theoretical status of non-phase heads (see next section) with the facts. Second, we can explain the fact that features present in phase heads are present in non-phase heads too.

4.2 The odd status of T in the lexicon

Chomsky's (2007, 2008) approach to phases is by and large an effort to sustain the idea that phase heads drive all operations except for External Merge. This leads to a restrictive scenario where phase heads and their concomitant non-phase heads act as a unit for Agree and IM processes. In different places, Chomsky has suggested that properties of T derive from C. This claim is well-known in the case of φ-features:

12 The labeling algorithm introduced in Chomsky (2008) states that whenever two syntactic objects are merged, the most accessible (simplest) one is determined as the label of the entire construction.

13 For similar approaches to this kind of head movement, I refer the reader to Saito (2012), Shimada (2007), Tonoike (2009), Kitahara (2013), Epstein, Kitahara, and Seely (2013).

> On the surface it seems to be T, not C, that is the locus of the φ-features that are involved in the Nominative-agreement system [. . .] There is, however, antecedent reason to suspect otherwise, confirmed (as we will see) by empirical phenomena [. . .] In the lexicon, T lacks these features.
> (Chomsky 2008:143–144)

But it seems that the same holds for iFF, in particular tense features. In this respect, Chomsky (2007) considers two scenarios, one of which highlights the odd (empty) status of T in the lexicon.

> What is true of agreement features appears to hold as well for tense: in clear cases, T has this feature if and only if it is selected by C, though C never (to my knowledge) manifests Tense in the manner of φ-features in some languages. If that is basically accurate, then there are two possibilities. One is that Tense is a property of C, and is inherited by T. The other is that Tense is a property of T, but receives only some residual interpretation unless selected by C (or in other configurations, e.g., in English-like modal constructions). One advantage of the latter option is that T will then have at least some feature in the lexicon, and it is not clear what would be the status of an LI with no features (one of the problems with postulating AGR or other null elements). Another advantage would be an explanation for why C never manifests Tense in the manner of φ-features (if that is correct).
> (Chomsky 2007:20)

Interestingly enough, and contrary to what Chomsky points out, languages like Irish do manifest tense on C. This can be seen in (23), where the shape of C depends on the tense of the embedded clause: *go* for non-past, *gur* for past.[14]

(23) a. Deir sé [$_{CP}$ **go** dtógfaidh sé an peann] (Irish)
say-PRES he C take-FUT he the pen
He says that he will take the pen

b. Deir sé [$_{CP}$ **gur** thóg sé an peann] (Irish)
say-PRES he C-PAST take-PAST he the pen
He says that he took the pen
(Cottell 1995:108)

These data could thus be taken to assume that, just like φ-features, tense is generated in C and then passed down to T. If so, T should be regarded as a feature-less

14 An anonymous reviewer asks whether English also manifests tense on C. In embedded clauses, tense seems not to be spelled-out in C, but in T-to-C scenarios (e.g., interrogatives), the auxiliary / modal could be analyzed as being the exponent of tense morphology (see 4.4. for more discussion). See Pesetsky & Torrego (2001) for an analysis where T moves to C also in embedded clauses.

lexical item, an empty element that exists so that the processes of FI can take place, and nothing else. This is precisely what Chomsky (2007:21) seems to assume, building on ideas by Richards (2007). If T is so strongly dependent on C, then one would expect for T to be selected only if C is. However, as noted above, standard approaches to ECM and raising environments utterly depart from this assumption, a position maintained in Chomsky (2007:21–22).

There are other suspect properties of T in Chomsky's (2007, 2008) phase-based system. Not only is T an empty LI (a mere 'feature receptacle,' as Richards 2007 puts it), it is further incapacitated to trigger operations. According to Chomsky (2007, 2008), both Agree and IM are triggered by T in a parasitic fashion, it is actually C that triggers those operations.[15]

In the next section I present some data that concern the featural bounds between phase and non-phase heads that will enhance the plausibility of the FIACT.

4.3 T-bound features in C and C-bound features in T

Chomsky's (2008) FI is conceived of as a mechanism that removes features from one head (a phase head) in order to give them to another head (a non-phase head). This has been assumed for φ-features (uFF) alone, but the possibility that the same holds for interpretable features remains to be explored in detail (but see section 4.2.). In this section I focus on two scenarios that strengthen the FIACT, namely cases where C shows T-like (L-related or argumental) features and cases where T shows C-like (non-L-related or operator) features.

That C can manifest inflectional features can easily be shown. In (10), (11), and (23) above, we saw that C can express tense and agreement distinctions, which could be taken to indicate that C has these features. The same is true in Spanish, if we take conjunctions and prepositions to be tensed and non-tensed versions of C:

(24) a. Le dije [$_{CP}$ **que** fuésemos al cine] (Spanish)
CL-to.him say-PAST.1.SG that go-3.PL to-the cinema
I told him to go to the cinema

15 Chomsky has offered two empirical arguments in favor of parallel operations. The first one concerns the Subject Condition (Chomsky 2008:147), while the second one comes from intervention effects (Chomsky 2008:152). In both cases, Chomsky argues that the facts can only be explained if all operations occur when C has been introduced in the derivation, not when T is.

b. Le dije [$_{CP}$ **de** ir al cine] (Spanish)
CL-to.him say-PAST.1.SG of go-3.PL to-the cinema
I told him to go to the cinema

Mood inflection is also known to be expressed in the complementizer of languages like Russian, Latin, and Old Salentino. Consider the Russian case (cf. Khomitsevich 2007):

(25) a. Ivan znaet [**čto** Maša ljubit Petra] (Russian)
Ivan know-3.sg that Masha love Petr
Ivan knows that Masha loves Petr
b. Ivan xočet, [**čtoby** Maša ljubila Petra] (Russian)
Ivan want-3.sg that-subj Masha love Petr
Ivan wants Masha to love Petr
(Khomitsevich 2007:8–9)

So far we have discussed cases where inflectional features are spelled-out in C, but evidence gathered from selection shows that such features are there even if they do not have a PF reflex. Mood and tense features are systematically spelled-out in a head lower than C in Romance, but they must also be present in C for selection purposes (cf. Shlonsky 2006): If C were not endowed with (covert) mood and tense features, it is not immediately obvious how the following asymmetries should be accounted for:

(26) a. Juan quiere [$_{CP}$ que María {*viene/venga}] (Spanish)
Juan want-3.SG that María come-IND/SUBJ-3.SG
Juan wants that María comes / for María to come
b. Juan entiende [$_{CP}$ {que tiene/*tener} problemas] (Spanish)
Juan understand-3.SG that have-3.SG/INF problems
Juan understands that he has problems / to have problems

The data in (26) indicate that *querer* (Eng. want) and *entender* (Eng. understand) impose inflectional constraints on the CP they select. If selection is a local process, then C should contain both tense and mood features, even though there is no PF reflex.[16]

16 Alternatively, one can argue that these features are generated in a lower head and then move upstairs through some feature sharing/percolation mechanism (perhaps at LF). Such move would however require postulating two mechanisms that are independent of Merge: labeling

Consider now the flipside of all this: C-like (meaning non-L-related or operator) features being present below C. Of course, such features (dubbed "criterial" by Rizzi 2004) have to be in C in order to capture the selection facts too:

(27) a. *I wonder [$_{CP}$ that Mary left]
b. I wonder [$_{CP}$ whether Mary left]

But at the same time it is an old observation that [Spec, TP] behaves as an A-bar position in Spanish and other Romance languages (cf. Torrego 1984, Uriagereka 1988, Ordóñez 1997, and references therein), which has sometimes taken to indicate that it is a cyclic category (cf. Rizzi 1982). The A-bar nature of this position can be seen in its optional status and in the fact that it yields a topic/categoric interpretation of the displaced subject:

(28) a. Benito Pérez Galdós escribe (Spanish)
Benito Pérez Galdós write-3.SG
Benito Pérez Galdós writes (Benito Pérez Galdós is a writer)
b. Escribe Benito Pérez Galdós (Spanish)
write-3.SG Benito Pérez Galdós
Benito Pérez Galdós writes (Benito Pérez Galdós is writing now)

Also relevant is the fact, noted by Torrego (1984:108–112), that preverbal subjects create a mild intervention effect in wh-movement scenarios. This is shown by the examples in (29):

(29) a. ?*Qué medidas$_i$ quiere Merkel [$_{CP}$ t$_i$ que Zapatero (Spanish)
what measures want-3.SG Merkel that Zapatero
adopte t$_i$]?
adopt-3.SG
What measures does Merkel want Zapatero to adopt?
b. Qué medidas$_i$ quiere Merkel [$_{CP}$ t$_i$ que adopte (Spanish)
what measures want-3.SG Merkel that adopt-3.SG
Zapatero t$_i$]?
Zapatero
What measures does Merkel want Zapatero to adopt?

and percolation. Given the arguments in Cable (2007) and Collins (2002), these operations are to be questioned on both empirical and conceptual grounds.

As can be seen, the preverbal position of the subject interferes with wh-movement, indicating that they have a common A-bar status. Along with these facts, it is also known that [Spec, TP] has *bona fide* A properties in Romance as well, for preverbal subjects can bind and control, which is actually reinforcing the FIACT.

In sum, this section has shown that both A and A-bar information is present in phase and non-phase heads, with parametric choices determining whether they are spelled-out in one head or in more, and if the former in which one (cf. Giorgi & Pianesi 1997).

5 Testing the FIACT: The C – T distinction

Having seen the empirical and theoretical problems of FI and assuming the conclusions reached in Chomsky (1995:262–263) with respect to lexical integrity (which preclude feature movement without pied-piping), we are led to conclude that if one feature is inherited, all features are. Consequently, if *all features* are in both C and T, the question arises how FI is to be understood precisely. Considering facts like (26), Chomsky (2013) qualifies the status of FI:

> [I]nheritance has to be understood as copying, leaving Q in its original position for selection and labeling. The basic argument for inheritance due to Richards is that if valued uninterpretable features remain at the phase head position, the derivation will crash at the next phase. That is not a problem here, because Q is interpretable. For φ-features it may mean that they are deleted or given a phonetic form (as in West Flemish), hence invisible at the next phase.
> (Chomsky 2013:47)

At first glance, this seems to bring back the possibility that there is some *Copy* operation, independent of *Merge*, an option dismissed in Chomsky (2008:158 fn.17).[17] Now, if there is no independent operation Copy, replicating C's features in T, then the question arises how C and T end up having the same features. As

17 Ouali (2008) puts forward the idea that FI can manifest itself in three variants: (i) donate, (ii), keep, and (iii) share. Although incorporating three variants of FI amounts to a considerable enrichment of UG, the share option is very close to what Chomsky (2013) appears to have in mind, as a reviewer points out. Likewise, as the same reviewer observes, a theory with three types of FI (which are empirically needed, according to Ouali 2008) is a more principled theory than a theory with only one type of FI but incorporating specific restrictions to block the other two types. Be that as it may, below I suggest that neither FI nor any of Ouali's (2008) variants are needed if we resort to the Copy Theory of Movement.

I have been arguing all along, a reasonable way out is offered by the FIACT: C and T have the same features because they *are* the same element. At this point it must, nonetheless, be acknowledged that many arguments have been offered in favor of the idea that C and T are LIs of their own right, thus threatening the plausibility of the FIACT; quite interestingly, there are also many arguments suggesting that C and T establish robust morpho-syntactic relationships (cf. discussion in Pesetsky & Torrego 2001). The FIACT suggests that these relationships be strengthened. With some perspective, similar claims have been made with categories like Agr_S / Agr_O (or more specific projections, like Gender and Number). In a nutshell, what I am arguing for is that collapsing C and T is a theoretical option – just like assuming that C is to be regarded as a cluster of different heads (Rizzi 1997, 2004) –, an option with interesting consequences.[18]

As noted at the outset, there are antecedents of the FIACT in the literature. Perhaps the most relevant one is Stowell's (1981) idea that C and T are a 'discontinuous object.'

> We might adopt a suggestion of Y. Aoun (personal communication) to the effect that the complementizer and Infl form a discontinuous element. The matching between complementizers and Infl would then follow from the fact that the two actually form a single unit at some level. So that selection for one implies selection for the other.
> (Stowell 1981:241)

What I want to argue, building on Stowell (1981), is that C and T are indeed a discontinuous element, not because they are related at some abstract level, but because they are two occurrences of the very same LI. In particular, I would like to reinterpret Stowell's proposal in order to argue that FI is to be seen as a process whereby T is a copy (lower occurrence) of C, as shown in (30).

(30) a. $[_{CP}\, C \ldots [_{v^*P} \ldots]\,]$
b. $[_{CP}\, C \ldots [_{CP}\, <C> \ldots [_{v^*P} \ldots]\,]\,]$ (typically, $[_{CP}\, C \ldots [_{TP}\, T \ldots [_{v^*P} \ldots]\,]\,]$)

This approach renders FI redundant and captures the morpho-syntactic connection between phase heads and non-phase heads. However, it also raises

18 One more caveat is in order. The FIACT is not arguing that C and T are to be ontologically or conceptually collapsed, since notions like "(illocutionary) force" and "tense" cannot be reduced to each other. The set of features provided by UG (tense, number, gender, mood, etc.) are indeed ontologically distinct (Chomsky 2000), but their combination gives rise to objects (LIs) that cannot be primitives at the conceptual-intentional level. It is important to clarify this point, for the FIACT could (wrongly) taken to indicate that C and T are literally treated as one and the same conceptual primitive.

questions. Due to space constraints, I would like to focus on what is plausibly the most problematic situation for the FIACT to be regarded as a valid hypothesis, namely the fact that C and T can appear simultaneously in the same clause. The possibility that T and C are (two occurrences of) the same LI seems most plausible in contexts where one of the two heads is not spelled-out – namely, root declarative clauses. Embedded environments provide a different scenario:

(31) John said [$_{CP}$ **that** Mary **will** arrive tomorrow]

The example in (31) is not an isolated fact of English, and the elements in C and T do not have to be necessarily complementizers and auxiliaries – other elements can occupy those positions (particles, clitics, modals, lexical verbs, etc.). Under the Copy Theory of Movement, the FIACT does not predict these facts, for lower occurrences of a non-trivial chain are usually silent, so something else must be said. In fact, there are two scenarios that the FIACT must be able to account for, one of which further involves three suboptions. The scenarios are listed in (32), where I use P and N to indicate phase and non-phase head respectively (Richards 2007):

(32) a. Either P or N is spelled-out
b. Both P and N are spelled-out
i. P and N are identical (morphologically)
ii. P and N are partially identical (morphologically)
iii. P and N are distinct (morphologically)

Option (32a) is the best-case scenario, expected under a Merge-based approach to movement where all copies but one are deleted (typically the highest one, but not necessarily; cf. Stjepanović 1999). The options in (32b) are illustrated by data like the following, and (31) too:

(33) a. [$_{CP}$ **Wen** glaubt Hans [$_{CP}$ **wen** Jakob gesehen hat <wen>]]? (German)
whom think Hans whom Jakob sen has
Whom does Hans think Jakob saw?
b. [$_{CP}$ **Was** glaubt Hans [$_{CP}$ **wen** Jakob <wen> anruft]]? (German)
what think Hans whom Jakob call
Whom does Hans think that Jakob called?
(McDaniel 1986:569, 573)

The first thing to note here is that we are dealing with heads, not phrases – as pointed out in the literature, multiple realization of elements appears to be mor-

phologically restricted to heads (Nunes 2004:38 and ff.). As the reader may see, this adheres to the FIACT, which focuses on the status of phase *heads*. The cases in (33a,b) are well-known in the literature for which different accounts have been provided (Fanselow 2006).[19] Given morphological similarities, situations like the ones in (33a,b) are amenable to a FIACT-based treatment, but (31) is not.[20]

There are at least two possible ways to reasonably account for (31) without abandoning the FIACT. One essentially goes back to Ross' (1969) idea that modals are main verbs. If that route is chosen, then (31) should be analyzed as in (34), *will* being a species of v.

(34) John said [$_{CP}$ **that**$_i$ Peter$_j$ t$_i$ [$_{vP}$ **will** [$_{vP}$ arrive t$_j$ late]]]

A second alternative involves taking *that* and *will* as being part of a complex C head, in a sense reminiscent of the 'big DP' analysis of clitics (Uriagereka 1988, 1995). From such perspective, *that* and *will* are two C heads that start off the derivation together, splitting later on.[21, 22] The fact that *that* and *will* share inflectional features is directly captured under a doubling analysis.

(35) John said [$_{CP}$ **that**$_i$ Peter$_j$ [$_C$ t$_i$ [$_C$ **will**]] [$_{vP}$ arrive t$_j$ late]]]

It must be noticed, however, that standard doubling accounts would take *that* and *will* to be different LIs that start the derivation together, which is a departure from the FIACT. Interestingly for our purposes, doubling clitics have also been analyzed as agreement morphemes, not actual independent heads (Strozer 1976, Suñer 1988, Sportiche 1993), which opens the door for a FIACT-abiding analysis

19 Other relevant cases involve wh-reduplication in child grammar and sign languages, as an anonymous reviewer points out (cf. Crain & Thornton 1998, Nunes & de Quadros 2008).

20 Morphological resemblance seems stronger when the phase head is spelled-out in different cycles. Descriptively, the morphological shape of the copies is more different if they are realized within the same cycle, which may be due to anti-locality or distinctness-abiding factors in Richards' (2010) sense.

21 An analysis very similar to this one is proposed by Pesetsky & Torrego (2001), who treat *that*, *do*, *will*, and other C morphemes as the spell-out of some T head moved to C.

22 Of course, one could also assume that local morpho-phonological rules change the shape of the different occurrences. This analysis seems compatible with the Italian facts discussed by Ledgeway (2005), but it seems to require more assumptions in cases like (31), where the change of shape is more severe.

of (31). From that point of view, the structure in (35) does not have to be created during the derivation, but in the lexicon. To be specific, I assume that Merge can also operate pre-syntactically in order to combine roots and functional morphemes to form LIs (cf. Hale & Keyser 1993, 2002 for discussion). Therefore, *will* in (35) can be analyzed as as sublexical component of C – the spell-out of a tense feature. Assuming this fairly agreed-upon view on lexical items, the FIACT can be kept intact, as *that* and *will* are nothing but the spell-out of different sub-lexical components of C. To see this in more detail, consider (36), where C is analyzed as a cover term for a more complex structure involving force, tense, mood, etc. features (Chomsky 2001):

(36) . . . [$_{CP}$ [$_{that}$ ***that*** [$_{will}$ *will*]] Peter$_j$ [$_{that}$ *that* [$_{will}$ ***will***]] [$_{vP}$ arrive t$_j$ late]]]

The basic assumption in (36) is that the features of the LI "[$_{that}$ *that* [$_{will}$ *will*]]" can be spelled-out in different positions (see section 4.3.), just like sublexical components of LIs are in cases where vocabulary insertion applies in a scattered fashion (unergative verbs, degree words, modals, tense morphemes, etc.; Giorgi & Pianesi 1997, Hale Keyser 2002). From this it follows that the identity between the elements of a non-trivial chain must be expressed featurally, which I will phrase as follows:

(37) *Lexical Identity Lemma* (LIL)
Given a lexical token α, taken from the lexicon, occurrences of α in a non-trivial chain are identical if they contain the same feature structure

Granted, *that* and *will* in (35) are not identical by the LIL, but they are part of an object that is, namely "C."

The analysis in (36) should also apply in Romance languages, where the verb moves to T, in principle occupying the position left by the copy of the phase head. A way to tackle this is to assume that V is internally merged with C when the CP phase starts (cf. Kitahara 2013 for related discussion), without occupying the copy left by this head. After the new unit, {C, V}, is created, C undergoes subsequent movement (a step that resembles excorporation, in the sense of Roberts 1991), as illustrated in (38):

(38) a. {C, {María, {dijo, {la, verdad}}}}
b. {{C, dijo}, {María, {<dijo>, {la, verdad}}}} V MOVEMENT
c. {C, {{<C>, dijo}, {María, {<dijo>, {la, verdad}}}}} C MOVEMENT

The copy left by C functions as what is called T, which can thus check the relevant inflectional features with the verb. This approach does not involve insertion of lexical material over a trace, and it further allows T-to-C movement in Romance (and V2 in germanic languages; cf. Koster 2003, Wiklund et al. 2007, Zwart 1993), assuming this process requires verb movement all the way to C – in such circumstances, all we have to assume is that the {C, V} unit moves as a whole, giving rise to verb pied-piping.

Different versions of the doubling analysis I am adopting have been applied to phenomena like resumption, focalization, reflexivization, or control (cf. Belletti 2005, Boeckx et al. 2007), so the fact that it is available in other corners of grammar should not be surprising. This approach can also accommodate the data provided by Haegeman & van Koppen (2010), which showed, as we saw in section 3.1., that C and T can manifest different φ-features. What is relevant for us is that this kind of mismatches has also been discussed for doubling: in certain varieties of Spanish, clitics can manifest agreement mismatches with their double in number (see (39a)), gender (see (39b)), and person (see (39a)) (Ordóñez 1997, Uriagereka 1995):

(39) a. **Le$_i$ / Les$_i$** dije la verdad **a los alumnos$_i$** (Spanish)
CL-to.him / them said the truth to the students
I told the truth to the students

b. **Le$_i$ / La$_i$** dije la verdad **(a María$_i$)** (Spanish)
CL-him / her said the truth to María
I told Mary the truth) mismatches

c. **Nos / Os / Los** vi **a los profesores$_i$** (Spanish)
CL-1.pl / 2.pl / 3.pl saw to the teachers
I saw we / you / them the teachers

One final piece of evidence I would like to incorporate to the discussion of the FIACT comes from so-called "complementizer reduplication," which has been reported in Romance languages (Uriagereka 1988), but also in old and colloquial English (Higgins 1988, McCloskey 2006)

(40) He thinks [$_{CP}$ **that** if you are in a bilingual classroom [$_{CP}$ **that** you will . . . not be encouraged to learn English
(McCloskey 2006:104)

Although there may be different solutions to account for facts like (40), an analysis that seems fully compatible with the FIACT is Ledgeway's (2005), which is

based on complementizer movement, an approach this author extends to the Italian instances of "complementizer reduplication:"

(41) Et e`y manifesta cosa [$_{CP}$ **che** homo che se ave a defendere a la . . .
. . . patria soa intre li amici e li canussienti suoy [$_{CP}$ **cha** ave a chesta . . .
. . . parte gran prerogativa e gran avantayo]] (Italian)
'And it is abundantly clear that (che), a man who has to defend himself in his own country among his friends and acquaintances, that (cha) he has in this respect considerable privilege and advantage' (LDT 126.2–4)
(Ledgeway 2005:380)

Data like (40) and (41) can be easily accounted for if the FIACT is entertained. Seen this way, complementizer reduplication (or recomplementation) is nothing but another construction where different occurrences of the same LI are spelled-out.

Let us take stock. In the preceding paragraphs we have seen arguments suggesting that FI can be regarded as involving a copying process, an idea I have pushed to support the FIACT. The proposal has the main advantage that it dispenses with an extra UG operation and straightforwardly captures the pervasive observation that C and T have an intimate dependency. Such dependency was explored by Stowell (1981), who suggested that these elements could form a discontinuous element, which the FIACT recasts as C and T forming a non-trivial chain. In the remainder of this section I have concentrated on how this chain can manifest in particular cases, given that there are many scenarios where the copy left by C is spelled-out, and when it is its morpho-phonological realization is not identical to that of the higher occurrence. I have argued that, in those circumstances, the system resorts to a strategy that seems to be available not only in different languages (Spanish, English, Hebrew, etc.) but also in different conditions (focalization, resumption, reflexivization, etc.).

6 Conclusions

The goal of this paper was to investigate an alternative to Chomsky's (2008) FI. On the light of both empirical and conceptual problems posed by this operation, I have argued for an alternative view according to which non-phase heads are copies of phase heads. I have phrased this hypothesis as (42), which adheres to the minimalism desideratum that "while (iterated) Merge comes free, any other operation requires justification" (Chomsky 2001:3).

(42) *Feature Inheritance as Copying Thesis (FIACT)*
Non-phase heads are (copies of) phase heads

If phase and non-phase heads are different occurrences of the same LI, then it is expected that they 'act as a unit' for the purposes of different syntactic operations. It also follows that FI, which revamps Chomsky's (1995) *Move F* approach, can be dispensed with, which simplifies the computational possibilities, and consequently limits the degree of cross-linguistic syntactic variation (Chomsky 2001).

The proposal in (42) is reminiscent of Chomsky's (1986) "Vacuous Movement Hypothesis," and has far-reaching consequences, only some of which have been explored. First and foremost, the FIACT allows us to eliminate the odd status of T (as argued in section 4.2.). Secondly, it also provides a plausible explanation for the morpho-syntactic connection between C and T that has been noted in the literature (cf. Stowell 1981, Bayer 1984, Haegeman 1992, Zwart 1993, Watanabe 2000, Pesetsky & Torrego 2001, Chomsky 2007, Miyagawa 2009, and references therein). Of course, the FIACT also raises questions. One such question concerns the lexicon, and the hypothesis that C and T are the same LI. Here I have limited myself to show that this theoretical option embodied by the FIACT – just like the option of analyzing agreement nodes as an independent LIs (Chomsky 1995) – does not aim at a conceptual collapse between C and T, and has advantages worth exploring, both theoretically and empirically.

There are many phenomena where the FIACT could be relevant in order to provide a different perspective (the left periphery, *do* insertion, *that*-trace effects, etc.), but in this paper attention has been restricted to situations in which both C and T are spelled-out. I have argued that those potentially problematic cases can be made consistent with the FIACT if a doubling-based account (Uriagereka 1988, 1995) is adopted. Future research will determine the relevance of this proposal in order to understand and better characterize the interactions between different functional heads (cf. Hale & Keyser 1993, 2003, Rizzi 1997). Whatever the ultimate take on those facts is, the FIACT states that they do not require postulating additional operations like FI.

Acknowledgments: For comments to previous versions of this paper I would like to thank Luis López and Juan Uriagereka. Thanks especially to Noam Chomsky for comments, constant support, and encouragement. I am also grateful to two anonymous reviewers, whose observations and questions were very useful. Usual disclaimers apply. This research has been partially supported by grants from the Ministerio de Ciencia e Innovación (FFI2011-29440-C03-01), and from the Generalitat de Catalunya (2009SGR-1079).

References

Bayer, Josef. 1984. COMP in Bavarian syntax. *The Linguistic Review* 3. 209–274.
Belletti, Adriana. 2005. Extended doubling and the VP periphery. *Probus* 17. 1–35.
Boeckx, Cedric, Norbert Hornstein & Jairo Nunes. 2007. Overt Copies in Reflexive and Control Structures: A Movement Analysis. *UMDWPL* 15. 1–46.
Bošković, Željko. 2007. Don't feed your movements: Object shift in Icelandic. *Working Papers in Scandinavian Syntax* 80. 1–15.
Bresnan, Joan. 1970. On complementizers: toward a syntactic theory of complement types. *Foundations of Language* 6. 297–321.
Cable, Seth. 2007. The *grammar of Q. Q-particles and the nature of wh-fronting, as revealed by the wh-questions of Tlingit*. Cambridge, MA: MIT dissertation.
Camacho, José. 2003. *The structure of coordination: Conjunction and agreement phenomena in Spanish and other languages*. Dordrecht: Kluwer.
Carstens, Vicki. 2003. Rethinking complementizer agreement: Agree with a case-checked Goal. *Linguistic Inquiry* 34. 393–412.
Carstens, Vicki. 2005. Agree and EPP in Bantu. *Natural Language and Linguistic Theory* 23. 219–279.
Crain, Stephen & Rosalind Thornton. 1998. *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge: MIT Press.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereja (eds.), *Step by step. Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond. The cartography of syntactic structures (vol. 3)*, 104–131. Oxford, NY: Oxford University Press.
Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland and Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, 1–30. Berlin: Mouton de Gruyter.
Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos Otero & M. Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory. Essays in honor of Jean-Roger Vergnaud*, 134–166. Cambridge, MA: MIT Press.
Chosmky, Noam. 2013. Problems of Projection. *Lingua* 130: 33–49.
Collins, Chris. 2002. Eliminating labels. In S. Epstein and T. Seely (eds.), *Derivation and explanation in the minimalist program*, 106–132. Malden, MA: Blackwell.
Cottell, Siobhan. 1995. The representation of tense in modern Irish. *GenGenP* 3. 105–124.
Donati, Caterina. 2006. On wh-head movement. In Lisa Cheng & Norbert Corver (eds.), *Wh-movement: moving on*, 21–46. Cambridge, MA: MIT Press.
Epstein, Samuel. 1998. Overt Scope Marking and Covert Verb-Second. *Linguistic Inquiry* 29. 181–227.
Epstein, Samuel. 1999. Un-principled syntax and the derivation of syntactic relations. In Samuel Epstein & Norbert Hornstein (eds.), *Working minimalism*, 317–345. Cambridge, MA: MIT Press.

Epstein, Samuel & T. Daniel Seely. 2002. Rule applications as cycles in a level-free syntax. In Samuel Epstein & T. Daniel Seely (eds.), *Derivation and explanation in the minimalist program*, 65–89. Malden, MA: Blackwell.

Epstein, Samuel, Hisatsugu Kitahara & T. Daniel Seely. 2012. Structure Building That Can't Be! In Myriam Uribe-Etxebarria & Vidal Valmala (eds.), *Ways of Structure Building*. Oxford: Oxford University Press.

Epstein, Samuel, Hisatsugu Kitahara & T. Daniel Seely. Simplest Merge generates set intersection: Implications for complementizer 'trace' explanation. Proceedings of Glow in Asia, Mie University, 5 September 2012.

Epstein, Samuel & T. Daniel Seely. 2006. *Derivations in Minimalism: Exploring the Elimination of A-Chains and the EPP*. Cambridge: Cambridge University Press.

Fanselow, Gisbert. 2006. Partial wh-movement. In Martin Everaert, Henk Van Riemsdijk, Rob Goedemans & Bart Hollebrandse (eds.), *The Blackwell Companion to Syntax*, 437–492. London: Blackwell.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT Press.

Gallego, Ángel J. 2010. *Phase Theory*. Amsterdam: John Benjamins.

Gallego, Ángel J. (ed.). 2012. *Phases: Developing the Framework*. Berlin: Mouton de Gruyter.

Giorgi, Alessandra & Fabio Pianesi. 1997. *Tense and aspect: From semantics to morphosyntax*. Oxford: Oxford University Press.

Guasti, Maria Teresa & Luigi Rizzi. 2002. Agreement and tense as distinct syntactic positions: Evidence from acquisition. In Guglielmo Cinque (ed.), *The structure of DP and IP*, 167–194. Oxford: OUP.

Guimarães, Maximiliano. 2000. In defense of vacuous projections in bare phrase structure. *UMD WPL 9*. 90–115.

Haegeman, Liliane. 1992. *Theory and description in generative syntax. A case study in West Flemish*. Cambridge: Cambridge University Press.

Haegeman, Liliane and Marjo van Koppen. 2010. Complementizer agreement and the relation between T° and C°. Ms., Ghent University, GIST / University of Utrecth.

Hale, Kenneth & Samuel J. Keyser. 1993. On the argument structure and the lexical expression of syntactic relations. In Hale, Kenneth & Samuel J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger,* 53–109. Cambridge, MA: MIT Press.

Hale, Kenneth & Samuel J. Keyser. 2000. *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.

Higgins, Roger. 1988. Where the Old English Sentence Begins. Ms., University of Massachusetts, Amherst.

Hornstein, Norbert & Juan Uriagereka. 2002. Reprojections. In Samuel Epstein & T. Daniel Seely (eds.), *Derivation and explanation in the minimalist program*, 106–132. Malden, MA: Blackwell.

Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9. 577–636.

Kayne, Richard S. 2011. Antisymmetry and the Lexicon. In Anne-Marie Di Sciullo & Cedric Boeckx (eds.), *The Biolinguistic Entreprise*, 329–353. Oxford: Oxford University Press.

Khomitsevich, Olga. 2007. *Dependencies across Phases. From Sequence of Tense to Restrictions on Movement*. Utrecht, Universiteit Utrecht dissertation.

Kitahara, Hisatsugu. 2013. Minimization of Universal Grammar, Reliance on Third Factor Principles, and Feasibility of Inquiry into Evolutionary Origins. *Reports of the Keio Institute of Cultural and Linguistic Studies* 44. 269–280.

Koppen, Marjo van. 2005. *One Probe – Two Goals: Aspects of agreement in Dutch dialects*. Leiden, Leiden University dissertation.
Koster, Jan. 2003. All languages are tense second. Ms., University of Groningen.
Larson, Richard. 1988. The double object construction. *Linguistic Inquiry* 19. 335–391.
Lasnik, Howard. 1999. *Minimalist analysis*. Oxford: Blackwell.
Ledgeway, Adam. 2005. Moving through the left periphery: The dual complementiser system in the dialects of southern Italy. *Transactions of the Philological Society* 103. 339–396.
Legate, Julie Anne. 2003. Some interface properties of the phase. *Linguistic Inquiry* 34. 506–516.
McDaniel, Dana. 1989. Partial and Multiple Wh-Movement. *Natural Language and Linguistic Theory* 7. 565–604.
McCloskey, James. 2006. Questions and questioning in a local English. In Rafaella Zanuttini, Héctor Campos, Elena Herburger & Paul Portner (eds.), *Negation, Tense and Clausal Architecture: Cross-linguistic Investigations*, 86–126. Georgetown: Georgetown University Press.
Miyagawa, Shigeru. 2009. *Why Agree? Why Move? Unifying Agreement-Based and Discourse Configurational Languages*. Cambridge, MA: MIT Press.
Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.
Nunes, Jairo & Ronice Müller de Quadros. 2008. Phonetically Realized Traces in American Sign Language and Brazilian Sign Language. In Josep Quer (ed.), *Selected papers from TISLR 2004*, Seedorf: Signum.
Obata, Miki. 2010. *Root, Successive-Cyclic and Feature-Splitting Internal Merge: Implications for Feature-Inheritance and Transfer*. Michigan, University of Michigan dissertation.
Obata, Miki & Samuel Epstein. 2011. Feature-Splitting Internal Merge: Improper Movement, Intervention and the A/A′-Distinction. *Syntax* 14. 122–147.
Obata, Miki & Samuel Epstein. 2012. Feature-Splitting Internal Merge: The Case of Tough-Constructions. In Myriam Uribe-Etxebarria & Vidal Valmala (eds.), *Ways of Structure Building*. Oxford: Oxford University Press.
Ordóñez, Francisco. 1997. *Word order and clause structure in spanish and other Romance languages*. New York, CUNY dissertation.
Ormazabal, Javier. 1995. The *syntax of* complementation: *On the connection of syntactic structure and selection*. Connecticut, University of Connecticut dissertation.
Ouali, Hamid. 2008. On C-to-T φ-feature transfer. In Roberta D'Alessandro, Susann Fischer & Gunnar Hrafn Hrafnbjargarson (eds.), *Agreement restrictions*, 159–180. Berlin: Mouton de Gruyter.
Ott, Dennis. 2011. A Note of Free Relative Clauses in the Theory of Phases. *Linguistic Inquiry* 42. 183–192.
Pesetsky, David & Esther Torrego. 2001. T-to-C movement: Causes and consequences. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 355–426. Cambridge, MA: MIT Press.
Richards, Marc D. 2007. On feature inheritance: An argument from the phase ompenetrability condition. *Linguistic Inquiry* 38. 563–572.
Richards, Norvin. 2010. *Uttering Trees*. Cambridge, MA: MIT Press.
Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar. Handbook in generative syntax*, 281–337. Dordrecht: Kluwer.

Roberts, Ian G. 1991. Excorporation and Minimality. *Linguistic Inquiry* 22. 209–218.
Roberts, Ian G. 2010. *Agreement and head movement: Clitics, incorporation and defective goals*. Cambridge, MA: MIT Press.
Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.
Ross, John Ross. 1969. Auxiliaries as main verbs. In William Todd (ed.), *Studies in philosophical linguistics* (Series 1), 77–102. Evanston, IL: Great Expectations Press.
Saito, Mamoru. 2012. Case Checking/Valuation in Japanese: Move, Agree, or Merge? *Nanzan Linguistics* 8. 109–127.
Shimada, Junri. 2007. Head Movement, Binding Theory, and Phrase Structure. Ms., MIT.
Shlonsky, Ur. 2006. Extended projection and CP cartography. *Nouveaux cahiers de linguistique française* 27. 83–93.
Stepanov, Arthur. 2001. *Cyclic domains in syntactic theory*. Connecticut, University of Connecticut dissertation.
Stjepanović, Sandra. 1999. *What do second position clitization, scrambling and multiple wh-fronting have in common?* Connecticut, University of Connecticut dissertation.
Stowell, Timothy. 1981. *The origins of phrase structure*. Cambridge, MA. MIT dissertation.
Sportiche, Dominique. 1993. Clitic Constructions. Ms., UCLA.
Strozer, Judith. 1976. *Clitics in Spanish*. Los Angeles, CA. UCLA dissertation.
Suñer, Margarita. 1988. The Role of Agreement in Clitic-Doubled Constructions. *Natural Language and Linguistic Theory* 6. 391–434.
Tonoike, Shigeo. 2009. Minimalist Program. *Research areas in linguistics* 1, Asakura Shoten.
Torrego, Esther. 1984. On inversion in spanish and some of its effects. *Linguistic Inquiry* 15. 103–129.
Uriagereka, Juan. 1988. *On government*. Connecticut, University of Connecticut dissertation.
Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26. 79–123.
Watanabe, Akira. 2000. Feature copying and binding. *Syntax* 3. 159–181.
Wiklund, Anna-Lena, Gunnar Hrafnbjargarson, Kristine Bentzen & Þorbjörg Hróarsdóttir. 2007. Rethinking Scandinavian Verb Movement. *The Journal of Comparative Germanic Linguistics* 10. 203–233.
Zwart, Jan Wouter. 1993. *Dutch syntax: A minimalist approach*. Groningen, University of Groningen dissertation.