Abstract

Recent proposals on clausal phases show that locality restrictions are obviated when the clause’s subject has certain syntactic and/or discourse properties. These obviation effects are taken as evidence that phasehood can be derived. Contrary to these proposals, I argue that phasehood does not need to be derived, presenting alternative analyses that are formally simpler and less stipulative. I also present similar obviation effects in nominals in English, expanding the empirical domain. The nominal data also constitute a novel argument for treating nominals as phases.

1 Introduction

Phases have played an important role in theories of locality since they were proposed in the late 1990s. A conventional assumption holds that phasehood is a property inherent to certain lexical items or their maximal projections. For instance, Chomsky (2000) suggests that phases are characterized by “propositional[ity]”; this definition identifies CPs and vPs to be phases. However, in recent work, Grano & Lasnik (2018) and Barros & Frank (2017) have challenged this view, essentially arguing that phasehood is derived and not necessarily an inherent property of certain syntactic categories. Evidence for this claim comes from what I will call “phase obviation effects” where a clause fails to behave like a phase. Setting aside technical differences, both proposals claim that clausal constituents may enter a derivation as non-phases, only becoming phases under specific conditions related to the subject of the clause.

These analyses of phase obviation effects pose interesting theoretical problems. For instance, while both Grano & Lasnik and Barros & Frank propose conditions under which a non-phase becomes a phase, these proposals in turn raise non-trivial questions about
phases: What is in the nature of these non-phases that allows them to turn into phases in a derivation? Why do these non-phases have such a property in the first place? What exactly is the role of the subject in turning non-phases into phases? These proposals also leave open the question of whether similar obviation effects can be observed in other domains. If such effects exist, they would constitute an argument that these domains are also phases/locality domains in the same way that clauses are.

This paper has two goals. The first goal is to simplify the analysis of these phase obviation effects. I identify and evaluate three alternative analyses of phase obviation effects, two of which build on independently-motivated principles, namely, the principle of Value-Transfer simultaneity (M. Richards 2007) and the Principle of Minimal Compliance (N. Richards 1997, 1998, 2001). The second is to provide additional motivation for the kind of phase theory proposed by Grano & Lasnik, Barros & Frank and in this paper, by pointing to similar effects in the nominal domain for gapping and wh-movement in English. I will further show that an extension of the analysis to the nominal data constitutes a novel argument for analyzing (definite) DPs as phases.

This paper is organized as follows. In Section 2, I describe the data that motivate Grano & Lasnik and Barros & Frank's proposals, and offer a critique of the proposals. In Section 3, I lay out the three analyses, ultimately arguing in favor of an analysis where locality effects are related to the valuation of \( \phi \)-features on a phase head. I argue that the data do not require that phases be derived in the course of a derivation, pace Grano & Lasnik and Barros & Frank. Consequently, the data are formally compatible with a simpler, more conventional set of assumptions about phases. In Section 4, I extend the proposal to gapping and wh-movement in the context of nominals. I argue that these have clear parallels with the clausal data, showing a bound possessor effect and what I will call a demonstrative effect. These effects can be accounted for if definite DPs are phases like CPs are. Finally, in Section 5, I discuss further implications and open questions before concluding.

2 Phase obviation effects

2.1 Empirical motivation

Grano & Lasnik and Barros & Frank begin with the observation that many dependencies in English are clause-bound, such as comparative deletion and too/enough constructions, illustrated below: compare (1a) and (1b), for instance. This clause-bound restriction is relaxed when the embedded clause (bracketed below) contains certain kinds of subject, for example, a bound pronoun subject (1c).

(1) Comparative deletion (Grano & Lasnik 2018: 472 exx. 9-11) (Struck-out material is silent but interpreted.)

a. More teachers gave the students pencils than gave the students pens.
b. *More teachers claimed that the principal gave the students pencils than claimed
c. ?More teachers claimed that they gave the students pencils than claimed [that they gave the students pens].

(2) Too/enough construction (ibid.)

a. This book is too valuable [for James to lend _ to Bill].

b. *This book is too valuable [for James to claim that Mark lent _ to Bill].

c. ?This book is too valuable [for James to claim that he lent _ to Bill].

Making the reasonable assumption that these clause-bound restrictions follow from the Phase Impenetrability Condition (PIC), Grano & Lasnik and Barros & Frank conclude that the relative acceptability of (1c) and (2c) indicates the absence of a phase boundary. In other words, in these examples, the clause is not a phase, contrary to standard assumptions. This obviation effect is also apparently restricted to subject positions; it is absent when the object (3a) or a possessor inside a subject is a bound pronoun (3b). I refer readers to Grano & Lasnik's paper for experimental verification of these contrasts.

(3) a. *More teachers claimed that the principal gave them pencils than [claimed that the principal gave them pens].

b. *More teachers claimed that their assistants gave the students pencils than [claimed that their assistants gave the students pens].

Grano & Lasnik and Barros & Frank differ in their analysis of these effects; key points of their proposals are summarized below in (4) and (5). These differences reflect the kind of data they were examining: Grano & Lasnik were primarily concerned with examples with bound pronoun subjects, while Barros & Frank were interested in a wider range of subjects that they collectively refer to with the label “non-shifty subjects.” I further note that Barros & Frank make clear that they are not committed to a specific syntactic analysis; they present the analysis summarized in (5) only for concreteness.

(4) Summary of Grano & Lasnik's proposal

1 A bibliographic note is in order: An earlier proposal about bound pronouns obviating locality restrictions can be found in Boyd 1992. Anticipating the phase-based approach of Grano & Lasnik, Boyd also proposes that the domain for the “binding of a wh-[race]” in an embedded clause is expanded when its AGR node is coindexed with the AGR of a higher clause, forming an “AGR-chain”; since AGR is assumed to be coindexed with its subject, this proposal entails that the domain is expanded when the embedded subject is bound (p. 51). It is interesting that domain expansion is stated in binding-theoretic terms, and not in terms of bounding nodes or barriers. As far as I can tell, Boyd does not discuss the impact of (non-)shifty subjects on domains. I will not discuss Boyd’s proposal in detail here: Grano & Lasnik’s proposal is no less well-motivated, and, importantly, framed in terms of phases, which I am interested in exploring in this paper.
a. A bound pronoun can enter a derivation with unvalued $\phi$-features; these unvalued features can get valued by a matrix binder (Kratzer 1998, 2009; Rullmann 2004; Heim 2008; Landau 2016).
b. When such a bound pronoun appears in a subject position, T fails to get its own $\phi$-features valued via subject–verb agreement.
c. C enters the derivation as a “candidate phase head” (following Felser 2004).
d. CP becomes a phase only if the $\phi$-features on T, the head of C’s complement, are valued.

(5) Summary of Barros & Frank’s proposal
a. Certain subjects, like R-expressions, are “shifty” such that “the embedded subject is not co-construed with the center of attention of the matrix clause” (p. 4, cf. Grosz et al. 1995). It is implicitly assumed that this discourse property has a syntactic reflex.
b. A functional head in the left periphery, Shift (Frascarelli 2007; Frascarelli & Hinterhölzl 2007), agrees with the subject’s “shift” feature.
c. When Shift agrees with a shifty subject, Shift’s clausal complement becomes a phase.

Despite technical differences, both proposals are similar in positing that (i) phasehood is a property that can be acquired in a derivation, (ii) there is an indirect relationship between the phasehood of a clause and the clause’s subject. This relationship is mediated by a functional head: T for Grano & Lasnik; Shift for Barros & Frank. It is the syntactic status of the functional head that determines whether a clause becomes a phase or not, which in turn determines whether the Phase Impenetrability Condition holds.

2.3 Three conceptual problems

Grano & Lasnik and Barros & Frank point out important data that bear on our theories of locality. They also show us how a system consisting of non-phases/candidate phases and phases can work. However, it seems to me that there are some open questions about a number of their assumptions. Broadly speaking, these questions can be grouped into three distinct (but related) problems.

The first is an ontological problem: from a phase theory perspective, how many classes of syntactic objects are there? Logically speaking, we need at least two: non-phases and phases. Barros & Frank implicitly make such an assumption in their proposal. In contrast, for Grano & Lasnik, who follow Felser (2004), there are three: candidate phases, phases, and an “elsewhere” class, which I will call “non-phases.” Candidate phases appear to be a hybrid: they are effectively non-phases that are distinguished by having the ability to become phases under the right conditions, so that the PIC can apply. While Barros & Frank have the more parsimonious (and hence desirable) assumption, their account is also less explanatory in that it does not account for why only some, but not all, non-phases
become phases.

The second problem is what I will call the indirect mechanism problem. In both proposals, a functional head agrees with the subject: for Grano & Lasnik, it is T, while Barros & Frank suggests that this head is Shift. This agreement process in turn has consequences for the phasal status of another constituent: CP for Grano & Lasnik, Shift’s clausal complement for Barros & Frank.

It is unclear, though, why the mechanism by which a constituent becomes a phase should be mediated by another head: why does T or Shift matter? As far as I can tell, this has to be stipulated in both accounts. Barros & Frank assert that Shift has the effect of turning its complement into a phase. Grano & Lasnik suggest that T is privileged because TP is C’s complement, but one could argue that this simply replaces the stipulation that T is privileged with the assertion that the head–complement relation is privileged, compared to other kinds of structural relations. Further, this argument rests on the presupposition that TP is C’s complement, which can be challenged. Assuming some variant of the split CP hypothesis of Rizzi (1997), it is plausible that C does not subcategorize for a TP; C might subcategorize for some other functional head, which in turn contains a TP. If a Rizzi-style analysis of clauses is correct, then there is no clear theoretical motivation for why CP should be sensitive to T’s syntactic properties.

Even if one were to resolve the indirect mechanism problem satisfactorily, one would still have to address a third problem: spelling out exactly what changes when a non-phase / candidate phase becomes a phase. In both accounts, the intuition presented is along the lines of “a functional head agrees with the subject.” While this is important, it is arguably unsatisfying. Ideally, a more explanatory, “deeper” account would explain why non-phases/candidate phases behave the way they do, for example, by connecting it to some broader representational or derivational principle. It seems fairly challenging to develop a empirically-supported account without stipulations.

3 Toward a simpler system

As I see it, an ideal theory of phases should resolve the first two problems in the following way: For the first (ontological) problem, parsimony leads us to favor the bare minimum, namely, the existence of two types of objects, phases and non-phases. As for the second (indirect mechanism) problem, one would ideally also have the relationship between phasehood and bound subjects be established directly between the head of the phase and the subject, rather than via another head like Shift or T.

It is more challenging to characterize a satisfactory solution to the third problem – specifying what changes when a non-phase/candidate phase becomes a phase and why these changes occur. However, I note that this problem is premised upon the assumption that that standard versions of the PIC are correct. In other words, this problem presupposes that syntactic operations are sensitive to the mere presence of a phase head. There is no a priori reason to think that way, though. I would therefore like to argue that we should take
a more careful look at how the PIC is defined: under what conditions do phases or phase heads exhibit the locality effects conventionally attributed to phase “impenetrability”? This re-framing expands the range of potential solutions. It also means that one can retain the conventional notion that a functional head enters the derivation as a phase head, *pace* Grano & Lasnik and Barros & Frank.

In this section, I present three proposals that address satisfy these criteria and evaluate them on empirical and conceptual grounds.

### 3.1 Proposal 1: a feature valuation account

In this first proposal ([6]) phase heads are necessary but not sufficient to give rise to the locality effects mentioned above. This outcome is achieved by positing features on C and an alternative version of the PIC.

(6) a. Phase heads can enter the derivation with unvalued $\phi$-features (cf. Chomsky 2008, M. Richards 2007), as can bound pronouns.
   b. A phase head agrees with the highest DP in its c-command domain for valuation: the subject.
   c. When the highest DP lacks valued $\phi$-features (e.g. certain bound pronouns), the phase head’s $\phi$-features remain unvalued.
   d. PIC: Syntactic operations may not cross a phase head with valued $\phi$-features.

To see how the proposal works, consider the contrast in (7). Following Grano & Lasnik, I assume that in *too/enough* constructions, there is a null operator that moves in a single step to the left edge of the non-finite clause. (This assumption also entails the $vP$ cannot be a phase.) In (7a)/(8a), C’s $\phi$-features get valued by agreement with *Mark*, indicated with the subscripted $\phi$. Movement across C therefore violates the PIC. In contrast, in (7b)/(8b), the bound pronoun enters the derivation with unvalued $\phi$-features, so C’s features do not get valued. Movement across the CP does not violate the PIC and so the resulting sentence is acceptable.

(7) a. *This book is too valuable [Op$_1$ for James to claim that Mark lent t$_1$ to Bill].
   b. ?This book is too valuable [Op$_1$ for James$_2$ to claim that he$_2$ lent t$_1$ to Bill].

(8) Note: $\phi$: valued $\phi$-features. *Italics*: bound pronoun with unvalued $\phi$-features
This proposal is partly empirically motivated. Barros & Frank observe that a similar phase obviation effect obtains when the embedded clause is an existential construction. This effect is not predicted by Grano & Lasnik’s account, which claims that a CP becomes a phase when T’s $\phi$-features are valued. In existential constructions, T’s $\phi$-features do get valued – via the post-copular NP – yet the CP does not behave like it is a phase.

(9) a. **Gapping**
Jill claimed that there was a problem with the heating, and Sally claimed there was a problem with the climate control in general.

b. **Comparative deletion**
More people claimed that there was a problem with the economy than claimed there was a problem with illegal immigration. (Barros & Frank 2017:9–10, exx. 21d and 23d)

To the extent that this existential *there* effect is also a case of phase obviation, it shows that locality effects cannot be linked to the syntactic status of T. One could follow Barros & Frank in assuming that what matters for locality is a discourse-related feature like shiftiness. However, it is also possible to accommodate this effect with a $\phi$-feature valuation approach (I return to this issue in Section 5.1). More specifically, I assume that the expletive subject *there*, like bound pronouns in Grano & Lasnik’s account, lacks valued $\phi$-features (contra Deal 2009, e.g.); this assumption about *there* can help explain why the finite verb in existential constructions exceptionally agrees with the post-copular NP and not with *there*, the subject. Because of the $\phi$-feature properties of expletive *there*, C’s $\phi$-features remain unvalued after C agrees with *there*. By (6d) syntactic operations may target C’s complement.

Also required is an assumption that unvalued features do not necessarily crash a syntactic derivation. To the extent that unvalued features are uninterpretable (e.g. Chomsky 2001), this assumption states that uninterpretable features do not always crash a derivation (see Preminger 2014 for arguments to that effect). In the examples above, it is unclear what else in the sentence might value C’s $\phi$-features; it is conceivable that C’s $\phi$-features remain
unvalued. Without such an assumption, one would incorrectly predict the examples to be ill-formed and unacceptable.

**Evaluation of Proposal 1**

This proposal satisfies the criteria set out in the start of this section. By redefining the PIC in terms of phase heads with valued $\phi$-features, we only require two types of syntactic objects, the bare minimum: non-phase heads and phase heads, pace Grano & Lasnik (2018) and Felser (2004). While this particular statement of the PIC might seem more complicated than the one(s) that Grano & Lasnik and Barros & Frank use, this proposal is arguably formally simpler in that it uses only the PIC, a necessary component of any theory of phases, and does not appeal to additional theoretical devices, such as candidate phases, to account for phase obviation effects. Since phasehood is not derived in this proposal, there is in turn no need for an explanatory account of how a candidate phase / non-phase becomes a phase.

In addition, the proposal postulates a direct relation (specifically, an agreement relation) between the phase head and the subject. It therefore eliminates the need to stipulate that CP’s phasal status is dependent on some functional head inside or outside CP. It also offers a straightforward account of why the obviation of locality constraints occurs only when the bound pronoun is the subject inside the CP. There is no bound pronoun object effect, for instance, because it would require C to bypass the subject to agree with the object, which one can plausibly assume to be a violation of Relativized Minimality. Nor is there a bound possessor effect; that would require C to agree not with the subject, but with a possessor inside it, which is farther away in structural terms.

**3.2 Proposal 2: a “delayed transfer” account**

I note that Proposal 1 resembles in several ways M. Richards’ 2007 proposal. In this section, I discuss the feasibility of re-framing Proposal 1 by building on the intuition behind his principle of Value–Transfer simultaneity.$^2$

M. Richards assumes that feature valuation and transfer only happen when phase heads are merged; locality effects arise as a consequence of transfer. More specifically, all uninterpretable $\phi$-features are thought to originate on phase heads. He further argues for a derivational principle dubbed Value–Transfer simultaneity (“VTS”): once C’s $\phi$-features are valued, they must be transferred to the interfaces right away. The argument for VTS is an argument by elimination: If valuation happens before transfer, the semantic component will encounter uninterpretable features that are valued, which it cannot remove because

---

$^2$In subsequent work, M. Richards develops a theory of phases based on the idea that phases are lexical sub-arrays, as originally proposed by Chomsky (2000). I will not discuss ways to re-frame Proposal 1 with this lexical sub-array theory, because of basic differences between these two approaches. Proposal 1 assumes that phase heads are primitives, while, as M. Richards points out in his 2011 paper, the lexical sub-array theory of phases does not attach any theoretical significance to these heads nor to feature valuation.
they are indistinguishable from valued interpretable features. If transfer happens before valuation, both interfaces will encounter (unvalued) uninterpretable features. In both scenarios, the derivation crashes. The logical conclusion is that valuation and transfer take place simultaneously.

He also gives a conceptual argument for the PIC: When C’s features are valued, C cannot be transferred. If C were transferred, there would no structure available for subsequent Merge operations. To satisfy VTS and the need for some structure, C’s features are first inherited by heads in its domain before they are valued (Chomsky 2008). When valuation happens, the domain is transferred, rendering it inaccessible to syntactic operations.

Proposal 1 and M. Richards’ proposal are thus similar in at least three ways. First, both posit the existence of only phases and non-phases, eliminating the need for candidate phases. Second, they posit features on phase heads, in particular, C. Third, both link locality effects to the valuation of C’s features.

In light of these similarities, it is worth trying to re-frame Proposal 1 using M. Richards’ proposal, which might let us derive phase obviation effects from more basic derivational principles. Following M. Richards, suppose that locality effects are due to feature inheritance and transfer. Building on his proposal, I posit the principle in (10) as a corollary of VTS 3

By hypothesis, delaying transfer entails that the phase — more accurately, the complement of the phase head — remains accessible to higher syntactic operations.

(10) Unvalued features of a phase head delay transfer.

Evaluation of Proposal 2

Compared to Proposal 1, Proposal 2 is attractive in that it tries to derive the effects of the PIC in (6) from a more basic principle of syntactic derivation, namely (10). It also does so without positing that phasehood is acquired in the course of a derivation. However, a closer inspection shows that this account runs into empirical and conceptual challenges.

First, it does not correctly predict the existential there effect. As noted above (see also Barros & Frank), T in the existential construction shows φ-feature agreement with the post-copular NP. Under Proposal 2 and the original VTS proposal that inspired it, these φ-features are inherited from C, so it is C’s features that are valued. If so, one predicts incorrectly that the CP should behave like a phase. Introducing (10) into this framework does not resolve the problem either, since one retains the original assumption that T’s features are inherited from C.

The solution to this problem is to assume that T’s features are distinct from C’s, a point independently made by Haegeman & van Koppen (2012) with data from West Germanic dialects. However, making this assumption weakens the argument for feature inheritance.

Further, if T’s features are distinct from C’s, and VTS holds, then T’s complement should

---

3 This principle has another precedent in a suggestion by Wurmbrand (2011:69): “[an unvalued interpretable feature] in a potential phase projection postpones transfer” (her emphasis).
get transferred as soon as T’s φ-features get valued. This would incorrectly rule out operations that move syntactic objects from T’s complement to a position above TP, like wh-movement of objects.

Lastly, as discussed in Section [3.3], a successful analysis of the existential there effect with Proposal 1 assumes that unvalued features do not crash derivations. This assumption is incompatible with Proposal 2. While unvalued features might delay transfer, they presumably cannot do so forever; and when transfer does happen, the unvalued features should crash the derivation, by hypothesis. This predicts that the existential constructions in question should be unacceptable, contrary to fact. I therefore conclude that there are significant obstacles to reducing Proposal 1 to the principle in [10].

3.3 Proposal 3: a Principle of Minimal Compliance account

In this section, I consider an analysis based on the Principle of Minimal Compliance (N. Richards [1997, 1998, 2001]). The PMC formalizes the intuition that a constraint, once satisfied by a particular syntactic operation applying in a certain domain, can be violated by subsequent operations applying in the same domain.

The PMC is used to account for a fairly diverse set of phenomena, but what is of interest for this paper is N. Richards’ discussion of a contrast involving wh-movement out of wh-islands (11) (N. Richards 2001:274–275). Adapting Boyd 1992, he notes that there is a dependency between the embedded (non-finite) T and the (finite) matrix T in (11a). He suggests that forming the dependency between the two Ts “extends the domain,” making wh-movement out of the wh-island possible, although he does not specify how the domain is extended or what constraint is satisfied so that wh-movement becomes possible under the PMC.

(11) a. What were you wondering how to fix t1?
   b. *What were you wondering how Mary fixed t1?
   (N. Richards 2001:274, ex. 140)

This argument assumes that [11a] contains a tense dependency and [11b] does not, but it is not clear that this assumption is valid. It is true that there is a tense dependency in [11a] Stowell (1982), Wurmbrand (2014), and others have observed that in non-finite clauses like the one in (11a), the embedded clause is interpreted as being in the future of the time denoted by the matrix tense. However, a tense dependency also exists in [11b], the embedded past tense must also be interpreted as being in the past of the matrix tense (the Upper Limit Constraint, see Abusch 1997, among others). In other words, [11b] cannot describe a scenario where the wondering temporally precedes Mary’s fixing something, which precedes the time of speech, such as a scenario where the listener wondered, “How will Mary fix this?” To the extent that both examples contain a tense dependency, wh-movement out of

---

4Richards also discusses the possibility of linking covert wh-movement with sequence of tense phenomena, which I will not address.
the embedded wh-question should be acceptable in both examples, contrary to fact.

Instead, I would like to consider an alternative explanation suggested by Grano & Lasnik: [11b] involves a PIC violation, while the relative acceptability of [11a] reflects a bound pronoun subject effect. I note that the PRO subject of the embedded clause is bound by the matrix subject; the finite paraphrase in (12b) is similarly acceptable. Suppose that in both (11a)/(12a) and (12b), the binding of the embedded subject happens syntactically and satisfies the PIC, and it does so before wh-movement. The PIC being satisfied by binding, we can invoke the PMC to explain why wh-movement may violate the PIC.

(12) a. What₁ were you₂ wondering how PRO₂ fix t₁?  (= [11a])
b. What₁ were you₂ wondering how you₂ can/should fix t₁?

For concreteness, I will assume the following:

First, wh-movement is triggered by a feature on C and v (McCloskey 2002, e.g.), and it is subject to the PIC as stated in (13), among other conditions.⁵

(13) a. “Weak” PIC (Chomsky 2001)
   In the configuration \([ZP \ldots [HP \alpha [H YP]]]\), where ZP and HP are phases, the domain of a phase head H (namely, YP), is not accessible to operations at ZP; only H and its edge \(\alpha\) are.
b. C and v are phase heads.⁶

The low acceptability of [11b] reflects (at least) a PIC violation: the wh-phrase what moves directly from the embedded vP to the matrix vP without first moving to the specifier of the embedded CP, which is occupied by how.

Second, I adopt the definition of the PMC in (14).

(14) Principle of Minimal Compliance (Richards 1998:601 exx. 6 and 7)
   a. For any dependency D that obeys constraint C, any elements that are relevant for determining whether D obeys C can be ignored for the rest of the derivation for purposes of determining whether any other dependency D' obeys C.
   b. An element X is relevant to determining whether a dependency D with head A and tail B obeys constraint C if
      (i) X is along the path of D (that is, X=A, X=B, or A c-commands X and X c-commands B).
      (ii) X is a member of the class of elements to which C makes reference.

We next need a theory of how pronouns are bound in a derivation. For purposes of exposition, I will assume that the binding of the embedded pronominal subject is the reflex

---


⁶One could also assume that v is not a phase head; making this assumption has no consequence for the analysis that follows.
of a syntactic operation, although I will remain agnostic about the nature of this operation (see e.g. [Kratzer 2009; Wurmbrand 2015]). Adapting Kratzer’s account somewhat, suppose that C enters the derivation and binds PRO, as indicated by the addition of a subscripted φ

(15a) When the matrix subject is merged into Spec,vP, the subject in turn binds C [15c]. Both binding operations satisfy the PIC. According to the PMC [14] both the matrix v and embedded C can be ignored for determining whether subsequent operations, such as wh-movement of what [15d] obey the PIC or not.

(15) A PMC-based derivation of What₁ were you₂ wondering how PRO₂ to fix t₁?
   a. [CP Cφ [PROφ to fix] . . . ] [C binds PRO]
   b. . . .
   c. [vP youφ v wondering [CP how Cφ [PROφ . . . ] [Matrix subject binds C]
   d. [vP what [youφ v wondering [CP how Cφ [PROφ . . . [Wh-movement out of wh-island]

In contrast, there is no binding in [11b] since there are no bound pronouns. The PIC is therefore not satisfied prior to wh-movement, and the matrix v and embedded C cannot be ignored for determining the well-formedness of wh-movement.

Evaluation of Proposal 3

This PMC-based account also satisfies the criteria set out in the start of this section. First, it only assumes the existence of phase and non-phase heads. Second, there is no need to give an account of how phasehood is gained or lost: this requirement is subsumed by the independently-motivated PMC.

That said, this account has several weaknesses. First, it is contingent on a theory of bound pronouns, and one major problem lies in constraining this theory so that the PMC account does not overgenerate. Recall that only bound pronoun subjects exhibit a phase obviation effect; bound pronoun objects and possessors do not. Applying the logic of the PMC, one must conclude that only the binding of a subject satisfies the PIC, while the binding of an object or a possessor does not [16]. (Since bound objects and possessors can appear in well-formed sentences, I follow Grano & Lasnik in assuming that these pronouns must have entered the derivation as a regular pronoun with φ-features.)

(16) a. More teachers₁ claimed that they₁ gave the students pencils than pens. [Satisfies PIC]
   b. *More teachers₁ claimed that the principal gave them₁ pencils than pens. [Does not satisfy PIC]
   c. *More teachers₁ claimed that their₁ assistants gave the students pencils than pens. [Does not satisfy PIC]

It seems relatively easy to account for the subject/object asymmetry. Suppose that binding of these bound pronouns depends on a feature transmission operation between binder and
bound pronoun (Kratzer 2009; Wurmbrand 2015). In the case of a bound pronoun object (16b), feature transmission might fail because the subject *the principal* intervenes between the matrix subject (the binder) and the object.

However, it is more difficult to explain why there is a subject/possessor asymmetry. In the example in (16c), there is no potential intervener between the binder *teachers* and the possessor. In fact, proposals on bound pronouns and fake indexicals like Kratzer’s and Wurmbrand’s are explicitly designed to generate examples like (17), where the subject *I* syntactically binds the possessor to derive *my*.

(17) Only I look after my children. (≈ “only I. λx.x looks after x’s children”.)

A second problem is that a PMC account does not provide a non-ad hoc account for the existential *there* effect. Consider the acceptable (18a). For the PMC account to work, one needs to assume that somehow the matrix binder *I* binds the expletive *there* subject in a way that satisfies the PIC, even though it is unclear why binding takes place at all. There is also no independent morphosyntactic evidence for such an operation. For the sake of argument, suppose binding involves feature transmission from binder *I* to the bound item *there*. However, there is no overt impact on agreement morphology in the embedded existential construction: the embedded verb still agrees with the pivot NP, not with the matrix subject (18b).

(18) a. Sally thinks there are problems with the heating, and *I* think there are problems with the climate control.
   b. Sally, *i* thinks that there {*is/are j* problems} . . .

4 Phase obviation effects in nominals

In the above section, I presented three alternative analyses of the phase obviation effects discussed by Grano & Lasnik and Barros & Frank. On conceptual and empirical grounds, I argued in favor of the first analysis, centered on $\phi$-feature valuation on phase heads and an alternative formulation of the PIC.

In this section, I present English data showing that similar obviation effects can be found outside of clauses. While the data do not let us to adjudicate between the three accounts, as mentioned in the introduction, they do constitute a new argument for treating (some) DPs as phases, providing support to the idea that both clauses and nominals are locality domains (see, e.g. Chomsky 1973). They also underscore the need for a theory of locality / phases as developed in Grano & Lasnik 2018, Barros & Frank 2017, and this paper.
4.1 The bound possessor effect

Clauses are not the only constituents whose boundaries dependencies are sensitive to. For instance, the dependencies in gapping and wh-movement cannot cross the boundary of a definite DP (a, b examples of (19) and (20)). Interestingly, as Davies & Dubinsky (2003) observed for wh-movement, this restriction is relaxed when the definite DP has a bound possessor (c examples).

This is not to claim that a bound possessor is sufficient to ensure the acceptability of gapping and wh-movement involving definite DPs. For example, Davies & Dubinsky observe that in wh-movement, there is an additional requirement that the verb be a “verb of creation,” for example, to tell rather than to hear, to write rather than to edit, to sing rather than to listen to. I will set aside the question about why the main verb matters, leaving that for future research. In this paper, I will only be concerned with the effect attributable to the bound possessor, which I will call the “bound possessor” effect for ease of reference.

(19)  
Gapping
  a. John joked about Obama, and Mary joked about Trump.
  b. *John told Colbert’s joke about Obama, and Mary told Colbert’s joke about Trump.
  c. ?John₁ told his₁ joke about Obama, and Mary₃ told her₃/₅ joke about Trump.

(20) Wh-movement (Davies & Dubinsky 2003)
  a. [Which president]₁ did John joke about t₁?
  b. *[Which president]₁ did Mary₃ tell Colbert’s joke about t₁?
  c. [Which president]₁ did Mary₃ tell her₃/₅ joke about t₁?

The bound possessor effect strongly resembles the bound pronoun subject effect. Therefore propose to unify the two, making two assumptions. The first is the assumption that possessive DPs are structurally analogous to clauses, an assumption independently motivated in the DP Hypothesis literature (Szabolcsi 1994, see also Abney 1987). Note that in (21a) “Poss” is intended as a syntactic category, although not necessarily one with possessive semantics. Second, I assume that at least possessive DPs are phases like clauses are. Both assumptions about DPs build on the intuition central to the DP Hypothesis that nominals and clauses are syntactically parallel to each other.

(21) a. CPs and DPs are isomorphic. Subjects and possessors are structurally analogous. (cf. other analyses where possessors are in Spec,DP, or are of the category D.)
For exposition purposes, I will analyze the bound possessor effect with the assumptions set out in Section 3.1. For a typical possessive DP, gapping across its boundary involves moving from a phase, namely, the DP itself. Consider the derivation when the possessor bears valued $\phi$-features, as Colbert’s in the case of (19b)/(22a). D agrees with the possessor and gets its own $\phi$-features valued. Movement of the remnant from the DP across D violates the PIC, resulting in low acceptability.

However, if the possessor is bound (19c)/(22b), it may enter the derivation with unvalued $\phi$-features. In that event, when D agrees with the bound possessor, D’s $\phi$-features remain unvalued. Movement of the remnant out of the DP phase does not violate the PIC, and the resulting sentence is (relatively) acceptable.

(22) Note: $\phi$: valued $\phi$-features. Italics: bound pronoun with unvalued $\phi$-features

A similar analysis applies to the bound possessor effect for wh-movement (20), but on the assumption that wh-phrases do not move to Spec, DP in English. Following McCloskey 2002 a.o., I assume that movement to the specifier of a head must be triggered by features on the head. In English, lexical idiosyncrasies mean that features that trigger wh-movement are found on C (maybe also v), but not on definite D. This assumption ensures that there is
no “escape hatch” in a DP that enables successive cyclic movement of a \( wh \)-phrase.

### 4.2 The demonstrative effect

[Davies & Dubinsky (2003)](#) also point out that, if the main verb is a verb of creation, \( wh \)-movement from a definite DP also becomes more acceptable when there is a demonstrative in the DP (23). I will refer to this contrast as the “demonstrative effect.”

(23)  
\[
\begin{align*}
\text{a. } & \text{ *[Which president]_1 \text{ did Mary tell } [DP \text{ Colbert’s jokes about } t_1]?} \\
\text{b. } & \text{ [Which president]_1 \text{ did Mary tell } [DP \text{ those jokes about } t_1]?}
\end{align*}
\]

I suggest assimilating the demonstrative effect with the existential \( there \) effect. Key to this analysis is the assumption that in demonstrative DPs, Spec,PossP is occupied by a morpheme that lacks valued \( \phi \)-features; this is the nominal analogue of expletive \( there \). For concreteness, I propose the following decompositional analysis of demonstratives (24), although other analyses are possible.

(24)  
\[
\begin{align*}
\text{a. } & \text{ Distal and proximal semantics are encoded on morphemes, of the category Poss, that bear unvalued \( \phi \)-features.} \\
\text{b. } & \text{ An expletive \( th \)- is inserted in Spec,PossP to satisfy an EPP feature on distal and proximal Poss.} \\
\text{c. } & \text{ This expletive lacks valued \( \phi \)-features.} \\
\text{d. } & \text{ Syntactic structure for } [DP \text{ those jokes } \ldots ]
\end{align*}
\]

\[
\begin{align*}
\text{DP} \\
\text{D } \quad \text{PossP} \\
\text{Agree } \quad \text{th-} \\
\text{Poss } \quad \text{-ose} \\
\text{NP } \quad \text{jokes } \ldots
\end{align*}
\]

I argue that English agreement morphology provides independent support for this assumption. For the sake of argument, suppose that the expletive \( th \)- morpheme comes with fixed \( \phi \)-feature values. If so, the distal or proximal Poss head should agree with it. In other words, demonstrative articles should be morphologically invariant regardless of the number of the NP complement, a prediction that is not borne out. Demonstrative articles in English show number agreement with the NP complement instead (25).

(25)  
\[
\begin{align*}
\text{a. } & \text{ Proximal Poss: this book-\( \emptyset \) (sg.) / these books (pl.)} \\
\text{b. } & \text{ Distal Poss: that book-\( \emptyset \) / those books}
\end{align*}
\]

In addition, analyzing \( th \)- as an expletive inserted to satisfy EPP features can explain why
demonstrative articles in English all have the same voiced *th* - initial, instead of showing greater variation in morphophonological form. The same analysis could be extended to French, whose demonstratives begin with a *c* - (/s/) initial.

### 4.3 Alternative hypotheses

#### 4.3.1 Davies and Dubinsky 2003

The above analyses of gapping and *wh* -movement in nominals offer several conceptual advantages over Davies & Dubinsky’s analysis of the bound possessor and demonstrative effects in *wh* -movement. In Davies & Dubinsky’s proposal (pp. 28–29), definite DPs can “incorporate” at LF onto the verb under specific circumstances: when the definite DP is a result nominal, modified with a bound possessor or a demonstrative that is hypothesized to contain a PRO (interpreted as the “agentive subject” of the result nominal), and when it is the object of a verb of creation. They further assume that a definite DP blocks government, and is thus a blocking category for *wh* -movement. However, when the definite DP incorporates at LF, the blocking effect is undone, according to the Government Transparency Corollary of Baker 1988.

Adopting a phase-based analysis eliminates the need to appeal to blocking categories or government, yielding an account more consistent with standard Minimalist assumptions. In addition, Davies and Dubinsky’s proposal presents a potential ordering paradox: incorporation, an LF operation, feeds *wh* -movement, which occurs in the overt syntax. This is inconsistent with conventional assumptions, where overt syntactic operations precede LF operations. In contrast, the current phase-based proposal does not depend on LF incorporation and so avoids the paradox.

#### 4.3.2 A weak definite analysis

An alternative approach relies on the intuition that there are actually two types of definiteness: strong and weak (26) (e.g. Schwarz 2009, 2014; Simonenko 2013, 2015, among others). Briefly, strong definites require an antecedent, while weak ones do not. Instead, weak definites have a uniqueness requirement, relativized to some situation.

(26)

a. Strong definite

Mary went to Washington, D.C., and met the city councilmember.

(Felicitous only if there is already a salient city councilmember in the context.)

b. Weak definite

Mary went to Washington, D.C., and met the mayor.

Boyd (1992) also suggests a similar analysis involving a PRO for acceptable *wh* -movement of arguments from clausal complements of definite deverbal NPs like *claim* (pp. 91–92).
(Felicitous when referring to the mayor of D.C., even when there is no prior mention of the D.C. mayor.)

Simonenko (2013, 2015) points out that weak definites allow wh-movement from within to form wh-questions, but strong definites do not; an example of this asymmetry is given in (27). One might therefore try to analyze the bound possessor and demonstrative effects in terms of weak/strong definiteness.

(27) a. Strong definite
   *[Which city]₁ did Mary meet [the city councilmember {of/for} t₁]?

   b. Weak definite
   [Which city]₁ did Mary meet [the mayor of t₁]?

There are two arguments against adopting this approach. The first argument is empirical: it is not clear that there is strong independent evidence for treating these DPs as weak definites. More precisely, the demonstratives involved in the demonstrative effect seem to require an antecedent. Consider (28) where there is no antecedent available for those jokes about . . . In this context, wh-movement from this definite DP feels odd. It improves, however, if an antecedent is available, for instance, if Alex had specifically said earlier that Mary told jokes about presidents (29).

(28) Alex: Were you at the comedy club last night? The theme was “U.S. Presidents” and Mary was the main performer.
   Ben: Unfortunately, I couldn’t make it.
   #*[Which presidents]₁ did she tell those jokes about t₁ ?

(29) Alex: Were you at the comedy club last night? Mary was the main performer and she told so many great jokes about U.S. presidents.
   Ben: Unfortunately, I couldn’t make it.
   [Which presidents]₁ did she tell those jokes about t₁ ?

The requirement for an antecedent is a standard diagnostic in the literature on weak/strong definites. To the extent that the demonstratives in the demonstrative effect require antecedents, they are strong definites, not weak ones.

The second argument is conceptual. Suppose we were able to extend a weak/strong definite analysis to account for the bound possessor and demonstrative effect. However, because the notions of antecedence and uniqueness do not have simple analogues in the clausal domain, this approach cannot be easily extended to the clausal domain to also account for the bound pronoun subject and existential there effects. Certainly, one could treat them as unrelated phenomena, one existing at a nominal level, and the other at the clausal level. But this analysis would miss generalizations about bound pronouns and bound possessors and about existential and demonstrative constructions.
4.3.3 Comments on the strong definite the

How, then, can we account for the fact that strong definite DPs with the article *the* are incompatible with *wh*-movement (30)? A thorough review of existing proposals and data on the interaction between *wh*-movement and *the* is beyond the scope of the paper; for recent proposals, interested readers should see Davies & Dubinsky 2003, Simonenko 2013, 2015 and references therein. In this section, I sketch an analysis that is consistent with the spirit of the current phase-based proposal, to show that the proposal can be extended to model the properties of the strong definite *the*.

(30) *(Which president)₁ did Mary tell {the / Colbert’s} jokes about t₁?*

For concreteness, I adopt Schwarz’s analysis of strong definite DPs (see Schwarz 2009:265, ex. 300, Simonenko (2015) adopts the same analysis). Schwarz proposes that strong definites contain an unpronounced indexical argument, while weak definites do not. This argument appears in the specifier of a head with an NP complement. Integrating this analysis with my assumptions about the syntax of DPs, I suggest that the head is of category Poss, and the indexical argument bears fully valued φ-features (31).

(31) Syntax of strong/weak definites (cf. Schwarz 2009)

a. Strong definite: \[DP \ D \ [\text{PossP} \ 1 \ [\text{Poss} \ Def \ [\text{NP} \ jokes \ldots]]]]

b. Weak definite: \[DP \ D \ [\text{PossP} \ Expl \ Def \ [\text{NP} \ jokes \ldots]]

Because the indexical argument (in (31a) “1”) has fully-valued φ-features, when D agrees with it, D gets its features valued. As a result, *wh*-movement out of DP violates the PIC. In contrast, there is no indexical in the weak definite (31b); if an expletive (Expl) without valued φ-features is inserted in the same position, then D’s features will remain unvalued upon agreement with the expletive. Subsequently, *wh*-movement does not violate the PIC.

5 Discussion and conclusion

5.1 φ-features or shiftiness?

In the three proposals presented in Section 5 I have followed Grano & Lasnik in assuming that locality restrictions associated with a phase head depend on φ-feature valuation. This assumption can be contrasted with Barros & Frank’s, which states that the phasehood of a CP ultimately depends on the discourse properties of a subject, specifically, its “shiftiness.”

It is technically straightforward to implement Barros & Frank’s idea for the three proposals: we make C’s features sensitive to the syntactic reflexes of discourse properties, instead of φ-features. I have not done so here, though, as I think that the relevant discourse properties are not well understood; in contrast, it is more straightforward to evaluate a φ-feature valuation account.
For example, there appears to be a difference between expletive *it* and *there* that a shiftiness account does not predict. A simple survey of ten native speakers of American English suggests that phase obviation effects are more restricted with the expletive subject *it* than with expletive *there*\(^\text{(32)}\) pseudogapping, comparative deletion, and multiple questions are quite clearly less acceptable in the presence of expletive *it* subjects.\(^8\) As far as I can tell, this asymmetry is unexpected under a shiftiness account, since both expletive subjects do not refer and should be equivalent in discourse terms. A proposal based on φ-feature valuation, such as the one discussed in Section 3.1, in contrast, predicts that the *it* sentences should be worse, assuming that *it* differs from *there* in bearing valued φ-features, namely, third person singular features.

(32) Phase obviation effects with expletive *there* and *it* (acceptability ratings on a 1–7 scale; 1 = unacceptable, 7 = acceptable)

a. Gapping (mean rating for *there* sentence = 5.7, for *it* sentence = 6.1, difference not significant, \(p\)-value = 0.68)
   (i) Jill claimed that there was a problem with Microsoft Excel, and Sally claimed there was a problem with Google Chrome.
   (ii) Jill claimed that it will rain three inches, and Mary claimed that it will rain four inches.

b. Pseudogapping (*there* sentence = 4.9, *it* sentence = 2.1, rating difference significant, \(p\)-value < 0.01)
   (i) Sally will probably claim that there is a problem with Microsoft Excel, whereas Bill will probably claim there is a problem with Google Chrome.
   (ii) Jill will claim that it will rain in Baltimore, and Mary will claim that it will rain in Washington, D.C. (where the PPs modify to *rain*)

c. Comparative deletion (*there* sentence = 6.3, *it* sentence = 5.0, rating difference significant, \(p\)-value = 0.04)
   (i) More students claimed there was a problem with Microsoft Excel than claimed there was a problem with Google Chrome.
   (ii) More meteorologists think that it will rain in Washington, D.C. than think that it will rain in Baltimore.

d. Antecedent-contained deletion (*there* sentence = 6.0, *it* sentence = 5.2, rating difference not significant, \(p\)-value = 0.35)
   (i) Joe claimed that there was a problem with every computer program Tim did claim that there was.

---

\(^8\)The native speakers all had a linguistics background. The sentences rated included the ones in (32), although the struck-out portions were omitted. To make the intended reading clear, a context was presented for each sentence. I will also set aside the fact that some of the *there* sentences do not seem to be perfectly acceptable, since my goal here is to highlight the differences between expletive *there* and *it*. Statistical analysis was done in R with the COIN package [R Core Team 2018][Hothorn et al. 2008], using a two-sided exact Wilcoxon-Mann-Whitney test.
(ii) John claimed that it rained in every city that Bill did claim it rained.

e. Multiple questions (there sentence = 4.6, it sentence = 2.6, rating difference significant, p-value = 0.04)

(i) Tell me which student claims that there’s a problem with which computer program.

(ii) Tell me who claims it will rain how many inches tomorrow.

5.2 On the Complex NP Constraint

Like the theory of subjacency, phase theory provides a way to derive the Complex NP Constraint (Ross 1967). When a CP and a (definite) DP are phases and there is no wh-movement to Spec,DP, a wh-phrase must move directly from the complement of D. Such a movement operation violates the PIC. However, according to the current proposal, there are also circumstances where movement out of a DP does not violate the PIC: when the DP has a demonstrative or a bound possessor, and when the main verb is a verb of creation. We predict wh-movement from complex NPs to be acceptable when these conditions are met.

This prediction is partially supported. As Davies and Dubinsky themselves noted (pp. 31–32, also Boyd 1992 and Ross 1967), under these circumstances, wh-movement of arguments is possible out of these complex NPs (33a). However, I note that wh-movement of adjuncts remains impossible; in (33b), the adjunct how angrily can only be understood as modifying the matrix VP write his report . . ., and not the VP in the complex NP criticized the assistant. This argument–adjunct asymmetry suggests that the Complex NP Constraint cannot be entirely reduced to the PIC or subjacency. Instead, a principle like the Empty Category Principle appears to be needed to account for the contrast.

(33)  a. ?Who did John write [DP his report [CP t1 that the mayor criticized t1]]?
    b. *[How angrily did John write [DP his report [CP t1 that the mayor criticized the assistant t1]]]?

5.3 An argument for the Activity Condition

Proposal 1 (Section 3.1) also bears on a debate on the Activity Condition (34). One of the arguments for the Activity Condition is the need to block hyper-raising (35). A-movement of a subject of an embedded finite clause to the matrix subject position. The condition claims that A-movement is impossible because at the point of movement, uninterpretable case features on the subject have been valued and eliminated.

Interestingly, Davies & Dubinsky (2003) give the following example, which contains a verb of creation but not a bound pronoun:

(i) Who did Kerry start the rumor that Kelsey is fond of? (ibid. pp. 31–32, ex. 78a, their judgment)

Boyd (1992) does not discuss wh-movement from complex NPs with bound pronouns either.
(34) Activity Condition (e.g. Chomsky 2000: 126–127; Chomsky 2001: 6)
Goal and probe must bear uninterpretable features for syntactic operations to apply.

(35) *There$_1$ seems $[CP$ (that) $t_1$ is someone at the door].

Nevins (2005) argues that hyper-raising can be independently ruled out with three constraints: (i) an appropriate formulation of the PIC that bans any syntactic operation from outside a phase from targeting elements in the complement of the phase head (Chomsky 2000), (ii) the “Single Case Constraint,” which bans nominals from receiving more than one case, and (iii) a ban on improper movement (from an A-bar position to A-position). Under Nevins’ proposal, (35) is unacceptable for two independent reasons: First, *there has moved directly from the embedded subject position to the matrix subject position, violating the PIC. Second, in both subject positions, *there receives nominative case from finite tense, thus violating the Single Case Constraint.

His proposal also rules out (36): the derivation is blocked because A-movement of *there violates the PIC, even though *there does not violate the Single Case Constraint, assuming that non-finite tense cannot assign any case to the subject. (For both sentences, there is another illicit derivation involving improper movement of *there to Spec,CP and then to the matrix subject position.)

(36) *There$_1$ to seem $[CP$ that $t_1$ is someone at the door].
(cf. There seems to be someone at the door.)

The theory of phases sketched in Proposal 1 and the assumption that expletive *there can lack valued $\phi$-features, however, predict hyper-raising to be possible in (36) even if we adopt the ban on improper movement and the Single Case Constraint. In this example, embedded C (that) agrees with *there, and C’s features do not get valued. As a result, A-movement of *there does not violate the PIC as defined in Proposal 1. Nor does this movement operation violate the ban on improper movement. In addition, *there only receives nominative case once. Since the low acceptability of (36) cannot be attributed to this version of the PIC, some other constraint on derivations — like the Activity Condition — is necessary, pace Nevins, although here I will remain agnostic about how this constraint should be stated.

5.4 Conclusion

In this paper, I argued that existing accounts of the bound pronoun subject effect and related phase obviation effects contain assumptions and stipulations that are conceptually problematic. I presented and evaluated three proposals that potentially resolve these problems, and argued in favor of the first proposal, where locality effects are linked to the valuation of $\phi$-features on a phase head. One consequence of this analysis is that there is no need to

---

10 Both Grano & Lasnik and Barros & Frank do not explicitly discuss hyper-raising or the Activity Condition.
posit that phases are formed in a derivation, contrary to recent proposals by Grano & Lasnik and Barros & Frank. I also underscored the relevance of these effects, by pointing out that similar effects are found in the DP domain. Finally, the DP data provide a new argument for the phasehood of (definite) DPs.

Of course, many open questions remain. For example, it would be interesting to show that the PIC falls out from some general principle of syntactic computation. In the context of this paper, one might ask why the PIC, as formulated in the first proposal, is the way it is. In the second proposal, I looked at the possibility of doing so with M. Richards’ Value–Transfer Simultaneity. While I ultimately argued against this approach, other approaches might exist. More work is also needed to clarify if similar effects are found with bound pronouns / non-shifty subjects in other languages. I hope that the questions and arguments presented here will lead us to improved accounts of phases and locality.

References


Wurmbrand, Susi. 2015. Fake indexicals, feature sharing, and the importance of gendered relatives. Handout of MIT colloquium talk.