Chapter 33

Strong and weak "strict cyclicity" in phase theory

Ángel J. Gallego
Universitat Autònoma de Barcelona

This paper explores the possibility that the No Tampering Condition (NTC) is eliminated in favor of a strong version of the Phase Impenetrability Condition (PIC). This possibility is welcome on theoretical grounds, given the redundant nature of NTC and PIC. We review empirical evidence indicating that the (original formulation of the) NTC is violated phase internally, a possibility that does not extend to the PIC. In so doing, we also consider the weak version of the PIC discussed in Chomsky (2016).

1 Efficient Computation

Generative Grammar has endorsed various economy principles (from Chomsky's 1975 [1955] *Traffic Convention* to Chomsky's (1995) *Minimal Link Condition*, going through many others). All such proposals adhere to a "least effort" desideratum attributed to the syntactic computation of the Faculty of Language. Within the Minimalist Program (MP), the basic structure-building operation is Merge – the only one that "comes free," without justification (Chomsky 2001: 3; 2008: 137).

Assuming it operates without bounds, Merge takes two objects, α and β, to construct a new object, γ. Additional applications of Merge target γ, which is the

Ángel J. Gallego. 2019. Strong and weak "strict cyclicity" in phase theory. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic Architecture and Its Consequences: Synchronic and Diachronic Perspectives, Volume 1: Syntax inside the grammar*, 653–670. Berlin: Language Science Press. DOI:??

only object left in the derivation (Chomsky 1995: 243), to yield γ′, and then γ′′, and so on and so forth—again, without bounds:[1,2]

(1) a. Merge(α, β) = {α, β}
b. Merge(λ, γ) = {λ, γ}
c. Merge(ψ, γ′) = {ψ, γ′}

A Merge-based system is enough to capture the property of CYCLICITY, that is, "in essence, the intuition that the properties of larger linguistic units depend on the properties of their parts" (Chomsky 2012: 1).[3] It is easy to see that a cyclic system will be largely compositional (Chomsky 2007: 5; 2012: 2): if computation is meaningful in an efficient manner, the interpretation of a given linguistic object will not be changed later on, which corresponds with "the general property of STRICT CYCLICITY" (Chomsky 2007: 5).[4] Therefore, whereas cyclicity follows from Merge alone, strict cyclicity requires something else—the mere existence of such operation does not in and of itself guarantee the conservation of the already assembled structure. This is the natural scenario where MP invokes so-called third factor conditions, which fall into two broad categories (Chomsky 2005):

(2) Third-Factor Conditions
a. Principles of data analysis that might be used in language acquisition and other domains;
b. Principles of structural architecture and developmental constraints that enter into canalization, organic form, and action over a wide range, including *principles of efficient computation*, which would be

[1] That α and β are no longer available was expressed in the following passage: "Applied to two objects α and β, Merge forms the new object K, *eliminating α and β*" (Chomsky 1995: 243, my emphasis).

[2] In Chomsky (2007: 11; 2008: 139) it is assumed that the free nature of Merge follows from LIs having an EDGE FEATURE (EF) that is undeletable and can thus give rise to an unbounded application of Merge. I will not assume EFs. Apart from the empirical advantage of dispensing with EFs (they have no realization in any language, so they are a purely theory-internal device), this allows us to dispense with the technical problems discussed in Narita (2014), related to the lack of EF percolation.

[3] As an anonymous reviewer observes, this is not the case if Merge allows, e.g., countercyclic infixing of SPEC-T after C has already been merged (see Chomsky 2008), or Parallel, Sidewards, Late, etc. Merge. Cf. Chomsky et al. (2019) and references therein for discussion.

[4] Of course, the interpretation of "Mary" is different in *Someone called Mary* and *Mary called someone*. That the interpretation of a given SO cannot be changed should thus be restricted to a post-Merge scenario, a possibility that is not entertained in feature-based approaches to theta-roles.

> expected to be of particular significance for computational systems such as language. It is the second of these subcategories that should be of particular significance in determining the nature of attainable languages. (Chomsky 2005: 6, my emphasis)

Different conditions have been put forward in order to capture the idea that linguistic objects generated by the syntactic computation cannot be changed (where *change* covers a wide range of possibilities: deletion, feature-valuation, late-insertion, tucking-in, etc.), especially by adding *ad hoc* symbols or performing operations that depart from least effort metrics. This is precisely the role played by the Inclusiveness Condition (IC, Chomsky 1995: 228), the No Tampering Condition (NTC, Chomsky 2008: 138), and the Phase Impenetrability Condition (PIC, Chomsky 2000). Putting details aside, IC, NTC and PIC all play a similar role in the current model, which was already noted by Juan Uriagereka in his annotated version of Chomsky (2001):

> So the Extension Condition [still holds]. This is somewhat surprising, given the [adoption of] "tucking-in" in Chomsky (2000). In effect, *we have several things ensuring the cycle.* The EC, in a radical way for the upward boundary of the phrase marker; the PIC for a kind of downward boundary, beyond which the system doesn't see any further operations; the idea of interpretation/evaluation at the strong phase in addition to both of these, as the derivation unfolds; and, finally, the phase-like access to the Numeration. *Much room for improvement and unification* … [my emphasis]

Such a redundant scenario is not expected, if only at a purely methodological level. This note argues that (the strong version of) the NTC can be subsumed under the PIC, given that local (phase-internal) modification is possible.[5] Discussion is divided as follows: §2 reviews the different conceptions of the NTC that have been entertained within MP and the empirical problems that have been observed for it; §3 turns its attention to the PIC, focusing on the recent possibility that the complement of a phase does not leave the computation (Chomsky 2008; 2016); in §4, I argue that (the strong) NTC can be eliminated adopting a strong version of the PIC, whereby transferred computation is forgotten (literally expunged), yielding a straight version of strict cyclicity; §5 summarizes the main conclusions.

[5] Probably, the same can be said of the IC, by simply observing that labels, indices, traces, and similar devices are not part of any I-language.

2 Merge and the NTC

There is a very close relationship between Merge and the NTC on the one hand, and between Transfer and the PIC on the other (as we will see in more detail in section 3). In fact, I would like to underscore the fact that, whereas Transfer and the PIC (as well as the operations of Feature Inheritance (FI) and Agree)[6] apply at the phase level, Merge and the NTC do not invariably so (Chomsky 2007: 17; 2008: 143; 2013: 40, 42). I state this correlation as follows, which I would like to build on to argue that there is a deep connection between the phase-based architecture and the (mildly) context-sensitive nature of the Faculty of Language (cf. Chomsky 1956; Uriagereka 2008):[7]

(3) a. EM = context-free
b. IM / Agree / Transfer = (mildly) context-sensitive

In what follows I would like to briefly review the different formulations of the NTC. As the reader will see, the conclusion will be that there are various situations where a weak version of the NTC must be assumed, not only for operations like FI or Agree (Chomsky 2007:19,fn.26)[8], but also Merge.

In Chomsky (2000; 2001; 2004; 2005), no explicit mention to the NTC is made. Instead, the Extension Condition (EC) is responsible for capturing the idea that Merge always applies to the edge of an SO. Thus, EC makes sure that, given {α, β}, a new element δ can only be merged as in (4a), not (4b), which would be counter-cyclic.

(4) a. {δ, {α, β}}
b. {{α, δ}, β}}

Chomsky (2000: 136) discusses these options, noting that (4a) satisfies the EC whereas (4b) does Local Merge. In the same breath, he notes that

[6] I assume that Agree actually implies a complex set of operations: Feature Inheritance, Match, Valuation and Deletion. Deletion is meant to cover erasure of uninterpretable φ-features, but it can also be applied to heads, as in Chomsky's (2015) analysis of *that*-deletion. Cf. Epstein et al. (2016) alternative in terms of phase-cancellation. Cf. Gallego (2014) for an alternative approach to FI, with interesting consequences for Chomsky's (2015) analysis of the EPP, discussed in Gallego (2017).

[7] It is typically assumed that all operations but EM apply at the phase level, simultaneously (Chomsky 2004: 116; 2005: 19; 2007: 17; 2008: 155). This raises questions for derivational systems, where the application of rules is ordered, as in Chomsky (2015).

[8] FI is reinterpreted as copying in Chomsky (2013: 47). This also departs from the strong NTC (unless we adopt the formulation in Gallego 2014).

> weaker assumptions suffice to bar [(4a)] but still allow Local Merge under other conditions. Suppose that operations do not tamper with the basic relations involving the label that projects: the relations provided by Merge and composition, the relevant ones here being sisterhood and c-command. (Chomsky 2000: 136)

Chomsky (2000) goes on to argue that "derivations then observe the condition [(5)], a kind of economy condition, where R is a relevant basic relation" (p. 137).[9]

(5) Given a choice of operations applying to α and projecting its label L, select one that preserves R(L, γ)

(5) holds in general, except for head adjunction. In the case of XP merger, Chomsky (2000) observes that EC must be satisfied for second-Merge, but not for subsequent applications or Merge—the creation of specifiers, which amounts to accepting tucking-in (N. Richards 1997).

In Chomsky (2004), it is explicitly noted that the EC can come in a strong and a weak version, the latter accepting deviations from (5) (my emphasis):

> Cyclicity of derivation requires that Merge to α always be at the edge of α, satisfying *an extension condition, strong or weak* ("tucking in") […] There appears to be one significant counterexample to cyclic Merge: late insertion of adjuncts […] Elementary considerations of efficient computation require that Merge of α to β involves minimal search of β to determine where α is introduced, as well as least tampering with β: search therefore satisfies [Local Merge], and Merge satisfies an EC, with zero search. One possibility is that β is completely unchanged (the *strong EC*); another natural possibility is that α is as close as possible to the head that is the label of β, so that any Spec of β now becomes a higher Spec ("tucking in," in Norvin Richards's sense). Further questions arise under Merge with multiple Specs. Assume some version of the EC to hold, in accord with SMT. (Chomsky 2004: 109, my emphasis)

The NTC is first introduced in Chomsky (2005), when discussing conditions of efficient computation. What I would like to capitalize on from the following quote is how similar NTC and PIC are, in the sense that the former appears to be related to the fact that what has been constructed in the course of a derivation *can be forgotten*; this is relevant, since this is typically the hallmark of the PIC.

[9] This is what Lasnik & Uriagereka (2005: Ch. 2) and Epstein et al. (2012: 256) refer to as Law of Conservation of Relations.

> One natural property of efficient computation, with a claim to extralinguistic generality, is that operations forming complex expressions should consist of no more than a rearrangement of the objects to which they apply, not modifying them internally by deletion or insertion of new elements. If tenable, that sharply reduces computational load: *what has once been constructed can be "forgotten" in later computations, in that it will no longer be changed*. That is one of the basic intuitions behind the notion of cyclic computation. The EST/Y-model and other approaches violate this condition extensively, resorting to bar levels, traces, indices, and other devices, which both modify given objects and add new elements. A second question, then, is whether all of this technology is eliminable, and the empirical facts susceptible to principled explanation in accord with the "no-tampering" condition of efficient computation [...] Assuming the NTC that minimizes computational load, both kinds of Merge to A will leave A intact. That entails merging to the edge, the EC, which can be understood in different ways, including the "tucking-in" theory of Richards (1997), which is natural within the probe-goal framework of recent work, and which can also be interpreted to accommodate head adjunction. (Chomsky 2005: 11, 13, my emphasis)

Notice that what this says is that the NTC is a third-factor condition on the way Merge operates.[10] More precisely, the NTC guarantees that when Merge applies to α and β, we obtain a new SO, γ, which can then be merged with further objects. So, for instance, if γ is merged with δ, given that α and β themselves are gone from the computation, the only way for this to happen is by forming {γ, δ}. This way, Merge must be to the edge as it cannot tamper with the objects it applies—in the case at hand, Merge cannot break up γ or tamper with it.

What is relevant about Chomsky (2008) is the discussion of certain situations that threaten the strong NTC: FI and the analysis of subject raising to SPEC-T.

> A natural requirement for efficient computation is a "no-tampering condition" (NTC): Merge of X and Y leaves the two SOs unchanged. If so, then Merge of X and Y can be taken to yield the set {X,Y}, the simplest possibility worth considering. Merge cannot break up X or Y, or add new features to them. Therefore Merge is invariably "to the edge" and we also try to establish the [IC] dispensing with bar levels, traces, indices, and similar descriptive technology introduced in the course of derivation of an expression [...] Note that SMT might be satisfied even where NTC is violated—if the

[10] This formulation states that the NTC is Merge-sensitive alone, which opens the door for conditions being sensitive on independent operations.

> violation has a principled explanation in terms of interface conditions (or perhaps some other factor, not considered here). The logic is the same as in the case of the phonological component, already mentioned [...] *The device of inheritance* [...] *is a narrow violation of NTC*. The usual question therefore arises: does it violate SMT? If it does, then the device belongs to UG (perhaps parametrized), lacking a principled explanation. But the crucial role it plays at the C-I interface suggests the usual direction to determine whether it is consistent with 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: 138, 144, my emphasis)

Chomsky (2007; 2008) assumes that φ-features are generated in phase heads, from which they are downloaded (downward percolation) to non-phase heads. Following M. Richards (2007), the process is taken to be mandatory, under the PIC: Since these features must be deleted, they must end up in the Transfer domain.[11] FI has consequences for the analysis of raising-to-subject, as discussed by Epstein et al. (2012). In particular, suppose the derivation of *Don Quixote fought the windmills* is as depicted in (6):

(6) a. {Don Quixote, {v*{fought, {the,windmills}}}} = v*P
b. Merge (T,v*P) = {T, {Don Quixote, {v*{fought, {the,windmills}}}}}
c. Merge (C,TP) = {Cφ, {T, {Don Quixote, {v*{fought, {the,windmills}}}}}}
d. FI (C,T) = {C, {Tφ, {Don Quixote, {v*{fought, {the,windmills}}}}}}
e. IM (*DQ*,TP) = {C, {Don Quixote, {Tφ, {*t*, {v*{fought, {the,windmills}}}}}}}

The problematic steps in (6) are (d) and (e), but (e) more clearly so. As Epstein et al. (2012) discuss, the original (SPEC-less) TP must be disconnected from C so that the external argument (EA) *Don Quixote* undergoes IM with it; when this new (SPEC-ful) TP is created, and it is then reconnected to C. The operation is thus ternary, in that Merge must target the EA, TP, and C. Noam Chomsky (p.c.) notes that this is a narrow extension of Merge, but does not depart from it in the way head movement does, since the EA is merged with TP, which it is a term of.

So far, as we can see, a key trait of NTC/IC-constrained Merge (α, β) is that α and β cannot be modified: they are left unchanged, no features, indices, etc. can

[11] As pointed out in footnote 8, Chomsky (2013) suggests that FI is actually a form of copying. If correct, FI could simply be reduced under the Copy Theory of Movement, as argued in Gallego (2014).

be added to them by Merge. Chomsky (2007) gives another twist by noting that while Merge cannot modify α or β, some subsequent operation might:

> Merge (X1,…,Xn) = Z, some new object. In the simplest case, n = 2, and there is evidence that this may be the only case (Richard Kayne's "unambiguous paths"). Let us assume so. Suppose X and Y are merged. Evidently, efficient computation will leave X and Y unchanged (the No-Tampering Condition NTC). We therefore assume that NTC holds unless empirical evidence requires a departure from SMT in this regard, hence increasing the complexity of UG. Accordingly, we can take Merge (X,Y) = {X,Y}. *Notice that NTC entails nothing about whether X and Y can be modified after Merge* […] Under NTC, merge will always be to the edge of Z, so we can call this an edge feature EF of W. (Chomsky 2007: 8, my emphasis)

This observation can probably be related to Chomsky's (2015: 10–11) analysis of phase-head deletion (de-phasing), which triggers a process that makes a non-phase head inherit all the properties of a phase head. De-phasing is put foward in order to account for the fact that subjects can be extracted from *that*-less clauses (an Empty Category Principle (ECP) violation in earlier terminology). So, as is well-known, subject extraction across a CP is ruled out if that is spelled out (cf. Chomsky 1986; Rizzi 1990):

(7) $[_{CP}$ Who does the book say $[_{CP}$ (*that) $[_{TP}$ t_{Who} stabbed Caesar]]]?

Chomsky (2015) reinterprets this phenomenon in order to argue that C can undergo deletion. This makes T inherit phasehood, which makes it strong, with no need for a DP to occupy SPEC-T for labeling reasons (cf. Gallego 2017). More to the point, Chomsky (2015: 11) argues that "The natural assumption is that phasehood is inherited by T […] along with all other inflectional/functional properties of C (φ-features, tense, Q), and is activated on T when C is deleted".[12]

Let us take stock. NTC is the formalization of the idea that computation applies in an efficient way, so that Merge (α, β) cannot modify α and β themselves. This strong formulation of the NTC, which bars tucking in and derives the Copy Theory of Movement (CTM), captures more than mere cyclicity. In particular, what I would like to emphasize is that by not letting Merge modify what it applies to, the NTC further captures some form of strict cyclicity too. To see this, let us go back to (1), repeated as (8) below:

(8) Merge (α, β) = {α, β} = γ

[12] Noam Chomsky (p.c.) elaborates on this by noting that the NTC states that an SO should not be modified by Merge, which doesn't literally imply that it cannot be deleted.

After (8), the workspace contains γ and nothing else, so α and β are no longer available (Chomsky 1995: 243). At this point, we may want to merge γ and a new object, δ:

(9) Merge(δ, γ) = {δ, γ}

δ is either internal or external to γ. If external, δ is drawn from the lexicon. This is External Merge (EM). If internal (e.g., δ = α), then δ is a term of γ. Assuming the NTC, γ cannot be modifed, so it must remain {α, β}, which yields {α, {α, β}}, and thus two copies (occurrences) of α. More importantly for our purposes, the strong NTC entails that {α, β} must be left as it is, so merger of α will not tamper with γ by removing α. There is no need for an extra operation (Copy) for IM, just like it is not needed for EM—if α were taken from the lexicon, it would not be copied.[13]

This said, there are two potentially problematic aspects about the NTC. The first one follows from the very fact that the strong NTC runs into the empirical problems in (10):[14]

(10) Violations of strong NTC

a. Feature Inheritance (Chomsky 2008)
b. IM to SPEC-T (after EM (C,TP)) (Chomsky 2008)
c. Tucking-in (N. Richards 1997)
d. Head Movement (Chomsky 2001)
e. De-phasing (Chomsky 2015)
f. Phase-cancellation (Epstein et al. 2016)

[13] The problem is more general if α and β remained in the workspace, along with γ. As Noam Chomsky (p.c.) points out, it has always been assumed that they do not, for the generative procedure constructs a single object, not a multiplicity of objects. Changing that convention would mean that instead of a generative process for expressions, we would be designing a generative process for an arbitrarily large collection of expressions. For instance, suppose that we hold that after EM(α, β) = γ = {α, β}, the workspace contains α, β, γ. We then have a new question: what is the relation between X in the workspace (call it α1) and α2 in γ = {α, β} (call it α2)? They are either copies or repetitions. If they are copies, everything goes haywire. Thus, if we continue to Merge to α1 finally yielding the finite clause FC, and to γ yielding the finite clause FC′, then the two clauses would contain the two copies α1 and α2, so one should be deleted, and if one enters into some relation (say anaphora) then the other does, etc. Things get much worse if, as this proposal allows, we construct simultaneously indefinitely many finite clauses. This is not only dubious, and in fact makes the notion of "copy" collapse.

[14] If the NTC is restricted to Merge, as Noam Chomsky (p.c.) notes, then only (10b) and (10c) are problematic.

Apart from these *local* (phase-bounded) violations of the NTC, there is another important observation to be made about the strong NTC, namely the redundancy between it and the PIC, as I discuss in the following section.

3 Transfer and the PIC

We have seen that the NTC has two formulations, strong and weak. Let me express this as follows:

(11) a. Strong NTC (NTC_S) = SOs cannot be changed by Merge
b. Weak NTC (NTC_W) = SOs can be changed locally, but not by Merge

What I would like to discuss is the fact that NTC_S is virtually analogous to the PIC. The PIC was proposed in order to capture strict cyclicity, so that "operations cannot 'look into' a phase below" (Chomsky 2000: 108). Chomsky (2004) relates the PIC to the operation Transfer (a wider version of Spell-Out, capturing the interaction between NS and both interfaces), which is defined in (12):

(12) TRANSFER hands D-NS over to Φ and to Σ. (Chomsky 2004: 107)

In Chomsky (2004), Transfer makes it impossible for the externalization systems to access what has been cashed out at previous phases. The possibility that the same happens in the case of the narrow computation is not so clear:

> When a phase is transferred to Φ, it is converted to PHON. Φ proceeds in parallel with the NS derivation. Φ is greatly simplified if it can "forget about" what has been transferred to it at earlier phases; otherwise, the advantages of cyclic computation are lost […] PIC sharply restricts search and memory for Φ, and thus plausibly falls within the range of principled explanation […] *It could be that PIC extends to NS as well, restricting search in computation to the next lower phase.* (Chomsky 2004: 107, my emphasis)

That the PIC does not carry over to the computation is connected to the existence of structures, in Icelandic or Spanish, like those in (13), where T can agree with the in-situ internal argument (IA):

(13) {T, {v*, {V, IA}}}
Agree

Empirically, (13) requires the φ-Probe to override the PIC and access the complement domain of v* (see M. Richards 2012). In order to tackle this, Chomsky (2001, 2004) adopts a weak version of the PIC, which led to a scenario analogous to that of the NTC, with both strong and weak versions:

(14) a. Strong PIC (PIC1 or PIC_S)
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)

b. Weak PIC (PIC2 or PIC_W)
[Given structure [$_{ZP}$ Z … [$_{HP}$ α [H YP]]], with H and Z the heads of phases]: The domain of H is not accessible to operations at ZP; only H and its edge are accessible to such operations. (Chomsky 2001: 14)

PIC2 is incompatible with FI, so in Chomsky (2008) it is discarded. Consider the following discussion, for suggests that phases that have been transferred can in principle be accessed (modulo intervention effects). Chomsky concludes that the effects of the PIC hold for the interfaces, but not necessarily NS:

> For minimal computation, as soon as the information is transferred it will be forgotten, not accessed in subsequent stages of derivation: the computation will not have to look back at earlier phases as it proceeds, and cyclicity is preserved in a very strong sense. Working that out, we try to formulate a PIC, conforming as closely as possible to SMT […] *Note that for narrow syntax, probe into an earlier phase will almost always be blocked by intervention effects.* One illustration to the contrary is agreement into a lower phase without intervention in experiencer constructions in which the subject is raised (voiding the intervention effect) and agreement holds with the nominative object of the lower phase (Icelandic). *It may be, then, that PIC holds only for the mappings to the interface, with the effects for narrow syntax automatic.* (Chomsky 2008: 143, my emphasis)

Chomsky (2016) in fact argues that Transfer should not eliminate anything from the NS. Otherwise, it would not be possible to explain how the structures in (15) are formed:[15]

(15) a. [$_{\alpha}$ The idea [$_{\beta}$ that the Earth is round]] was rejected t_{α}

b. [$_{\alpha}$ That [$_{\beta}$ I kept my job]] seems to t_{α} bother Mary

[15] I put aside another situation where the PIC is strongly violated: covert movement. This matter is pointed out (not addressed) in Chomsky (2004: 111; 2005: 13).

The problem here is as follows: in both cases, β is a phase, so it should be transferred before α is raised to matrix SPEC-T. But how can β be pronounced along with α if it is gone from the computation? Chomsky (2016) claims β is never gone from the workspace, but rendered inaccessible by Transfer. There are two ways to interpret this version of the PIC, which I will call PIC3: what's been processed is either (i) totally inaccessible or (ii) can not be changed.[16] Given the data in (15), (i) must be dismissed. We therefore expect that violations of the PIC do not change whatever is inside the transferred phase. This crucially allows us to change what is outside it, including the φ-Probe of matrix T in (16), taken from Fernández-Serrano (2016):

(16) Spanish

Me encantan [$_{CP}$ PRO escuchar [$_{v*P}$ t_{PRO} t_{v*} [$_{VP}$ V truenos]]]
to.me love-3.pl listen thunder
'I love to listen to thunder.'

Let us therefore assume the PIC3 allows access into a lower phase, as long as it is not modified. This makes it difficult to keep the copy / repetition distinction. Take (17), call it K, where the lower phase complement containing β, that is {α, β}, has already been transferred:

(17) K = {…{P, {α, β}}

Imagine we now merge β with K. β could be taken from the lexicon, so it would be a repetition. Can it be a copy? Given that {α, β} is not expunged from the derivation, the question is whether NS can tell whether β is taken from the lexicon or it is interpreted as an occurrence of the β contained within P's complement. If {α, β} can be accessed, the system cannot tell the difference. But we want to exclude this, or successive cyclic movement would go away. Island conditions would be affected too. Notice that the logic here is clear: the copy / repetition distinction does not require changing anything within the already passed phase. So, it should be possible to do that, given Chomsky's (2016) PIC3.

A way out would be to assume, as Noam Chomsky (p.c.) suggests, that if β raises from {α, β}, then both {α, β} and β itself have been modified: {α, β} by now

[16] A reviewer points out that what I call PIC3 is actually a conception of Transfer and its effect on transferred material, not the PIC, which "describes the timing of Transfer and the size of the transferred object". For the purposes of this paper, I will not dwell on this (to me, largely terminological) issue. The PIC was meant to state what is accessible and what is not after Transfer (a mapping operation) applies. All I am assuming is that the PIC3 says that everything is actually accessible after Transfer as long as it is not changed.

containing a copy that is part of chain, and β by the mere fact of becoming a discontinous object. Now, if this is correct, even the application of IM to *Who* changes the v*P and *Who* in (18).

(18) {Who, {Samson, {v*, {defeated, *t*}}}}

Presumably, this has not been considered problematic, for it does not violate the PIC, but it does the NTC_S. Now, we have seen that NTC_S and PIC are remarkably similar in that they both capture strict cyclicity. If nothing else, (18) shows another scenario where we depart from the NTC_S. I take this to indicate that the NTC_S is to be dispensed with entirely. More controversially, I also argue that the NTC_W is dispensable, if (IF) the PIC can play its role. Under PIC1, which I repeat here as (19), this replacement is possible:

(19) Strong PIC (PIC 1 or PIC_S) 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)

What (19) says is enough to capture the effects of the NTC. In particular, the fact that the objects generated in the course of the derivation cannot be tampered with. Notice that this *does* allow tampering *before Transfer applies*, but we have seen that this is empirically sustained. To the cases listed in (10), we can add a sixth one, which follows from the PIC3:

(20) Violations of NTC_S

a. Feature Inheritance (Chomsky 2008)
b. IM to SPEC-T (Chomsky 2008)
c. Tucking-in (N. Richards 1997)
d. Head Movement (Chomsky 2001)
e. De-phasing (Chomsky 2015)
f. Phase-cancellation (Epstein et al. 2016)
g. IM (chain creation)

In the next section, I would like to summarize the main ideas of the previous pages and, at the same time, argue that the PIC3 can be eliminated in favor of the PIC1. In so doing, I also discuss how the data mentioned in Chomsky (2016) can be handled under such proposal. The proposal entails that Transfer eliminates material from the workspace, yielding a more effective reduction of computational load—the original motivation behind Phase Theory (cf. Chomsky 2000).

4 NTC eliminated: some consequences

Let me spell out the interim conclusions so far. I will phrase them as questions:

(21) a. Do we need both NTC and the PIC?
b. If we need the PIC, do we need the PIC3?

Both NTC and PIC express an efficiency desideratum, namely that a given SO should not be changed (manipulated, tampered with, altered, etc.) once it has been created. This creates a redundancy, as I have pointed out.[17] At the same time, we have seen different phenomena indicating that the strong version of the NTC cannot be maintained. Should the weak version be? I think it should not, just like the weak PIC (the one in Chomsky 2001). This raises the more general question whether the strong PIC could be the only cyclic principle. If so, then the derivation can allow tampering up to the phase level, when Transfer applies. Suppose the derivation has assembled α and β to yield this:

(22) {α, β}

Suppose next that we apply IM to β. If the NTC does not hold, this could yield (23), potentially affecting the CTM.

(23) {β, {α}}

Note that this derivation is not forced (thus, the CTM does not go away), but the question is whether the step in (23) creates a problem. It is not clear that it does, at least if something like (23) is at stake for de-phasing (cf. Chomsky 2015).

If the only cyclic condition is the PIC, the next question is (21b). Recall that there are two empirical arguments to sustain it. The agreement facts (cf. (16)) could be tackled if Agree takes place at the border of NS-externalization, not in NS. This would have two welcome consequences. On the one hand, we could explain the parametric nature of Agree, which I would like to relate to Chomsky's (2014) *Thesis T*:

(24) Language is optimized relative to the conceptual-intentional (CI) interface alone, with externalization a secondary phenomenon. (Chomsky 2014: 7)

[17] A reviewer does not see the redundancy, as (s)he takes the NTC to be a third-factor condition on Merge (defining a Merge-cycle that adds stuff to the derivation) and the PIC to be a natural result of Transfer (which removes stuff from the workspace). Given the (empirical) arguments given below (and in Chomsky et al. 2019) it is unlikely that the PIC actually removes stuff from the workspace.

The *Thesis T* tells us that efficiency of operations should be found in the NS → SEM channel, not in the NS → PHON one, which is further consistent with the claim that "language is primarily an instrument of thought, with other uses secondary" (Chomsky 2014: 7). If Agree is pushed to NS → PHON, then the fact that its effects are subject to parametrization (as appears to be the case), would fall into place, and would also be compatible with the idea that language variation and parametrization are to be found only there (Chomsky's 2001 Uniformity Principle; cf. Chomsky 2010; Berwick & Chomsky 2011).

Another consequence of this concerns the very nature of Agree, which is a complex operation, consisting of Match, Valuation, Transfer and Deletion. Chomsky (2004 et seq.) takes these operations to somehow apply simultaneously (at the phase level), but this is hardly consistent with a derivational system, for operations must be ordered (as in Chomsky 2015).[18] Plausibly, the operations should be ordered as follows:

(25)
1. Match (NS)
2. Valuation (NS)
3. Transfer (NS → SEM/PHON)
4. Deletion (PHON)

As noted in Epstein & Seely (2002), this timing is problematic, since it entails that uninterpretable features will be valued before Transfer, becoming undistinguishable from interpretable ones. Unless Deletion could apply at SEM too somehow deleting the uninterpretable, but valued, φ-features of v* and C, operations would have to apply simultaneously, which, as noted, is odd within a derivational system. A way out is at hand if the derivation can somehow remember that φ-features were introduced as unvalued. This should be possible, given the relevance of phase-level memory to distinguish trivial / non-trivial chains, which in its most direct interpretation would entail revamping the long-abandoned idea or feature chains (Chomsky 1995: 262, 270–271, 383, fn. 27, abandoned in Chomsky 2000 due to the intricacies of head movement). So, if Merge could apply not only to LIs, but also to features—more precisely, to their values, which is what seems to be copied from one LI to another, then this would assimilate Valuation to Merge, making it possible for the system to remember that a valued feature was introduced as unvalued, which would signal it as uninterpretable. The technical solution I am sketching would not be too different from FI itself. In brief,

[18] If Transfer is part of Externalization, then it can be subject to parametrization (for the same reasons Agree would be). This opens the possibility that the effects of Transfer vary from language to language (cf. Uriagereka's 1999 radical or conservative Spell-Out).

we could dispense with the simultaneity of operations and perhaps the need for Agree to apply in NS alone if Merge could apply to LIs, features and values.

Obata's (2010) data are different. Consider (26):

(26) [α That [β Judas left the dinner]] seemed [to *t*α worry everyone]]

Here β is transferred before α is raised to matrix SPEC-T, which makes it impossible for it to be spelled-out where we see it. However, even if we assumed that the PIC leaves β accessible (through the PIC3), this does not cover IM. That is, it is only α (presumably its head, *that*) that can raise to matrix SPEC-T, so how can β be pied-piped along with α? If we allowed that, then we would also be changing the already transferred object, as noted for (18) above. A possible way out for these cases is that what is transferred is turned into a pair <X,Y>. I would like to connect this to Chomsky's (2004) analysis of adjuncts, which adopted (27):[19]

(27) In ⟨α, β⟩, α is spelled out where β is (Chomsky 2004: 199)

If Transfer converts the structure into some kind of pair, then when IM targets α, the actual pronunciation of β (or some part of it) could be possible. This would have the effect of placing β in a "secondary plane" (Chomsky 2004), but we want α (the phase edge), and α alone, to remain in the primary plane. This is what the PIC1 bought us, which brings back the possibility that Transfer can yield (28), removing the complement domain from NS (cf. Ott (2011)):

(28) a. {Edge, {P, {β}}}
b. Transfer (β) = {Edge, {P}} or {Edge, P}

If Transfer applies this way, there would be tampering, but locally. (28) would make it possible for P to be the head of the entire phase, with consequences for the v*-EA relation (cf. Epstein & Shim 2015).

5 Conclusions

This paper has discussed the nature of different conditions put forward to capture computational efficiency within minimalism, most importantly, the NTC and the PIC. Given their redundant nature (they both aim at capturing the idea behind the strict cycle, namely that SOs formed in the course of a derivation cannot be changed at subsequent stages), one of them should be dispensed with. I

[19] Cf. Chomsky (2008: 139) for similar ideas in the case of Merge.

have argued that strict cyclicity effects follow from the PIC alone. The decision is justified on methodological and empirical grounds. The former have to do with the multiplicity of conditions favoring strict cyclicity. The latter concern the empirical evidence showing that the strong version of the NTC cannot be maintained.

The strong PIC (or PIC1 cf. Chomsky 2000), which is the one that should be adopted, forces successive cyclic movement (SCM). Since nothing is left in the (primary plane of) computation after Transfer, that's the only way for a chain to be created. It also follows that the SO that has been cashed out cannot be modified: it is gone from the workspace. Interestingly, there are no violations of the PIC analogous to those of the strong NTC, which is another argument to stick to the former. Interestingly, it seems that only CP and vP give rise to SCM—NPs, PPs and other categories lack it (cf. Gallego 2012; van Urk 2016), which may provide yet another reason to defend that only CP and vP are phases.

Abbreviations

CI = conceptual-intentional, **CTM** = Copy Theory of Movement, **EA** = external argument, **EC** = Extension Condition, **ECP** = Empty Category Principle, **EF** = edge feature, **EM** = External Merge, **EPP** = Extended Projection Principle, **EST** = Extended Standard Theory, **FI** = Feature Inheritance, **IA** = internal argument, **IC** = Inclusiveness Condition, **IM** = Internal Merge, **MP** = Minimalist Program, **NS** = narrow syntax, **NTC** = No Tampering Condition, **PIC** = Phase Impenetrability Condition, **SCM** = successive cyclic movement, **SMT** = Strong Minimalist Thesis, **SO** = syntactic object, **UG** = Universal Grammar.

Acknowledgements

I am very happy to contribute to this volume in honor of Ian Roberts, a key figure in the field of generative grammar. I had the opportunity to work with Ian back in 2008, when he supervised a British Academy Visiting Fellowship I was awarded with, right after I became a doctor. I remember that experience (with long appointments at Ian's office) as a very important one in my career and in my personal growth too.

A previous version of this paper was presented at the University of Michigan, in a talk organized within the UMich Linguistics Colloquium Series on 10 March 2017. I would like to thank the audience of that talk for questions and comments. For discussing these matters with me, I am also indebted to Noam Chomsky, Sam

Epstein, Hisa Kitahara, Dennis Ott, and Daniel Seely. This research has been partially supported by grants from the Ministerio de Economía y Competitividad (FFI2017-87140-C4-1-P), the Generalitat de Catalunya (2017SGR634), and the Institució Catalana de Recerca i Estudis Avançatis (ICREAAcadèmia 2015). Usual disclaimers apply.

References

Berwick, Robert & Noam Chomsky. 2011. The biolinguistic program: The current state of its evolution. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 19–41. Oxford: Oxford University Press.

Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2. 113–124.

Chomsky, Noam. 1975 [1955]. *The logical structure of linguistic theory*. New York: Plenum.

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 Uriagereka (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*, vol. 3 (The cartography of syntactic structures), 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language?: Chomsky's minimalism and the view from syntax–semantics*, 1–30. Berlin: De Gruyter.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2010. Some simple evo-eevo theses: How true might they be for language? Alice V. and David H. Morris Symposium on Language and Communication, Stony Brook University, New York, USA.

Chomsky, Noam. 2012. Foreword. In Ángel J. Gallego (ed.), *Phases: Developing the framework*, 1–7. Berlin: De Gruyter.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI:10.1016/j.lingua.2012.12.003

Chomsky, Noam. 2014. Minimal recursion: Exploring the prospects. In Thomas Roeper & Margaret Speas (eds.), *Recursion: Complexity in cognition*, 1–15. Dordrecht: Springer.

Chomsky, Noam. 2015. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, 1–16. Amsterdam: John Benjamins. DOI:10.1075/la.223.01cho

Chomsky, Noam. 2016. A dialogue with Noam Chomsky. Barcelona, 6 November 2016.

Chomsky, Noam, Ángel J. Gallego & Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions and challenges. *Catalan Journal of Linguistics*.

Epstein, Samuel David, Hisatsugu Kitahara & T. Daniel Seely. 2012. Structure building that can't be. In Myriam Uribe-Etxebarria & Vidal Valmala (eds.), *Ways of structure building*, 366–384. Oxford: Oxford University Press.

Epstein, Samuel David, Hisatsugu Kitahara & T. Daniel Seely. 2016. Phase cancellation by external pair-merge of heads. *The Linguistic Review* 33. 87–102.

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

Epstein, Samuel David & Jae-Young Shim. 2015. Two notes on possible approaches to the unification of theta relations. *Linguistic Analysis* 40(1–2). 3–20.

Fernández-Serrano, Irene. 2016. Parameters of hyper-agreement in spanish dialects. Talk at Workshop on Formal Approaches to Romance Microvariation, University of Bucharest, Bucharest, 24–26 November 2016.

Gallego, Ángel J. (ed.). 2012. *Phases: Developing the framework*. Berlin: De Gruyter.

Gallego, Ángel J. 2014. Deriving feature inheritance from the copy theory of movement. *The Linguistic Review* 31(1). 41–71.

Gallego, Ángel J. 2017. Remark on the EPP in labeling theory: Evidence from romance. *Syntax* 20(4). 384–399. DOI:10.1111/synt.12139

Lasnik, Howard & Juan Uriagereka. 2005. *A course in minimalist syntax: Foundations and prospects*. Malden, MA: Wiley-Blackwell.

Narita, Hiroki. 2014. *Endocentric structuring of projection-free syntax*. Amsterdam: John Benjamins.
Obata, Miki. 2010. *Root, successive-cyclic and feature-splitting internal merge: Implications for feature-inheritance and transfer*. University of Michigan dissertation.
Ott, Dennis. 2011. A note on free relative clauses in the theory of phases. *Linguistic Inquiry* 42(1). 183–192.
Richards, Marc. 2007. On feature inheritance: An argument from the phase impenetrability condition. *Linguistic Inquiry* 38(3). 563–572.
Richards, Marc. 2012. On feature inheritance, defective phases, and the movement-morphology connection". In Ángel J. Gallego (ed.), *Phases: Developing the framework*, 195–232. Berlin: De Gruyter.
Richards, Norvin. 1997. *What moves where in which language*. MIT dissertation.
Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
Uriagereka, Juan. 1999. Multiple spell-out. In Norbert Hornstein & Samuel David Epstein (eds.), *Working minimalism*, 251–282. Cambridge, MA: MIT Press.
Uriagereka, Juan. 2008. *Syntactic anchors: On semantic structuring*. Cambridge: Cambridge University Press.
van Urk, Coppe. 2016. On the distribution of reflexes of successive cyclicity. Talk at BCGL 9, Brussels (Belgium).