Alan Hezao Ke

Michigan State University

Department of Linguistics, Languages and Cultures

619 Red Cedar Road B-230 Wells Hall, East Lansing, MI 48824 USA

kehezao@msu.edu

A note on the domain of transfer

July 30, 2021

Abstract. This paper argues that transfer of the full phase, rather than the phase head complement, is theoretically and empirically preferable. Full phase transfer solves various problems that the standard theory of transfer faces. In addition, the main concern of the full phase transfer approach, that is, movement across phase boundaries, can be overcome with assumptions that are independently motivated.

Keywords: theory of transfer, full phase, phase head complement, phase theory, domain of transfer, successive cyclic movement

1.   Introduction

An influential hypothesis in the Minimalist Program is that structures are built phase by phase, respecting cyclicity. According to Chomsky (2001b), once a structure with two phases are built, at the time when the higher phase head merges into the root structure, the complement of the lower phase head is transferred to the following interfaces, PHON and SEM, with its phonological material handed over to the former and the semantic material to the later.[1] As a consequence, the lower phase head complement becomes not accessible to operations once the higher phase head merges in. This restriction on the accessibility of a transferred domain is reflected in the definition of the Phase Impenetrability Condition (PIC):

(1)   *Phase Impenetrability Condition (PIC)* (Chomsky 2001b:14):

Given [$_{ZP}$ Z... [$_{HP}$ [H YP]]], where Z and H are phase heads, the domain of H (=YP) is not accessible to operations at ZP; only H and its edge are accessible to such operations.

Chomsky (2000a, 2008) and subsequent work adopt a different formalization regarding the timing of transfer and thus changes the size of the syntactic object (SO) that is subject to PIC: the complement of a phase head YP is transferred when its local phase head H, instead of the higher phase head Z, merges in, rendering YP inaccessible even to operations at HP. Although this change generally reduces the number of accessible syntactic objects in Syntax, it does not reduce the number of accessible phases or phase heads in the workspace. In the rest of this paper, I will use the term "the classical/standard theory of transfer" to refer to both Chomsky's (2000a) and (2001b) theories of transfer, which take the phase head complement rather than the full phase as the domain of transfer.

A reasonable development along the lines of phase theory is that the domain of transfer is also a phase, as Chomsky (2004:108) has acknowledged: "[Spell-out](=transfer) must be able to Spell-out PH(ase) in full, or root clauses would never be Spelled Out." The standard theory of transfer is forced to transferred the

---

[1]Chomsky (2008, 2013, 2015) refers to these two interfaces as Sensory-motor (SM) and Conceptual-intentional (C-I), respectively.

full phase for the root structure in addition to the transfer of the phase head complement for all other phases. This disunity between full phase transfer and phase head complement transfer is a property of the theory that I seek to eliminate: this paper argues that transfer should uniformly apply to full phases indistinguishably to all phases. That is to say, theoretically, the most reasonable unit for transfer should be the full phase rather than the phase head complement, because full phase transfer is at least a necessary mechanism to transfer the root CP in every derivation.

Although various authors such as Chomsky (2000a) and Fox (2000) have suggested a full phase transfer approach, they do not provide an analysis. This idea quickly fades away due to some empirical reasons that favor phase head complement transfer.[2] A more developed argument regarding the full phase transfer approach can be found in Groat (2015) and Bošković (2016). This paper, however, being independently developed originally, is based on different assumptions and offers a distinctive analysis that reserves most merits of the standard theory of transfer (see Section 3 for details). For example, This paper deviates from Groat (2015) in at least two important aspects: (i) I adopt the idea that CP and vP are phases, but TP is not, whereas Groat (2015) considers every head that can check features, T included, as a phase head; (ii) Groat (2015) assumes an SO with unvalued features can stay in-situ without causing the derivation to crash upon transfer, which makes it always available for probing for Agree purposes, whereas this paper assumes transferred phase is not available for minimal search in general, Agree (and labeling) as a special implementation of minimal search included (Chomsky 2013; Ke 2019). On the other hand, Bošković (2016) adopts Chomsky's (2001b) transfer system where the lower phase is transferred after a higher phase head enters into the derivation; by contrast, the current approach can transfer a full phase either right away after it is completed or until the completion of the next phase. In addition, both Groat (2015) and Bošković (2016) hold the viewpoint that successive cyclic movement does not exist, whereas this paper provides

---

[2]One of the most important empirical reasons is that the edge of a phase P should be available to the higher phase Q for elements in P to move to Q. I will return to this issue later.

an alternative approach that can derive successive cyclic movement, which preserves the relevant classical insights established in many previous studies. In fact, the approach that is advocated here is a hybrid of the full phase transfer versions of Chomsky's above two theories of transfer. I will argue that this small yet significant change based on the standard theory is actually conceptually natural and empirically preferable.

The current full phase transfer approach solves various problems that the standard transfer theory faces, and gives correct predictions concerning edge effects, as well as long-distance wh-movement and head movement. Empirical motivations underlying the standard theory of transfer will be properly handled under the full phase transfer approach.

The structure of the paper is as follows: in Section 2, I highlight several problems of the standard transfer theory, to which full phase transfer can provide potential solutions. Section 3 discusses a main concern with regard to the full phase transfer approach: it may not be able to account for the edge effects and derive movement across phase boundaries in general. To address this concern, I lay out independent assumptions about the feature system and the non-crash proof nature of the system, then the Spell-out Order Preservation (SOP) principle (Fox and Pesetsky 2005a,b) is employed to derive the edge effects observed in successive cyclic movement. In that section, I also give examples to demonstrate how the system works for long-distance wh-movement and head movement. Section 4 addresses the problems of the standard theory of transfer raised in Section 2. Finally, Section 5 concludes the paper.

2.   Problems with the standard transfer theory

As mentioned above, Chomsky (2001b) considers phase head complement the primary domain of transfer. This is empirically motivated. The edge of a phase is kept in the workspace, exempted from transfer, until the next phase to allow raising of phase head and raising of phrases in the edge to the next phase (e.g. raising of the *v*P-internal subject to SpecTP). The edge also functions as an "escape hatch" for successive cyclic movement via phase edges (see also Chomsky 2001a). For example, the underlined positions in (2a) indicate

the positions where the wh-phrase *who* appears during the derivational process: it moves through the edge of each of the *v*P phases and the lower CP phase, until it arrives at the edge of the matrix CP phase (the SOs inside the pointed brackets "<>" are lower copies that are not directly relevant to the discussion).

(2) a. [$_{CP}$ who do you [$_{vP}$ <you> ___ think [$_{CP}$ ___ C [$_{TP}$ John saw [$_{vP}$ <John> ___ <see> ___]]]]]

   b. [$_{CP}$ ___ C [$_{TP}$ John saw [$_{vP}$ <John> ___ $v_{<see>}$ [$_{VP}$ V$_{<see>}$ ___]]]]

(2b) shows the derivation step where the complement of the lower phase head *v* is transferred (highlighted in gray) at the point when the embedded C merges in. If instead the *v*P is transferred in full, then the wh-phrase at the specifier of the *v*P might be trapped inside the *v*P if it does not raise to the specifier of the CP. Therefore, for the full phase transfer approach, we have to allow the wh-phrase to move to the specifier of the CP before the *v*P is transferred. I will turn back to this issue in Section 3.[3]

   With the empirical motivations of the classical transfer theory in mind, below I would like to highlight several theoretical and empirical problems caused by transfer of the phase head complement. Some of these problems have been noted, by Chomsky and others, but to my knowledge, they have never been aggregated for comprehensive and comparative evaluation of the classical transfer theory. In fact, most of them are assumed (incorrectly in my view) to be irrelevant to the domain of transfer. I would also like to show that most of the advantages of the standard transfer theory can be captured if we transfer the full phase, and that the disadvantages, on the other hand, can be overcome by the proposed analysis.

   To make some basic assumptions clear before we proceed, in this paper, I adopt Chomsky's assumption that CP and *v*P are phases (see e.g. Chomsky 2008:154-155 for reasons behind this assumption). The discussion is restricted to the so-called strong phases only, and I use *v*P to indicate a strong phase, which is

---

[3]Chomsky (2000a) mentions another reason why the phase head should not be transferred together with its complement: the phase head must be visible for selection and head-movement. This is crucial when the phase head complement is transferred at the point when its phase head merges in, like in Chomsky (2000a), whereas it is trivial for Chomsky's (2001b) transfer system. I will also derive head movement and address the selection issue in Section 3.

equal to Chomsky's (2000a) *v*⋆P.[4]

The first potential problem of the standard transfer theory concerns the conflict between the definition of phases and the domain of transfer. Chomsky (2000a, 2001b) argues that phases can be identified because they are propositional. Being propositional is a property at the SEM interface, and therefore it should be phases that are interpreted at the interface. However, as many researchers have noted, it is not the phase that is transferred and interpreted at SEM (e.g. Grohmann 2000; Epstein 2007; Boeckx and Grohmann 2007; Ott 2009). It may be argued that there are SOs that are propositional yet not phases (e.g. TP), and there are SOs that are not propositional (e.g. passive *v*P) but are phases based on other defining properties (Legate 2003). We agree with Gallego (2012:20) that SEM independence serving as a defining property of phases is problematic and insufficient, yet it still provides us useful information about a general distinction between regular TP, VP versus CP, *v*P. To be more specific, as Chomsky (cited in Gallego 2010:54-55) points out, interface motivation (including SEM independence) is more of a consequence than a cause. If we take this viewpoint, we should expect that as a natural consequence of the theory of phases, the interface would interpret CPs and *v*Ps rather than TPs and VPs (Problem [1]). To put it differently, whatever reason is being used to motivate the phasehood of CPs and *v*Ps instead of TPs and VPs, as long as we argue that CPs and *v*Ps are phases, whereas TPs and VPs are not, a natural expectation is that CPs and *v*Ps should be relatively more independent than TPs and VPs at the interfaces. Consequently, we expect CPs and *v*Ps, which do seem to be more independent in terms of their completeness of information structure or theta-roles in general, to be interpreted at the interfaces. That is, the domain of transfer should be full phases for phases to obtain an independence status. The issue with another relevant feature of phases, the phonological isolability of phases, can be considered along the same lines, in support of full phase transfer.[5]

In addition, when defining the concept "phase," Chomsky (2008) writes, "As discussed elsewhere (Chomsky 2001b), the size of phases is in part determined by uninterpretable features. . . A stronger principle

---

[4]It has been argued that the distinction between strong and weak phases is not necessary (Richards 2011).

[5]See Bošković (2016) for arguments in favor of full phase transfer building upon the syntax-phonology interaction.

would be that phases are exactly the domains in which uninterpretable features are valued, as seems plausible." (p. 155) This has been considered one of the most uncontroversial criteria that defines phases/phase heads (see relevant discussion in Gallego 2010). By itself, this diagnostic of phase heads does not cause problems. However, this idea conflicts with the other idea that the edge of a phase is a locus for lexical items with uninterpretable features (uFs) that must be accessible to further operations outside of that phase: the edge of a phase is a part of that phase; consequently a phase becomes simultaneously a domain where uFs are valued and a domain where uFs are kept active for future operations (Problem [2]). Therefore, if uFs are defining properties of phases, the status of phase edge cannot be justified. One may argue that if the quote is interpreted as "phases are the domains in which *any/some* specific uFs (e.g. ϕ-features) are valued," the conflict may be avoided. This is not the end of the problem, because an important point here is that under this definition of phases, the phase edge does not have any theoretical status to be a special escaping hatch for SOs with uFs.

There is also a timing problem between uF-valuation and transfer. Transfer cannot apply after uFs are syntactically valued, because such valued uFs will not be distinguishable from inherently, lexically valued features (in the absence of invoking some diacritic marking and/or a lookback device that is forbidden by the Inclusiveness Condition). But transfer also cannot happen before uFs are valued as this will cause crash (Chomsky 2001a). Thus, it is not clear when transfer can possibly apply (Problem [3]). Chomsky (2004) suggests a "simultaneity" approach, which eliminates intra-phasal "ordering," by taking valuation of uFs and transfer to happen concurrently. However, this approach encounters many serious analytical problems as well, depriving us of derivational (computationally efficient) explanation (see e.g. Epstein and Seely 2002), because we are not able to exploit the intermediate representations when operations apply simultaneously. With the simultaneity approach assumed, it becomes unclear why the narrow Syntax should build up SOs phase by phase, that is, in an ordered derivational manner, while inside a phase, operations should be applied simultaneously.

Indeed, there seems to be deep inconsistency between the current definition of (simplest) merge as a derivational operation and the simultaneity approach. An extension of the simultaneity approach is to use the simplest merge simultaneously to create an entire phase all at once. The basic operation would then become "build phase," but that is certainly not a combinatorial primitive operation as simplest merge is.

Moreover, the simultaneity approach is technically difficult, if not impossible, to implement as far as the relevant operations are concerned, if "simultaneity" implies no ordering between the operations.[6] That is, a simultaneity approach requires there to be no case where the output of an operation is fed into another operation as its input. By contrast, the operations that were argued to apply simultaneously in fact must be correctly ordered, as shown in (3).

(3)   a.   Valuation of uFs, including ϕ-features and Case feature.

b.   Deletion of the *valued* ϕ-features and Case feature.

c.   Transfer the phase complement with the uFs that have been *valued* and *deleted*.

The only order that does not result in crash is (3a)→(3b)→(3c), because (3b) takes the output of (3a) as its input given that it takes the valuation of uFs as its prerequisite, and (3c) takes the output of (3b) as its input because it must be applied after the valued uFs are deleted. Again, as long as some operations need to take the output of other operations as their input, these operations cannot be applied simultaneously. In a word, the simultaneity approach is inconsistent with the derivational nature of the language system and is technically not attainable in the specific situation where it is argued to apply (Problem [4]).

It must be noted that the timing problem and the simultaneity approach are relevant to the purpose of this paper because it motivates the transfer of phase head complement right at the point when the local phase head enters into the derivation. The timing problem forces the model of transfer established in Chomsky (2000a).

---

[6]The other way around, i.e. free of ordering implying simultaneity, seems logically flawed. For instance, if the operation X applies either before or after operation Y, we cannot conclude that these two operations must apply simultaneously.

If the problems due to the timing of transfer and the simultaneity approach can be solved independently, we would be enabled to adopt also (a full phase transfer version of) the model of transfer in Chomsky (2001b), namely, transfer can be applied to the lower phase when a higher phase head enters into the derivation.

The standard theory of transfer also encounters some non-trivial empirical problems. For example, under the standard transfer theory, the edge of a matrix CP can never be transferred to the interfaces for semantic and phonological interpretation unless we stipulate a special rule to transfer the edge of the matrix CP (Chomsky 2004, Problem [5]).[7] For example, in the case of (4), the standard transfer approach should stipulate a rule to deal with the transfer of the matrix CP specifically, by hypothesizing that "unlike all other phases, the highest phase head and its edge are transferred 'for free' along with the phase head complement...," and "[this might be] a mere albeit empirically motivated stipulation which is in effect...construction or category specific..." (Obata 2010:113, 134).

(4)   [CP What did [TP you buy]]?

There seems to be no principled reason why our language faculty should transfer the complement of the phase-head instead of the full phase. The distinction between the edge and complement of a phase seems to be not explicable in terms of computational efficiency or any other third factors (Chomsky 2005) or axioms of UG.

Finally, Matushansky (2005) and Citko (2014), among others, point out that an entire CP phase (as in 5a), rather than its complement TP (as in 5b), can move to a higher phase. The underlined blank space indicates the original base position of the moved CP or TP (or DP in (5c)).

(5)   a.   [CP That [TP John caught a fish]] was denied [CP ___]

       b.   *[TP John bought the book] was denied [CP that [TP ___]]

---

[7]As argued by Uriagereka (1999) and Obata (2017), this special rule needs to be extended to adjunct CPs to transfer adjunct CPs in full.

c. [DP The man who guesses [CP which book C [they think that the storekeeper hopes that the customers will buy <which book>]]] was found [DP ___].

They use the contrast between (5a) and (5b) as a diagnostic for the phase status of CPs and the non-phase status of TPs. As Chomsky (2012; 2013) acknowledges, the system should allow the moved CP inside the DP in (5c) to be interpreted "in the surface position, not the base position in which it entered the computation (Chomsky 2012:5)." This implies that the CP inside DP in (5c) needs to be kept in the workspace for future operation, although the interior of the CP phase is not accessible for further modification due to PIC.

Let us take a closer scrutiny of what the examples in (5) reveal regarding the theory of transfer. If the complement of the phase-head rather than the full phase is obligatorily transferred, it is incorrectly predicted that complete CPs in (5a) can never move. This is because the phase head complement (TP) will be transferred separately, making CP movement impossible (Problem [6]). The reasoning is that a transferred SO is an opaque unit that is not accessible to further operations, although the transferred SO itself as a giant lexical compound is accessible to the Syntax (cf. Uriagereka 1999; Chomsky 2013). Therefore, movement is predicted to be unable to target a phase which is transferred in two separate pieces. For instance, the embedded CP in (5a) is transferred in two units, with its edge transferred together with the higher phase (the matrix VP), and its phase head complement TP transferred by itself. Now movement of this CP would need to combine part of the transferred matrix VP and the phase head complement TP together, which seems impossible if the matrix VP is an opaque unit after transfer. If on the other hand, full phases are transferred, movement of full phase becomes a natural and empirically desirable consequence: a transferred phase can be moved as an opaque unit.

Adding to the examples in (5), topicalization seems also to be applied to whole phases rather than to complements of phases or to a secondary level internal phase. For example, (6a) is a case of *v*P topicalization, and it is allowed; on the other hand, (6b) and (6c) are cases where moved constituents are TPs, and they are not acceptable.

(6)   a.   [$_{vP}$ Buy a car], John will [$_{vP}$ ——].

      b.   *[$_{TP}$ Will buy a car], John [$_{TP}$ ——].

      c.   *[$_{TP}$ Bought a car], John [$_{TP}$ ——].

To summarize this section, a series of conceptual and empirical problems of the standard theory of transfer are highlighted, alongside its advantages. The problems are widespread over issues regarding the definition of phases, the distinction between the phase edge and the interior, the timing of transfer, the theoretical status of edges, the transfer of matrix CP, and the movement of full phases. It turns out that all these problems are related to the domain of transfer. I will address main concerns of the full phase transfer approach in the next section. I argue that the concerns can be dispelled if the full phase approach is equipped with some independently motivated assumptions. I will then turn back to address the problems raised in this section.

3.   Movement across phase boundaries

The theory of phases is proposed to reduce computational complexity. If we adopt the phase theory, we expect that transfer applies to phases rather than any other SOs. Otherwise, we need justification of why transfer should apply to these SOs (see also Bošković 2016). Full phase transfer is thus the null hypothesis that requires minimal burden of proof. However, a few problems immediately arise under the full phase transfer approach: How can SOs ever move to a higher phase if the edge of the lower phase is transferred together with the phase head complement? For example, how does successive cyclic movement work? In the case of (7), is it ever possible for the wh-phrase to undergo successive cyclic movement and finally land at the Spec of the matrix CP?

(7)   [$_{CP}$ who do you [$_{vP}$ —— think [$_{CP}$ —— John [$_{vP}$ —— saw ——]]]]

How can v-to-T movement ever apply? How can we derive head movement across phase-boundaries? Can

this proposal derive the edge effects without stipulating a special status for phase edge? These are potential questions and concerns regarding the full phase transfer approach, which I will address in this section.

3.1   Feature system

As noted, we need to derive successive cyclic movement under the full phase transfer approach. The features on the movers turn out to be relevant. I will thus first spell out the assumptions about the feature system that this paper adopts before we dip into the problems caused by successive cyclic movement.

Recent studies suggest that features have both semantic and phonological/morphosyntactic sides (Shen and Smith 2019; Ke 2019; and citations therein). In this paper, I will distinguish SEM features from PHON features on lexical items (and both of them are different from categorical features), and I assume they will be "interpreted" or read by the SEM and PHON interfaces, respectively. Interpretability here is thus not as "interpretable to only the SEM interface," which is what "interpretable" means in Chomskyan system (e.g. Chomsky 2000a). When I say a feature is "interpretable," I would also mention whether the feature is interpretable to the SEM interface or to the PHON interface. A semantic feature is interpretable to the SEM interface, and a phonological/morphosyntactic feature is interpretable to the PHON interface. Some features may be interpretable to both interfaces (cf. Epstein 2007). Therefore, the interpretability of a feature is relevant to both the SEM and PHON interfaces. This conception of interpretability is close to the "readability" of the features to the interfaces (Chomsky 2000b:17; see also Epstein et al. 2010).

What is "uninterpretable," or unimplementable, to both interfaces are unvalued features.[8] This is because unvalued features cannot give sufficient information to the interfaces to know how to either pronounce the feature at PHON or evaluate its semantic contribution to SEM. For example, if a T node does not value its unvalued ϕ-features (including the number feature), it would present insufficient instructions to the PHON interface because the PHON interface cannot translate the unvalued features to corresponding phonological

---

[8]This feature system is hence similar to a feature system that makes a four-way distinction, such as the one proposed by Pesetsky and Torrego (2007), except that the definition of interpretability is different here.

representations: e.g. should the unvalued number feature at T be pronounced as plural or singular? The following feature criterion summarizes this assumption:

(8)   *Feature criterion*

   The SEM and PHON interfaces are unable to implement unvalued features.

   The feature criterion has important implications on the timing of transfer. We now do not need to delete any valued feature at the time of transfer because they are implementable either to the SEM or to PHON interface (or to both). The interfaces will be required to implement only relevant features in their domain, and they will map irrelevant features to vacuous sound or meaning. For instance, after the number feature on a T head is valued as [+singular], it may be mapped to, e.g. [-s], at PHON, and be mapped to a vacuous meaning at SEM.

   As a result, under the current feature system, we do not need to trigger transfer of the phase head complement as soon as the phase head merges in, and the theoretical basis for the simultaneity approach, which I have argued to be problematic, is removed. We can then transfer full phases at times after they are complete.

   For the purpose of this paper, I assume, following the standard assumption in the literature (e.g. Chomsky 1995), that T acquires ϕ-feature values from an agreeing DP. The ϕ-features on T do not have much semantic content by themselves, thus these are unvalued PHON ϕ-features. As another example, a wh-phrase that is interpreted as interrogative at SEM and is pronounced with a rising intonation at PHON, e.g. the wh-phrases in English, holds an unvalued question feature [-Q] (Cable 2007), which is both an SEM and PHON feature.

3.2   Transfer and non-crash proof derivation

I assume a non-crash proof derivational system, different from Frampton and Gutmann (1999, 2002). This is resonant with the idea that merge is free and is not crash proof (e.g. Epstein et al. 2010; Chomsky 2013,

2015). Transfer applies whenever a full phase is complete, and if the transfer domain, a full phase, includes unvalued features when transfer applies, the derivation crashes. Nevertheless, if the SO with unvalued features moves to a higher phase, I assume, following Nunes (1995) and Bošković and Nunes (2007), that the lower copy of the SO must be deleted at PHON due to independent requirement on linearization (Kayne 1994). Consequently, the lower copies of a mover do not cause any problems to the PHON interface as they are phonologically null. This is relatively easy to understand. However, what about unvalued SEM features? I assume when an SO with unvalued SEM features moves to a higher phase, the unvalued SEM features are also deleted, because the lower copies will be turned into variables semantically (Heim and Kratzer 1998; Groat 2015). When the lower copies are variables semantically, they are necessarily bound by their higher copies if the movement involved is legitimate. That is, the variables induced by movement, unlike regular pronouns, do not need to have semantic formal features such as number and gender: their semantic content is determined by their binder (similar to minimal pronouns; see Heim 2008; Kratzer 2009). Such an analysis may explain why the lower copies of a quantified phrase has a different semantic type than the quantified phrase: while quantified phrases are of type $<<e,t>,t>$, their traces (or traces of DPs in general) as bound variables are always of type e (Barwise and Cooper 1981; Heim and Kratzer 1998; Poole 2017). Quantifier raising is in fact based on this type difference between quantified phrases and their traces to solve a type mismatch problem (cf. for example, May 1985). Finally, this is also consistent with Chomsky's (1995:301) suggestion that traces, i.e. lower copies of a mover, although cannot be fully erased, their "formal features" (e.g. ϕ-features and phonological features) are deleted where possible.[9]

---

[9]The deletion of formal features on lower copies is supported by the anti-superiority effects observed both in Bangla (Simpson and Syed 2016) and Icelandic (Kučerová 2016): in the case where T probes for a target object NP, if there appears an intervening NP bearing a feature identical to the relevant one borne by the target NP, a blocking effect occurs (either blocking movement of the target NP or blocking agreement between the prober and the target NP); however, if the intervening NP moves to the phase edge, its lower copy no longer causes an intervention effect. This is expected if the formal features that can match the probe are removed after movement.

Categorical features, serving as pure syntactic features, may be different from both PHON and SEM features, in that they must be retained on the lower copies. This is reasonable because categorical features can satisfy c(ategory)-selection of the relevant selecting head and satisfy the theta criterion in their base-generated positions. The categorical features of DP traces, which seems to be related to their semantic types, together with their variable status, may also give rise to scope reconstruction (Fox 2000; Groat 2015; Poole 2017). What's more, the well-known "wanna contraction" test also suggests that at least some features on traces, likely the categorical features, are preserved.

So far we have discussed two assumptions that seem natural theoretically and empirically: (i) transfer is not crash-proof, and (ii) the removal of PHON and SEM features on an SO after the SO moves to a higher phase and the lower phase containing the copy of the SO is transferred. That is, transfer of a mover's lower copy that bears unvalued features will not cause crash. Now we are ready to derive successive cyclic movement.

3.3   *Deriving successive cyclic movement*

Successive cyclic movement occurs after a phase head enters the derivation. I believe that this is another way to interpret the idea that "IM (internal merge) should be driven only by phase heads,"(Chomsky 2008:143) and it is consistent with the principle that "phases are exactly the domains in which uninterpretable features are valued." (Chomsky 2008:155) A phase head bears unvalued features and triggers probing. After all feature valuation operations apply, if the unvalued features are still not valued, they are not able to be valued in their local phase. At this point, there are two possibilities. The first possibility is that the local full phase containing the unvalued features gets transferred. This will cause the derivation to crash. The second option is to apply transfer after the SO with the unvalued features moving to a higher phase to be valued later.[10] This instance of movement is a case of internal merge, and it is free and optional (Chomsky 2013).

---

[10]The reader may wonder why the lower phase can be kept in the workspace such that a higher phase can be built upon it. Indeed, the workspace here can hold up to two phases and no more. Conceptually, in order to build a sentence

Therefore, movement of an SO with unvalued features may occur before or after the application of transfer. Successive cyclic movement of an SO with unvalued features is possible only when movement applies after transfer. For example, at the derivational step shown in (9), the wh-phrase with an unvalued feature [-Q] moves to merge with the root after the phase head *v* enters the structure. The wh-phrase then lands at the edge of the *v*P phase such that it can move further to a higher phase edge before *v*P is transferred.[11]

(9)   Who do you think John saw?

   [$_{vP}$ who$_{[-Q]}$ see [$_{VP}$ <see> <who>]]

A question that the reader may have is what if we have a phase that can be transferred when complete without causing the derivation to crash? Will the head that c-selects the phase (head) not be able to access the transferred full phase? Let us take (10) as an example.

(10)   John said [$_{CP}$ C Mary [$_{vP}$ <Mary> visited Bill]]

The embedded CP phase can be transferred when it completes. Now the question is, when *said* (or more accurately, *say*) enters the derivation, how can it be attached to a transferred phase? As mentioned in

structure that involves movement across a phase boundary, the minimum (and maximum) number of phases that the workspace needs to keep is two (excluding phases that have been transferred as only their labels are relevant to future computation; see below for discussion). Now it should be no surprise that in Chomsky's theories of transfer, just like the current full phase transfer approach, exactly two phase heads or materials from two phases are always kept in the workspace. Empirically, following Chomsky (2001b), this is also necessary to account for long-distance agreement across languages (Bhatt 2005; Boeckx 2004; Bošković 2007; Etxepare 2011; Keine 2020; Polinsky and Potsdam 2001; among others). It seems long-distance agreement across one phase boundary is well-attested, whereas that connects two SOs across two or more phase-boundaries is rarely observed (but see Bošković 2007 for relevant discussion). Such a contrast is well captured by the two-phase "vision" of Chomsky's (2001b) theory of transfer and the current approach.

[11]This analysis thus implies that an SO with unvalued features can move to the edge of its local phase and proceeds from there. Island effects will therefore not be analyzed as syntactic effects, but instead as consequences of sentence processing difficulties (cf. Sprouse and Hornstein 2013; Sprouse et al. 2016).

Section 2, Obata (2010) argues convincingly that transferred full phases are not completely removed from the workspace although the interior of the transferred phase are not accessible due to PIC.[12] I assume that once being transferred the full phase is labeled and the label of that phase can be selected by a head such as *said* in (10).

It is worth noting that the current analysis predicts an interesting difference between the CP and *v*P phases as exemplified in (10). The embedded *v*P phase cannot be transferred without causing crash before the *v*P-internal subject moves higher. That is, for the derivation to be successful (and it could fail), *v*P is transferred when the next higher phase CP is complete and the *v*P-internal subject *Mary* has moved to the SpecCP.[13] By contrast, the CP phase will be transferred right after it is completed. This difference between *v*P and CP may (partially) account for the opacity of CP and transparency of *v*P with regard to probing in Agree observed in the literature (cf. Keine and Zeijlstra 2021 and citations therein), a topic that is intriguing for future research.

3.4   Deriving the edge effects

Since we assume that merge applies freely and optionally without imposing additional restrictions which require substantial justifications (Chomsky 2013), movement does not need to always target phase edges; thus, we cannot explain why successive cyclic movement targets each phase edge, as previous studies of edge effects have suggested (Fox 2000; McCloskey 2000; Legate 2003; Sauerland 2003; Felser 2004; van Urk and Richards 2015; among others; although the conclusion is recently challenged by, e.g. Bošković

---

[12]See Chomsky et al. (2019) for a formalization of workspace that is potentially compatible with this paper.

[13]One could also assume that the transfer of *v*P occurs as early as possible, that is, right after *Mary* moves to the embedded SpecTP. However, in this paper, I will stick to the principle that transfer is triggered by the completion of phases. I will set aside an extensive evaluation of the theoretical and empirical consequences of this difference, due to space limitation, but see Ke (2019) for a discussion on a relevant difference these two choices can make in the analysis of cyclic agreement. In addition, not directly relevant to the purpose of the this paper, I will assume that the default hypothesis that transfer applies to one single phase each time, and leaving the consequences of this choice to future research.

2016). Given we have argued that phase edge does not have an independent theoretical status, we have to derive the edge effects without resorting to the concept of phase edge. This subsection is devoted to this purpose.

Let us start with something that is permitted by the current assumptions but does not give us the edge effects. The following example shows a possible derivational step: the wh-phrase can move to the specifier of the embedded CP directly without stopping over in the lower phase edge (Spec$v$P).

(11)   Who do you think John saw?

$[_{CP}$ who$_{[-Q]}$ C $[_{TP}$ John saw $[_{vP}$ <see> $[_{VP}$ <see> ___]]]

How can we then derive the edge effects? I argue below that edge effects follow from Fox and Pesetsky's (2005b) Principle of Spell-out Order Preservation (SOP), and point out that Fox and Pesetsky's system is compatible with the full phase transfer approach but not with the standard transfer theory.

Fox and Pesetsky (2005a,b) propose that SOP in (12) is imposed on each phonological spell-out domain (equivalent to phases, i.e. CP and vP) due to a linearization requirement.

(12)   *Spell-out Order Preservation (SOP)* (Fox and Pesetsky 2005b:6)

Information about linearization, once established at the end of a given Spell-out domain, is never deleted in the course of a derivation.

SOP requires the relative linear order between SOs in a spelled-out phase to be preserved after transfer; otherwise, the PHON interface would not be able to properly linearize the spell-out domains given an ordering contradiction. Again, use the wh-movement in (13) (equivalent to example (3) in Fox and Pesetsky (2005a) with minor changes on the labeling of the structure to accommodate the analysis to its equivalent under the current framework) to illustrate how SOP independently motivates successive cyclic movement of SO bearing unvalued features via each phase edge.

(13)   To whom will he say that Mary gave the book?

$[_{CP2}$ to whom$_{[-Q]}$ will he $[_{vP2}$ — say $[_{CP1}$ — that Mary $[_{vP1}$ — gave the book —]]]]
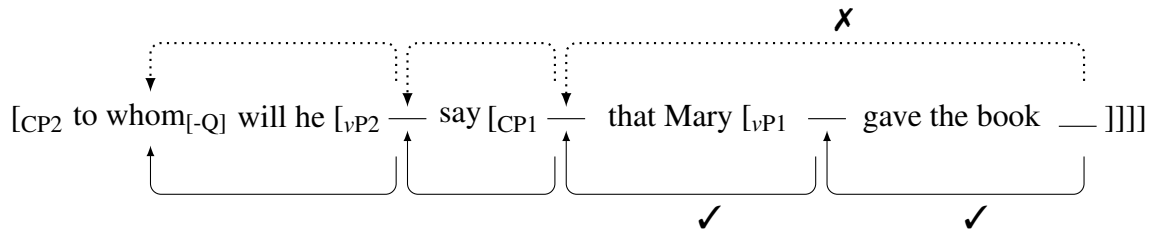
According to previous studies of edge effects, the wh-phrase in (13) must move through each phase edge; therefore, the derivation steps indicated by the dashed lines on the top, which skipped the phase edge of vP1, are prohibited, whereas those represented by the solid lines on the bottom are legitimate. Fox and Pesetsky (2005b) argue that this is because the former violates SOP. According to the former analysis, the wh-phrase moves directly from its base-generated position to the CP1 phase, skipping the edge of vP1. When vP1 is spelled out and is linearized, *gave* (or *the book*) precedes *to whom*. However, when the CP1 phase is spelled out and is linearized, the relative linear order between *gave* and *to whom* changes, because now *to whom* precedes *gave*, given *to whom* precedes the vP phase. According to SOP, the relative linear order of SOs cannot be changed after Spell-out. Thus, the impossibility of the analysis of (13) indicated by the dashed lines receives a principled explanation.

For the sake of completeness, I will take this opportunity to give an illustration of the whole derivation process, taking (13) as an example. As we have mentioned, *to whom* moves to Spec$v$P1 rather than SpecCP1 due to SOP. If it does not move but stay in-situ, when vP1 is built and transferred, the derivation crashes. That is, unless it moves to a higher phase, the derivation will never converge. It also needs to move to Spec$v$P2 rather than any other positions below, as required by SOP. Because if it moves to a position below *say*, when vP2 is transferred and linearized, *say* precedes *to whom*. *To whom* will then move on to a higher phase for the derivation to not crash.[14] This movement will change the linear order between *to whom* and *say*: in the

---

[14]Again, this movement is not imposed by a crash-proof system, but simply an option that is available, as move is optional and free.

*v*P2 phase, *say* precedes *to whom*, whereas when the CP2 phase is linearized, *to whom* will precede *say*, thus creating a conflict according to SOP. The possibility that the wh-phrase moves to a position below the phase edge of each phase is thus ruled out by SOP. Finally, the [-Q] feature on the wh-phrase will be valued by the C head of CP2.

We have seen that the full phase transfer approach with SOP can neatly handle successive cyclic wh-movement, which has caused notorious problems for the standard transfer theory (see Richards 2019 for a recent discussion). The full phase transfer approach captures the edge effects and correctly predicts that wh-phrases must move through the "agreementless" intermediate phase edge positions. It also solves another problem noticed by Felser (2004), i.e. the Convergence Problem. The Convergence Problem says that when *v*P1 in (13) is transferred, it contains a wh-copy which bears an unvalued feature given the wh-phrase moving to CP1 cannot be valued there (the unvalued [-Q] feature is valued only at the completion of movement at SpecCP2), and the derivation must crash at that point. However, it is no longer a problem under the current assumptions because after the wh-phrase moves upward, the PHON and SEM features of the lower copies, including the unvalued feature, are deleted, with the categorical feature being an only exception.[15]

It is important to note that SOP works under a full phase transfer/spell-out system.[16] As acknowledged by Fox and Pesetsky (2005b), "there is no need to distinguish phases from spell-out domains." (p. 15) Fox and Pesetsky (2005b) in fact have to assume that the unit of spelled out is full phases. If the spell-out domain is the phase-head complement instead, even the legitimate successive cyclic movement in (13) indicated by the solid lines is excluded by SOP, because in the lowest spell-out domain (the phase head complement of *v*P1), *the book* would precede *to whom*, and then in a higher domain (the phase head complement of CP1), *to whom* would precede *the book*, violating SOP.

---

[15]SOP then needs to apply before the deletion of the PHON features if it concerns about the PHON features besides the categorical features.

[16]See Fox and Pesetsky (2005a) for discussion of cross linguistic variation of the size of phases, i.e. the periodicity of spell-out in their terms, defending the idea that spell-out domains are equal to phases.

To summarize this subsection, I have shown that a constraint of phonological linearization proposed by Fox and Pesetsky (2005a,b), namely, SOP, independently drives the edge effects without the need to assume the special status of phase edge. In addition, SOP presuppose a full phase transfer system rather than the standard theory of transfer.

3.5 *Deriving head movement across a phase boundary*

We have just derived successive cyclic wh-movement and the edge effects under the full phase transfer approach. Now let us switch our attention to head movement across a phase-boundary. Since transfer of a lower phase can occur until the next higher phase head enters in the derivation, the lower phase may remain in the workspace without being transferred. Consequently, *v*-to-T movement in French, for example, is not a problem. We can derive *v*-to-T movement without the *v*P phase edge as an escaping hatch for *v* to move to T. As shown in (14) with the relevant derivational steps in (15), especially (15d), after the subject *I* and the verb *eat* move to TP, the lower phase *v*P can be transferred, highlighted in (15e), without causing crash.

(14) *v-to-T movement in French* (Iatridou 1990:554)

     Je mange souvent des pommes
     I   eat   often   apples

     'I often eat apples.'

(15)   a. $[_{vP}$ eat $[_{VP}$ <eat> apples$_{[-Case]}$]]

       b. $[_{vP}$ apples$_{[-Case]}$ eat $[_{VP}$ <eat> <apples>]]

       c. $[_{vP}$ I$_{[-Case]}$ often apples$_{[+Case]}$ eat $[_{VP}$ <eat> <apples>]]

       d. $[_{TP}$ eat $[_{vP}$ I$_{[+Case]}$ often apples <eat> $[_{VP}$ <eat> <apples>]]]

       e. $[_{CP}$ C $[_{TP}$ I eat $[_{vP}$ <I> often apples <eat> $[_{VP}$ <eat> <apples>]]]]

Therefore, we have seen that both successive cyclic movement and head movement across a phase boundary can be derived under the full phase transfer approach without much difficulty. The theory of

transfer that we end up with is a hybrid of the full phase transfer versions of Chomsky's (2000a) and that of Chomsky's (2001b) models of transfer: if a phase does not contain unvalued features, it can be transferred without causing crash right after it is completed; if a phase contains SOs with unvalued features, the phase can be transferred, leading to crash, but it can also be transferred when the next higher phase is complete, by which the SOs with unvalued features could have moved to the higher phase.

4.　Addressing the problems

Let us now return to the problems listed in Section 2. As is shown in Table 1, Problem [1] can be solved with full phase transfer: phases are the natural domain of transfer and there is no burden to motivate phase head complement as the transfer domain instead. In addition, the relatively more semantic and phonological independence and isolability of phases at interfaces are also expected. We also derived the edge effects in successive cyclic wh-movement with SOP, which solves Problem [2]. The current approach thus has the merit of not stipulating a special status of phase edges. Importantly, we have also shown that SOP works under the full phase transfer approach rather than the standard theory of transfer. The timing problem (Problem [3]) and the simultaneity approach (Problem [4]) were once used to argue for Chomsky's (2000a) theory of transfer, which I have contended to be not sustainable. This frees us from Chomsky's (2000a) system and I ended up adopting a hybrid of the full phase transfer versions of Chomsky's (2000a) and (2001b) theories of transfer: a phase can be transferred in full if it is free of unvalued features, otherwise it can be transferred after a higher phase head enters in the derivation and the SO with unvalued features are moved to the (edge of the) higher phase. However, the system is not crash proof, allowing full phases with unvalued features to be transferred, which will lead to crash. The current approach also does not require a special rule to transfer the edge of the matrix CP (Problem [5]) but correctly predicts that phases rather than phase head complements are accessible for movement after they are transferred (Problem [6]).

Table 1: Solutions to the problems

	Problem	Solution
1	Conflict between the definition of phases and the domain of transfer	Full phase transfer
2	The unmotivated status of phase edges	Derivation of phase edges by SOP
3	The timing problem of transfer	Transfer applies when a full phase is built
4	The simultaneity approach is technically impossible	The simultaneity approach is removed from the system
5	Need a special rule to transfer the edge of the matrix CP	No such rule needed
6	Phases but not phase head complements (TPs) can move	Full phases are opaque units available for movement after being transferred

5. Conclusions

In this paper, I first investigated potential conceptual and empirical problems of the standard transfer theory, which takes the phase head complement as the domain of transfer. It is argued that transfer of the full phase is both theoretical and empirically preferable. With very minimal assumptions that were independently justified in previous studies, the current approach can derive movement across a phase boundary without much difficulty. The full phase transfer approach captures the edge effects without stipulating a special status for phase edge, or making a distinction between phase edge and interior. Notorious problems associated with successive cyclic wh-movement are also addressed accordingly. In addition, various other problems of the standard theory of transfer are resolved with this slight but significant change on the domain of transfer.

References

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219. URL http://dx.doi.org/10.1007/BF00350139https://link.springer.com/content/pdf/10.1007%2FBF00350139.pdf.

Bhatt, Rajesh. 2005. Long distance agreement in hindi-urdu. *Natural Language & Linguistic Theory* 23:757–807. URL http://www.jstor.org.proxy.lib.umich.edu/stable/4048115.

Boeckx, Cedric. 2004. Long-distance agreement in hindi: Some theoretical implications. *Studia Linguistica* 58:23–36.

Boeckx, Cedric, and Kleanthes K. Grohmann. 2007. Remark: Putting phases in perspective. *Syntax* 10:204–222. URL http://dx.doi.org/10.1111/j.1467-9612.2007.00098.x.

Bošković, Željko. 2007. Agree, phases, and intervention effects. *Linguistic Analysis* 33:54–96.

Bošković, Željko. 2016. What is sent to spell-out is phases, not phasal compliments. *Linguistica* 56:25–66.

Bošković, Željko, and Jairo Nunes. 2007. The copy theory of movement: A review from PF. In *The copy theory of movement*, ed. Norbert Corver and Jairo Nunes, volume 107, 13–74. Amsterdam/Philadelphia: John Benjamins Publishing.

Cable, Seth. 2007. The grammar of q: Q-particles and the nature of wh-fronting, as revealed by the wh-questions of tlingit. Thesis, Massachusetts Institute of Technology.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000a. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000b. *Noam Chomsky: The architecture of language*. Oxford: Oxford University Press.

Chomsky, Noam. 2001a. Beyond explanatory adequacy. In *MIT Occasional Papers in Linguistics 20*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Working Papers in Linguistics (MITWPL).

Chomsky, Noam. 2001b. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond*, ed. Adriana Belletti, volume 3, 104–131. Oxford: Oxford University Press.

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

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

Chomsky, Noam. 2012. Forword. In *Phases: Developing the framework*, ed. Ángel J. Gallego, 1–7. Berlin/Boston: De Gruyter Mouton.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.

Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, strategies and beyond : Studies in honour of Adriana Belletti*, ed. Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, volume 223, book section 1, 3–16. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Chomsky, Noam, Ángel J Gallego, and Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics* 229–261.

Citko, Barbara. 2014. *Phase theory: An introduction*. New York, NY: Cambridge University Press.

Epstein, Samuel David. 2007. On i(nternalist) functional explanation in minimalism. *Linguistic Analysis* 33:20–53.

Epstein, Samuel David, Hisatsugu Kitahara, and T Daniel Seely. 2010. Uninterpretable features: What are they, and what do they do? In *Exploring crash-proof grammars*, ed. Michael Putnam, volume 3, 125–142. Philadelphia, PA: John Benjamins Publishing.

Epstein, Samuel David, and T Daniel Seely. 2002. *Rule applications as cycles in a level-free syntax*, 65–89. Oxford, UK: Blackwell Publishing.

Etxepare, Ricardo. 2011. Number long distance agreement in (substandard) Basque. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"* 40:303–349.

Felser, Claudia. 2004. Wh-copying, phases, and successive cyclicity. *Lingua* 114:543–574. URL http://www.sciencedirect.com/science/article/pii/S0024384103000548.

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

Fox, Danny, and David Michael Pesetsky. 2005a. Cyclic linearization and its interaction with other aspects of grammar: a reply. *Theoretical Linguistics* 31:235–262. URL //www.degruyter.com/view/j/thli.2005.31.issue-1-2/thli.2005.31.1-2.235/thli.2005.31.1-2.235.xml.

Fox, Danny, and David Michael Pesetsky. 2005b. Cyclic linearization of syntactic structure. *Theoretical linguistics* 31:1–45.

Frampton, John, and Sam Gutmann. 1999. Cyclic computation, a computationally efficient minimalist syntax. *Syntax* 2:1–27. URL http://dx.doi.org/10.1111/1467-9612.00012.

Frampton, John, and Sam Gutmann. 2002. Crash-proof syntax. In *Derivation and explanation in the minimalist program*, ed. Samuel David Epstein and Daniel Seely, 90–105. Oxford, UK: Blackwell Publishing.

Gallego, ángel J. 2010. *Phase theory*, volume 152. Amsterdam/Philadelphia: John Benjamins Publishing.

Gallego, Ángel J. 2012. *Phases: Developing the framework*. Berlin/Boston: De Gruyter Mouton.

Groat, Erich. 2015. Total transfer, dynamic labeling, and transfer remnants. In *Remnant movement*, ed. Günther Grewendorf, volume 123, 257–320. Berlin/Boston: Walter de Gruyter.

Grohmann, Kleanthes K. 2000. Prolific peripheries: A radical view from the left. Dissertation, University of Maryland, College Park.

Heim, Irene. 2008. Features on bound pronouns. In *Phi theory: Phi-features across modules and interfaces*, ed. Daniel Harbour, David Adger, and Susana Béjar, 35–56. Oxford, United Kingdom: Oxford University Press.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Blackwell textbooks in linguistics. Malden, MA: Blackwell.

Iatridou, Sabine. 1990. About agr(p). *Linguistic Inquiry* 21:551–577. URL http://www.jstor.org/stable/4178695.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Ke, Hezao. 2019. The syntax, semantics and processing of agreement and binding grammatical illusions. Dissertation, University of Michigan, Ann Arbor.

Keine, Stefan. 2020. *Probes and their horizons*. MIT Press.

Keine, Stefan, and Hedde Zeijlstra. 2021. Morphology of extraction: Reassessing vp phasehood URL https://stefankeine.com/papers/Keine-Zeijlstra-vP.pdf, manuscript.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40:187–237. URL https://muse.jhu.edu/article/263446.

Kučerová, Ivona. 2016. Long-distance agreement in Icelandic: Locality restored. *The Journal of Comparative Germanic Linguistics* 19:49–74. URL http://dx.doi.org/10.1007/s10828-016-9077-6.

Legate, Julie Anne. 2003. Some interface properties of the phase. *Linguistic Inquiry* 34:506–516. URL http://www.jstor.org.proxy.lib.umich.edu/stable/4179245.

Matushansky, Ora. 2005. Going through a phase. In *Perspectives on phases*, ed. Martha McGinnis and Norvin Richards, volume 49, 157–181. Cambridge, MA: MIT Working Papers in Linguistics (MITWPL).

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

McCloskey, James. 2000. Quantifier float and wh-movement in an irish English. *Linguistic Inquiry* 31:57–84.

Nunes, Jairo Morais. 1995. The copy theory of movement and linearization of chains in the minimalist program. Dissertation, University of Maryland at College Park.

Obata, Miki. 2010. Root, successive-cyclic and feature-splitting internal merge: Implications for feature-inheritance and transfer. Dissertation, University of Michigan, Ann Arbor.

Obata, Miki. 2017. Is transfer strong enough to affect labels? In *Labels and roots*, ed. Leah Bauke and Andreas Blümel, volume 128, book section 5, 117—126. Berlin: Walter de Gruyter.

Ott, Dennis. 2009. The conceptual necessity of phases: Some remarks on the minimalist enterprise. In *Explorations of phase theory: Interpretation at the interfaces*, ed. Kleanthes K. Grohmann, volume 17, 253–275. Berlin/Boston: Walter de Gruyter.

Pesetsky, David Michael, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture: Syntactic derivation and interpretation*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 262–294. Philadelphia, PA: John Benjamins Publishing Company.

Polinsky, Maria, and Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language & Linguistic Theory* 19:583–646.

Poole, Ethan. 2017. Movement and the semantic type of traces. Doctoral Dissertation, University of Massachusetts at Amherst.

Richards, Marc. 2011. Deriving the edge: What's in a phase? *Syntax* 14:74–95. URL http://dx.doi.org/10.1111/j.1467-9612.2010.00146.x.

Richards, Marc. 2019. Problems of" problems of projection": breaking a conceptual tie. *Catalan journal of linguistics* 139–152.

Sauerland, Uli. 2003. Intermediate adjunction with a-movement. *Linguistic Inquiry* 34:308–314. URL https://muse.jhu.edu/article/42196.

Shen, Zheng, and Peter W. Smith. 2019. Morphological and semantic agreement beyond hybrid nouns. In *Proceedings of the 49th Annual Meeting of the North East Linguistic Society (NELS 49)*.

Simpson, Andrew, and Saurov Syed. 2016. Blocking effects of higher numerals in bangla: A phase-based analysis. *Linguistic Inquiry* 47:754–763. URL https://muse.jhu.edu/article/634148.

Sprouse, Jon, Ivano Caponigro, Ciro Greco, and Carlo Cecchetto. 2016. Experimental syntax and the variation of island effects in english and italian. *Natural Language & Linguistic Theory* 34:307–344. URL https://doi.org/10.1007/s11049-015-9286-8.

Sprouse, Jon, and Norbert Hornstein. 2013. *Experimental syntax and island effects*. Cambridge University Press.

Uriagereka, Juan. 1999. Multiple spell-out. In *Working minimalism*, ed. Samuel David Epstein and Norbert Hornstein, volume 32, 251–282. Cambridge, MA: MIT Press.

van Urk, Coppe, and Norvin Richards. 2015. Two components of long-distance extraction: Successive

cyclicity in dinka. *Linguistic Inquiry* 46:113–155. URL https://muse.jhu.edu/article/566098.