

Improper case

Ethan Poole

University of California, Los Angeles

July 2020

This paper argues that case assignment is impossible in configurations that parallel generalized improper-movement configurations. Thus, like improper movement, there is “improper case”. The empirical motivation comes from (i) the interaction between case and movement and (ii) crossclausal case assignment in Finnish. I propose that improper case is ruled out by the *Ban on Improper Case*: a DP in [Spec, XP] cannot establish a dependent-case relationship with a lower DP across YP if Y is higher than X in the functional sequence. I show that this constraint falls under a strong version of the *Williams Cycle* (Williams 1974, 2003, 2013; van Riemsdijk and Williams 1981) and is derived under Williams’s (2003, 2013) analysis of embedding.

Keywords: case · Williams Cycle · locality · Finnish

1	Introduction	1
2	Background on dependent case	3
3	Movement and case	5
3.1	Some movement can feed dependent case	5
3.2	Some movement must not feed dependent case	7
3.3	Section summary	11
4	Finnish crossclausal case assignment	12
4.1	Background on Finnish case	13
4.2	Case in nonfinite clauses	17
5	Improper case	21
5.1	The Williams Cycle	22
5.2	Proposal	25
5.3	Application to Finnish	27
5.4	Application to movement	28
5.5	Remarks on Baker (2015)	34
6	Deriving the Williams Cycle	35
7	Discussion	38
7.1	Timing of case assignment	39
7.2	Phases	40
7.3	Potential exceptions to the Williams Cycle	41
	References	45
	Appendix: AGREE-based dependent-case assignment	53

1 Introduction

Some syntactic positions can be targeted by some movement types, but not by others. The classical example of this phenomenon is *HYPERRAISING*, whereby \bar{A} -movement can leave a finite clause (1), but A-movement cannot (2). The traditional analysis of hyperraising involves a conspiracy of two constraints: (i) movement out of a finite clause must proceed through the intermediate [Spec, CP] position (Chomsky 1973, 1977, 1981, 1986) and (ii) a ban on “improper movement”, according to which \bar{A} -movement but not A-movement may proceed from [Spec, CP] (Chomsky 1973, 1981; May 1979).

- (1) Who does it seem [CP has left]? (2) *Alex seems [CP has left].
-

A growing body of work has shown that these movement asymmetries are not limited to the binary distinction between A-movement and \bar{A} -movement (e.g. Williams 1974, 2003, 2013; Müller and Sternefeld 1993, 1996; Abels 2007, 2009, 2012a,b; Neeleman and van de Koot 2010; Müller 2014a,b; Keine 2016, 2019, 2020). Thus, there needs to be a more general theory of “improper movement” that restricts what movement types are available to what positions.

One particularly general and therefore interesting account of these asymmetries stems from Williams (1974, 2003, 2013) and van Riemsdijk and Williams (1981); I will refer to it as the *Williams Cycle* (WC).¹ The core analytical intuition behind the WC is that one and the same node is a barrier to some movement types, but not to others, and that this distinction correlates with the structural height of the landing site in the functional sequence. In Williams (2003), the WC is formulated as the *Generalized Ban on Improper Movement*, given in (3) (to be discussed in greater detail in section 5.1).

- (3) **GENERALIZED BAN ON IMPROPER MOVEMENT (GBOIM)**
 Movement to [Spec, XP] cannot proceed from [Spec, YP] or across YP, where Y is higher than X in the functional sequence. [based on Williams 2003]

The WC accounts for the ban on hyperraising (2) as a prohibition on moving from inside a CP to [Spec, TP]. According to the WC, CP is a barrier for movement to TP, but not for movement to CP, as schematized in (4), because C is higher than T in the functional sequence.

- (4) [CP \checkmark [TP \times ... [CP] DP ...
-

This account extends beyond hyperraising to the other kinds of movement asymmetries that have been documented in the literature. In addition to the work by Williams and van Riemsdijk and Williams, various versions of the WC have been developed by Abels (2007, 2009), Müller (2014a,b), and Keine (2016, 2019, 2020), amongst others.

¹ For discussion of the other approaches and arguments in favor of the WC, see Müller (2014a).

While the WC has traditionally been proposed on the basis of movement, [Keine \(2016, 2019, 2020\)](#) argues that analogous restrictions also govern agreement. This generalizing of the WC raises the question of whether other syntactic dependencies are also subject to the WC. This paper investigates the locality of case assignment and argues that it too is constrained by the WC.² Therefore, in line with movement and agreement, there is IMPROPER CASE.

The paper is couched in terms of dependent-case theory (DCT) ([Marantz 1991](#); [Bittner and Hale 1996](#); [McFadden 2004](#); [Baker 2015](#)). The reason for this choice is that the paper draws heavily on Finnish, which I will argue requires the notion of dependent case ([Poole 2015a](#); also [Maling 1993](#); [Anttila and Kim 2011, 2017](#)). However, the main arguments in this paper equally apply to functional-head case theory (FHCT) (e.g. [Chomsky 2000, 2001](#); [Legate 2008](#)); see fn. 12 and 33 in particular for discussion.

The motivation for improper case, i.e. that case assignment is subject to the WC, comes from two puzzles that the previous literature has not investigated in depth. The first puzzle involves the interaction between dependent case and movement, namely that some movement may feed dependent-case assignment, but other movement crucially must not do so. The second puzzle is crossclausal case assignment in Finnish, where a subject, but not an object may license dependent case on another DP across a nonfinite clause boundary. I show that both puzzles crucially do not fall under the purview of standard syntactic locality constraints, e.g. phases. I argue instead that both problems receive a unified analysis if case assignment is subject to the WC, which I formulate for case as the *Ban on Improper Case* in (5).

(5) **BAN ON IMPROPER CASE**

A DP in [Spec, XP] cannot establish a dependent-case relationship with a lower DP across YP, where Y is higher than X in the functional sequence.

According to the Ban on Improper Case, the heights of two DPs relative to one another on the functional sequence dictate whether they can establish a dependent-case relationship. For movement, (5) means that the height of a movement's landing site determines the range of positions from which another DP can license a dependent-case relationship with that moved DP. For clausal embedding, (5) means that the size of an embedded clause dictates which DPs in higher clauses can establish a dependent-case relationship across that clause boundary.

The Ban on Improper Case brings the locality of case into line with movement and agreement, in that the WC applies to all three. The question that follows then is how to uniformly derive WC effects in all three of these empirical domains. Crucially, improper case does not follow from recent proposals that analyze WC effects as the result of a constraint on AGREE or MERGE ([Abels 2007, 2009](#); [Müller 2014a,b](#); [Keine 2016, 2019, 2020](#)), because dependent-case assignment does not seem to involve either one of these operations. I argue that a unified analysis of WC effects for case, movement, and agreement becomes available if we adopt [Williams's \(2003\)](#) analysis of clausal embedding. [Williams](#) proposes that a ZP can only be embedded in a clause that has itself

² Note that throughout this paper, I am concerned with case and not nominal licensing (i.e. so-called "abstract Case").

been built up to ZP, which he calls the *Level Embedding Conjecture*. The crucial consequence of this proposal is that a root XP containing an embedded YP, where Y is higher than X in the functional sequence ($Y > X$), never exists in the course of a derivation (6).

(6) *Williams’s Level Embedding Conjecture*

- a. * $[_{XP} X^0 \dots [_{YP} \dots$ where $Y > X$
 b. $\checkmark [_{YP} Y^0 \dots [_{XP} \dots [_{YP} \dots$ where $Y > X$

Any movement, agreement, or case assignment between XP and YP that would violate the WC is in turn impossible because the relevant structure where X and elements in XP would have access to YP—under the strict cycle—is simply not created by the grammar, as schematized in (6a). Because this constraint follows from the way that syntactic structures are built, the key consequence of this account is that *all* syntactic dependencies are subject to the WC, regardless of whether they share the same operational core or not.

The argumentation proceeds as follows: Section 2 briefly overviews the paper’s assumptions about DCT. In sections 3 and 4, I present two locality puzzles for dependent-case assignment: the interaction of case and movement and Finnish crossclausal case assignment. To account for these two seemingly disparate locality problems, in section 5, I propose that dependent-case assignment is subject to the Ban on Improper Case. In section 6, I then argue that Williams Cycle effects for case, movement, and agreement can be uniformly analyzed in terms of clausal embedding. Section 7 concludes by discussing several purported exceptions to the Williams Cycle and further ramifications of the paper’s proposals.

2 Background on dependent case

In DCT, the calculus of case follows the algorithm in (7) (Marantz 1991; Bittner and Hale 1996; McFadden 2004; Baker 2015; and its predecessor Yip et al. 1987).

(7) *Case calculus in dependent-case theory*

1. Assign idiosyncratic lexical and inherent cases.
2. Take the remaining DPs. If DP_α c-commands DP_β , assign dependent case either to DP_α (\approx “ergative”) or to DP_β (\approx “accusative”). This directionality is parameterized.
3. If a DP was not assigned case in the previous two steps, then assign it unmarked case (\approx “nominative” and “absolute”).³

“Ergative” and “accusative”—in their canonical textbook definitions—are collapsed into the unified notion of DEPENDENT CASE. Whenever two DPs presently unvalued for case stand in a

³ Unmarked case is not necessarily morphophonologically unmarked, although there is a strong tendency for it to be. Furthermore, unmarked case is distinct from the notion of default case (see Marantz 1991; Schütze 2001).

c-command relationship in the same local domain, one of the DPs is assigned dependent case, though which one depends on the language’s parameterization. When the c-commanding DP is assigned dependent case, this corresponds to what would traditionally be called “ergative”. When the c-commanded DP is assigned dependent case, this corresponds to what would traditionally be called “accusative”. I will refer to this process as establishing a DEPENDENT-CASE RELATIONSHIP, and to the higher DP in the pair, i.e. the one that initiates the relationship, as the LICENSOR.

For the sake of concreteness, I adopt the syntactic implementation of DCT from Preminger (2011, 2014) throughout the paper: (i) DPs enter the derivation with an unvalued case feature, [CASE: \square], which can be valued as either DEP (for dependent case) or a particular lexical case.⁴ (ii) Lexical cases are assigned locally by (typically, lexical) heads, e.g. P⁰ and V⁰, to their sibling upon first-merge (8).⁵

$$(8) \quad [PP/VP/XP \quad P^0/V^0/X^0 \quad \underline{DP}_{[CASE: LEX]}] \quad \text{where LEX = the relevant lexical case}$$

The diagram shows a head $P^0/V^0/X^0$ c-commanding its complement $\underline{DP}_{[CASE: LEX]}$. A horizontal line connects the head and the DP, with a vertical line from the head pointing down to the line, and another vertical line from the DP pointing up to the line, forming a U-shape.

(iii) Dependent case is assigned whenever two DPs with unvalued case features ([CASE: \square]) stand in a c-command relationship (9).⁶ The language’s parameterization determines whether it is the higher or lower DP in the pair that gets assigned dependent case; throughout the paper, I indicate the assignee with an underline. Concerning the timing of dependent case, in section 3.2, I will argue that dependent-case relationships are established as early as possible, which I take to be upon merging the licensor into the c-commanding position. The morphological exponence of dependent case, e.g. as “accusative” and “ergative”, is determined at PF.

$$(9) \quad [\underline{DP}_{[CASE: \square]} \dots [\dots \underline{DP}_{[CASE: \square]} \dots]] \begin{cases} \rightarrow [DP_{[CASE: \square]} \dots [\dots \underline{DP}_{[CASE: DEP]} \dots]] \\ -or- \\ \rightarrow [\underline{DP}_{[CASE: DEP]} \dots [\dots DP_{[CASE: \square]} \dots]] \end{cases}$$

The diagram shows a c-command relationship between two DPs: $\underline{DP}_{[CASE: \square]}$ c-commands $[\dots \underline{DP}_{[CASE: \square]} \dots]$. A horizontal line connects them, with vertical lines from each DP pointing to the line. From the right DP, two arrows branch out: one pointing to $[DP_{[CASE: \square]} \dots [\dots \underline{DP}_{[CASE: DEP]} \dots]]$ and another pointing to $[\underline{DP}_{[CASE: DEP]} \dots [\dots DP_{[CASE: \square]} \dots]]$. The word “-or-” is placed between the two arrows.

(iv) If [CASE: \square] is still unvalued when Spellout occurs, it is realized as unmarked case at PF (10). Thus, unmarked case (\approx “nominative” and “absolutive”) is the absence of any otherwise assigned case (see Kornfilt and Preminger 2015).

$$(10) \quad [CASE: \square] \leftrightarrow \text{UNMARKED CASE}$$

These detailed mechanics will be abstracted over when not relevant to the discussion at hand. One advantage of this syntactic case calculus is that the structure consisting of a head and the DP that it c-selects is necessarily built before any larger structure containing that DP and

⁴ For the sake of simplicity, I collapse the distinction between lexical and inherent case (Woolford 2006).

⁵ Following Bare Phrase Structure, where what projects is the head itself (Chomsky 1995a), lexical case can also be assigned in a specifier–head relation as siblinghood agreement (à la Rezac 2003).

⁶ Baker (2015) proposes several enhancements to dependent-case assignment, including reciprocal relationships (where both DPs are assigned dependent case; see also Deal 2015), null relationships (where neither DP gets dependent case), and ‘keying’ dependent-case rules to particular domains. These enhancements are compatible with the proposals in this paper, but are not directly relevant to the matters at hand, so I have set them aside.

another DP in a c-command relationship. Therefore, the precedence relations in (7) fall out intrinsically based on how structure is built, and do not need to be stipulated, as, e.g., the original implementation in [Marantz \(1991\)](#) does.

Before proceeding, there are two important points about DCT worth emphasizing: First, dependent-case assignment is *in addition* to case assignment by designated heads (which in the terms here, falls under lexical case).⁷ As such, the expressive power of DCT is a proper superset of the expressive power of FHCT ([Preminger 2017](#)). The crucial question then is whether the additional expressive power of DCT is warranted, i.e. whether there are case patterns that call for the notion of dependent case. Several such patterns have been identified in the literature (see e.g. [Marantz 1991](#); [Baker and Vinokurova 2010](#); [Baker 2014, 2015](#); [Baker and Bobaljik 2017](#); [Jenks and Sande 2017](#); [Yuan 2018, 2020](#)), and in this paper, I argue that the distribution of Finnish accusative case is another such pattern. Second, the descriptive label given to a case should not be taken to entail a particular case-assignment mechanism. That is, just because a given case in a given language is pretheoretically called “accusative” or “ergative” does not mean that it is necessarily dependent case—and likewise for “nominative” or “absolute” and unmarked case (hence the scare quotes).

3 Movement and case

This section shows that some movement may lead to dependent-case assignment, but other movement must not do so. This dichotomy will be shown not to follow from standard conceptions of locality, e.g. phases, and thus it presents a challenge for DCT.

3.1 Some movement can feed dependent case

The DCT literature has identified a number of case patterns as involving movement feeding dependent-case assignment. Let us consider three representative examples.

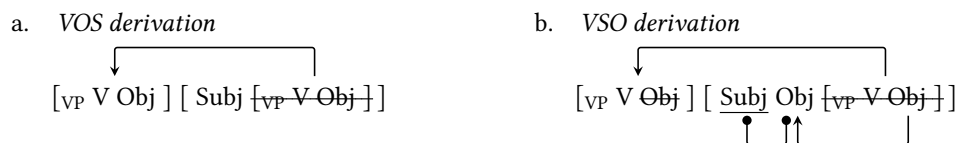
The first example involves object shift: a dependent-case relationship between the subject and the object is allowed only if the object has raised out of VP ([Bittner and Hale 1996](#); [Baker 2015:125–130](#); [Woolford 2015](#)). This pattern is illustrated in (11) with Niuean, where the case markings correlate with both the specificity of the object and the clausal word order ([Massam 2000, 2001](#)). If the object is nonspecific, the subject is nominative, and the clause has VOS word order (11a). If the object is specific, the subject is ergative and the clause has VSO order (11b).

⁷ Traditionally, in DCT, lexical-case assignment (i.e. case assignment by heads) is considered to be very local. Under [Preminger’s \(2011, 2014\)](#) system, which I am adopting here, it is restricted to siblinghood. However, [Baker \(2015\)](#) proposes that lexical case (though he does not call it such) can in fact be assigned under closest c-command (see also [Baker and Vinokurova 2010](#); [Preminger 2017](#)). Because lexical case does not factor into the arguments in this paper, I will assume the traditional, more restrictive view of lexical case.

- (11) a. *Nonspecific object* → VOS, *nominative subject* [Niuean]
 Ne inu [kofe] [a Sione]
 PAST drink coffee NOM Sione
 ‘Sione drank coffee’
- b. *Specific object* → VSO, *ergative subject*
 Ne inu [e Sione] [e kofe]
 PAST drink ERG Sione NOM coffee
 ‘Sione drank the coffee’ [Massam 2000:98]

Massam (2000, 2001) analyzes this alternation in terms of object shift and VP fronting. The VOS order is produced by the object remaining in its base position and thus fronting along with the VP (12a). By contrast, the VSO order is produced by the object raising out of the VP prior to VP fronting (12b). The additional correlation with ergative case is then captured by assuming—in pretheoretical terms—that ergative case requires that the object leave VP.

- (12) *Massam’s (2000, 2001) analysis of Niuean word order*



In DCT terms, VP blocks dependent-case assignment in Niuean (see section 7.1). Thus, only when the object raises out of VP can a dependent-case relationship between the subject and the object be established. Other languages exhibiting object shift feeding dependent case include Dyirbal, Eastern Ostyak, Ika, Inuit, Nez Perce, Sakha, and Tagalog (see Baker and Vinokurova 2010; Baker 2015:125–130; Woolford 2015; and references therein).

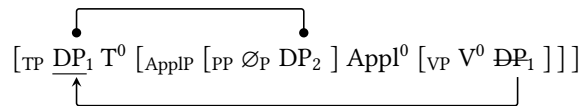
The second example is Shipibo applicatives of unaccusatives (Baker 2014). Baker shows that in Shipibo, an unaccusative subject is ordinarily nominative (13a), but adding an applicative argument causes the subject to become ergative (13b). He argues that ergative in Shipibo is dependent case and that in (13b), the subject is ergative because the applicative provides the additional DP needed for dependent case.

- (13) a. **Kokoti**-ra joshin-ke [Shipibo]
 fruit-PCL ripen-PFV
 ‘The fruit ripened’
- b. **Bimi-n**-ra **Rosa** joshin-xon-ke
 fruit-ERG-PCL Rosa ripen-APPL-PFV
 ‘The fruit ripened for Rosa’ [Baker 2014:345–346]

However, if the unaccusative subject (i.e. the theme) is base-generated inside VP and the applicative is base-generated above VP—both standard assumptions—, then it should be the applicative

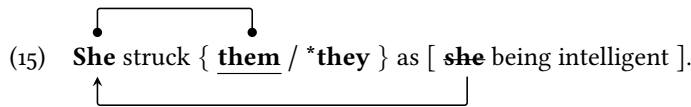
that gets dependent ergative. That is, upon merging into the structure, the applicative should establish a dependent-case relationship with the theme; in the pair, the applicative would be the higher DP and hence should be the one assigned dependent case (= ergative). Baker (2014) handles this problem by proposing that the applicative is encased inside a null PP, so that (i) it does not c-command the theme and (ii) it is ineligible for movement to subject position, [Spec, TP]. Due to the latter, the theme is able to raise over the applicative to [Spec, TP]. From [Spec, TP], the theme c-commands the applicative and establishes a dependent-case relationship with it; in this version of the pair though, the theme is the higher DP and thus gets assigned dependent case. This analysis is schematized below in (14).

(14) Baker's (2014) analysis of Shipibo applicatives of unaccusatives



Assuming that Baker's (2014) analysis is on the right track, Shipibo applicatives of unaccusatives are an instance of movement feeding dependent-case assignment.⁸

The third example is the English raising predicate *strike as* (Marantz 1991), which is traditionally analyzed as the matrix subject starting out as the embedded subject and raising into matrix-subject position, as schematized in (15). Taking accusative on objects in English to be dependent case, the licensor of dependent case on the internal argument of *strike* must be the subject after it has raised to matrix [Spec, TP], because there is no other possible licensor.



Note that something needs to be said about why the internal argument of *strike* does not license dependent case on the subject before it has raised; I will return to this point in section 5.4.

In sum, taken together, these three examples crucially show that there are instances of movement that feed dependent-case assignment.

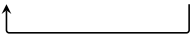
3.2 Some movement must not feed dependent case

While the previous section showed that some movement may feed dependent-case assignment, there is also other movement that *must not* feed dependent-case assignment. I will illustrate this

⁸ Nez Perce also has applicatives of unaccusatives, which behave similarly to Shipibo (Deal 2019). Deal argues though that the applicative does not raise to [Spec, TP] because of antilocality, rather than an encasing null PP, as Baker (2014) proposes. Deal's analysis is in principle compatible with the proposals in this paper, but it would be necessary to find another explanation for why the applicative is not assigned dependent case before the theme moves.

problem using *wh*-movement, though it holds generally for nonlocal movement, and I will use English for ease of illustration.⁹ Let us take the problem in two parts.

The first part of the problem is that dependent case cannot be assigned based on the surface structure alone. For example, the structure in (16a) with *wh*-movement must be mapped to the string in (16b) and cannot be mapped to (16c). Descriptively, dependent case needs to be calculated *before wh*-movement has occurred.

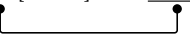
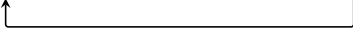
- (16) a. Who did she see ~~who~~?

 b. **Who(m)**_{DEP} did **she**_{NOM} see?
 c. ***Who**_{NOM} did **her**_{DEP} see?

One potential solution that can be immediately set aside is to assume that case is assigned at PF and that *wh*-movement ‘reconstructs’ for case at PF. This solution would face the problem that unlike canonical reconstruction at LF, this hypothetical PF reconstruction would have to be for case alone and not uniformly for all PF processes, in particular not for linearization. As such, it would be nothing more than a restatement of the empirical generalization that *wh*-movement does not affect case. Rather, I propose that dependent-case assignment is interspersed with structure building, so that dependent case is assigned as early as possible in the derivation (see also Baker and Vinokurova 2010:604; Preminger 2011, 2014). I call this principle *Earliness* (17).

(17) **EARLINESS**

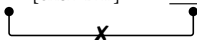
Upon (re)merging α into the structure, if α c-commands β and both α and β have unvalued case features, establish a dependent-case relationship between α and β .

Earliness crucially forces dependent-case assignment to happen prior to *wh*-movement. The derivation of (16) under this analysis is illustrated in (18): First, a dependent-case relationship is established between *she* and *who* immediately upon first-merge of *she* into the structure (18a). Second, *wh*-movement happens later in the derivation, after dependent case has been assigned (18b). Third, at PF, [CASE: DEP] is realized as “accusative” and [CASE: \square] as “nominative” (18c).


- (18) a. **she**_[CASE: \square] saw **who**_[CASE: DEP]? *Assign dependent case*

 b. **who**_[CASE: DEP] did she_[CASE: \square] see **who**_[CASE: DEP]? *Move wh-element*

 c. **who**_{ACC} did she_{NOM} see **who**_{ACC}? *PF: Realize case features*

⁹ Note that I am not claiming that *whom* is an accusative or dependent-case form in English (see Lasnik and Sobin 2000). I am using *who/whom* merely for ease of presentation, but the key empirical properties generalize to languages in which *wh*-elements have visible dependent case.

The formulation of Earliness in (17) also reiterates the restriction that two DPs can enter into a dependent-case relationship only if they both presently have unvalued case features ([CASE: □]); this is part of the case calculus laid out in section 2. This restriction prevents DPs with dependent case from reparticipating in dependent-case assignment, either after having themselves moved or with DPs that have moved above them. For example, in (19), this restriction prevents the moved *wh*-element that has itself been assigned dependent case from turning around and licensing dependent case on the subject from the higher position to which it has *wh*-moved.

(19) **who**_[CASE: DEP] did **she**_[CASE: □] see **who**_[CASE: DEP]?


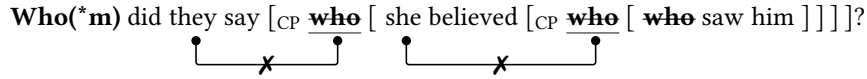
While Earliness is a necessary component to solving the problem imposed by *wh*-movement, it is not sufficient. This brings us to the second part of the problem: dependent-case assignment in the context of successive-cyclic movement. When a *wh*-element that is itself unvalued for case—and should surface with unmarked case at PF—moves successive cyclically, it passes through intermediate [Spec, CP] positions from where it *should* in principle affect the calculus of dependent case, but does not. (For the moment, let us set aside the possibility of successive-cyclic movement through [Spec, *v*P] until section 7.2.) Consider (20), where *who* undergoes successive cyclic *wh*-movement to matrix [Spec, CP] and must surface with unmarked case.

(20) **Who**(*m) did they say [_{CP} **who** [she believed [_{CP} **who** [**who** saw him]]]]?


Sentences like (20) present two complications for DCT. For convenience, I will discuss these complications in terms of an accusative alignment, where the lower DP in a dependent-case pair is assigned dependent case, but the problem extends to an ergative alignment too, where the higher DP in the pair is assigned dependent case. The first problem is that a *wh*-element does not have its own case altered from its intermediate landing sites.^{10,11} From these intermediate positions, there may very well be another DP unvalued for case that c-commands the *wh*-element. All else equal, the *wh*-element should be assigned dependent case in such configurations—but crucially, it is not. Descriptively, the moving *wh*-element cannot have the case overwritten that would have been assigned to it if it had not moved. In DCT terms, the *wh*-element cannot be the *lower* DP in a dependent-case pair when it is in an intermediate landing site, as schematized in (21). As such, I will refer to this problem as the *Lower-DP Problem*.

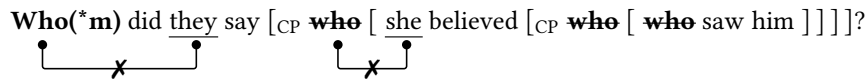
¹⁰ A *wh*-element also cannot enter into a dependent-case relationship in its final landing site in an embedded question, but this instantiates the same relevant configuration as an intermediate landing site.
¹¹ This is not to imply that there cannot be a dedicated lexical case for \bar{A} -moved elements, e.g. as is found in Dinka (van Urk 2015). Because the assignment of such case is not contingent on the presence of another DP, it does not qualify as dependent case and thus falls outside the purview of the present discussion. However, for discussion of movement and lexical case, see section 5.4.

(21) *Lower-DP Problem*



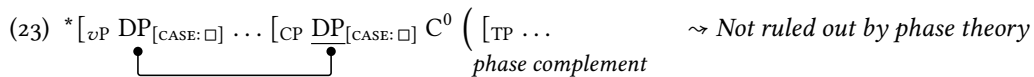
The second problem is that a *wh*-element does not alter the case of other DPs from its intermediate or final landing sites. From these positions, the *wh*-element may very well c-command another DP unvalued for case, and thus it should, all else equal, be able to license dependent case on it—but it cannot do so. In other words, the moving *wh*-element cannot be the *higher* DP in a dependent-case pair (modulo from its base-generated position with, e.g., an object). As such, I will refer to this problem as the *Higher-DP Problem* (22).

(22) *Higher-DP Problem*



The standard dependent-case calculus does not offer an explanation for why successive-cyclic movement does not affect case in these two ways (for a discussion of Baker 2015, see section 5.5). Crucially, in light of the data in section 3.1, it would not suffice to simply stipulate that movement does not affect case assignment. Thus, a more nuanced account is called for.

Neither of the problems that successive-cyclic movement raises for DCT fall under the purview of phases (or its predecessor, subjacency). First, because phase edges remain accessible at the next highest phase, per the Phase Impenetrability Condition (Chomsky 2000, 2001), the locality enforced by phases permits precisely the configurations that give rise to the Lower-DP Problem, as schematized in (23). In other words, a DP unvalued for case may c-command the edge of the lower phase, thereby satisfying the criteria for establishing a dependent-case relationship with a DP in that edge position, if it too is unvalued for case.¹²



Second, because movement to [Spec, CP] takes place before phasal Spellout, such movement should, all else equal, be able to affect the case of elements in the CP-phase domain.¹³ Otherwise, establishing any relation between the phase edge and the phase complement would be impossible, and such relations are minimally necessary for movement dependencies. Thus, the Higher-DP Problem is also not solved under the locality afforded by phases. Note that I am *not* claiming that these considerations provide evidence against phases; rather, the point is that they do not follow from phase theory itself.

¹² The same problem is faced by FHCT if caseless DPs are permitted to move through a phase edge or if one assumes that nominative is unvalued case or that there is case stacking. A *v*⁰ head could then assign accusative case to a DP in a phase-edge position, e.g. an intermediate [Spec, CP] position, thereby making the same incorrect prediction that DCT does in (21).

¹³ Baker (2015), however, proposes an analysis that stipulates that material at the phase edge cannot affect the case of elements inside the phase domain; see section 5.5 for discussion.

In sum, successive-cyclic movement leads to the generalization that some movement crucially must *not* feed dependent-case assignment.

3.3 Section summary

Some, but not all movement affects dependent-case assignment. Assigning dependent case as early as possible, Earliness (17), already filters out many of the undesirable interactions between dependent case and movement. For example, in a simple transitive clause, moving an object over a subject will have no effect on dependent case because a dependent-case relationship will have already been established between the two DPs upon first-merge of the subject. This was shown for *wh*-movement in section 3.2, but it holds for A-movement as well, in particular for A-scrambling. The question then is about all the interactions that Earliness does *not* capture. These interactions include some instances of A-movement that can feed dependent case, but no such instances of \bar{A} -movement. It also includes some instances of \bar{A} -movement that cannot affect dependent case, but which are not ruled out by Earliness, e.g. with successive cyclicity. This state of affairs is summarized with the generalization in (24).¹⁴

(24) **MOVEMENT–CASE GENERALIZATION**

A-movement can feed dependent case, but \bar{A} -movement cannot.

To the best of my knowledge, (24) appears to be without a clear exception, though there are three phenomena that warrant discussion. Two of these phenomena, I discuss later: Koryak in section 5.4 and Sakha in section 7.3. The third phenomenon is Hungarian long focus movement (which includes *wh*-movement): when an embedded nominative element is long-focused, it can surface as accusative if the matrix predicate is transitive, i.e. could itself assign accusative, as shown in (25) (e.g. Massam 1985; É. Kiss 1987; Gervain 2009; den Dikken 2009, to appear; Jánosi 2013; Jánosi et al. 2014). This pattern is (mostly) general: a long-focused element can bear the case associated with the embedded clause (25a) or the matrix clause (25b) (Jánosi 2013).¹⁵

(25) *Hungarian long focus movement*

- a. *Long-focused element with ‘embedded’ case*

AUTÓ₁ mondta [CP hogy ___₁ áll a kapunál]
 car.NOM said.3SG.DEF that stand.3SG the gate.ADE
 ‘He said that there was a CAR standing at the gate’

- b. *Long-focused element with ‘matrix’ case*

AUTÓ-T₁ mondott [CP hogy ___₁ áll a kapunál]
 car-ACC said.3SG.INDEF that stand.3SG the gate.ADE
 ‘He said that there was a CAR standing at the gate’

[Jánosi 2013:57]

¹⁴ It should be noted that (24) is distinct from the traditional GB definition of A-movement as movement to receive case. In DCT, there is never a need to *move* to receive case; case can always be received in situ. Rather, (24) embodies a very different kind of interaction between movement and case, one particular to dependent case.

¹⁵ Lexical cases display somewhat different behavior in Hungarian long focus movement; see Jánosi (2013).

Massam (1985) analyzes this pattern as case being assigned to the long-focused element in the intermediate [Spec, CP] position that it \bar{A} -moves through; if true, it would be an exception to (24) (assuming that accusative in Hungarian is dependent case).^{16,17} However, in the literature on Hungarian, it has been argued for reasons entirely independent of considerations like (24) that this alternation actually stems from two distinct structures (den Dikken 2009, to appear; Jánosi 2013; Jánosi et al. 2014). When the long-focused element bears embedded case, as in (25a), there is genuine long movement: the element is base-generated in the embedded clause, is assigned case locally, and \bar{A} -moves into the matrix-focus position. By contrast, when the long-focused element bears matrix case, as in (25b), it is base-generated in the *matrix* clause, is assigned case locally, locally \bar{A} -moves to the matrix focus position, and is indirectly linked to the embedded gap via resumption. Discussing the arguments in favor of this analysis would take us too far afield; the reader is referred to Jánosi (2013). Crucially for the purposes of this paper though, under this independently-motivated analysis of Hungarian long focus, it is not an exception to (24), as there is no \bar{A} -movement feeding dependent-case assignment.

Explaining (24) requires a way of teasing apart movement types. In minimalist syntax, because there is only a single primitive movement operation (i.e. MERGE), there is no principled way to distinguish A-movement and \bar{A} -movement. A goal of this paper is thus to *derive* the locality constraint in (24) without reference to separate primitives for A-movement and \bar{A} -movement. In the next section, I show a pattern from Finnish crossclausal case assignment that also does not follow from any binary notion of locality, e.g. phases. Despite not involving movement, this pattern will be shown to parallel movement configurations that are accounted for under the Williams Cycle. I will argue that adopting the Williams Cycle as a constraint on dependent-case assignment, in the form of the Ban on Improper Case, provides a unified account of both crossclausal case assignment in Finnish and the Movement–Case Generalization in (24).

4 Finnish crossclausal case assignment

This section shows that in Finnish, dependent case may be licensed across a nonfinite clause boundary, but only by a subject and not by an object (or an adjunct). As with movement,

¹⁶ Kayne (1981, 1984) and Rizzi (1982) propose a similar derivation for French and Italian ECM, which is possible only if the embedded subject undergoes \bar{A} -movement. On their analyses, the embedded subject is licensed in [Spec, CP] by the matrix predicate. Though they call this relation “Case”, no actual case is involved. See Koopman and Sportiche (2014) for a recent reanalysis of these facts in French that does not involve crossclausal movement.

¹⁷ Kayne (1984) also proposes a similar derivation for the use of *whom* in sentences like (i.a) (acceptable for some English speakers), where *whom* corresponds to a subject position and thus should be nominative *who*. On his analysis (in FHCT), the higher predicate assigns accusative to *who(m)* in an intermediate [Spec, CP] position. However, Lasnik and Sobin (2000) argue that this is a mischaracterization of the data. In particular, *whom* can still appear if the higher predicate is passive and hence would be unable to assign accusative (irrespective of FHCT or DCT), as shown in (i.b). See Lasnik and Sobin (2000) for an alternative analysis of *whom*.

(i) a. the man [_{CP} whom₁ [I believe [_{CP} ___₁ has left]]]

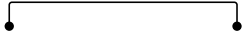
b. the man [_{CP} whom₁ [it was believed [_{CP} ___₁ had left]]]

[Lasnik and Sobin 2000:358]

this dichotomy will be shown not to fall under the purview of standard conceptions of locality, e.g. phases. The section begins in subsection 4.1 with some background on Finnish structural case and arguments that accusative in Finnish is dependent case, and then continues in subsection 4.2 with the crucial case patterns in embedded nonfinite clauses.

4.1 Background on Finnish case

Finnish has three structural cases: nominative, accusative, and partitive.¹⁸ For the sake of simplicity, I set aside partitive case and focus on the distribution of nominative and accusative (for a more comprehensive analysis, see Poole 2015a). In a simple transitive clause, the external argument is nominative and the internal argument is accusative (26). To simplify the exposition, let us refer to the external argument as the ‘subject’ and the internal argument as the ‘object’. Whenever the subject is absent, e.g. in a passive (27a) or in an imperative (27b), or the subject bears lexical case (i.e. a quirky subject) (27c), the object is nominative.^{19,20}

- | | | |
|---------|---|---------|
| (26) |  | NOM-ACC |
| | Pekka.NOM bought book-ACC | |
| | ‘Pekka bought the/a book’ | |
| (27) a. | Kirja oste-ttiin | NOM |
| | book.NOM buy-PASS.PAST | |
| | ‘The book was bought’ / ‘People bought the book’ | |
| b. | Osta kirja! | NOM |
| | buy.IMP book.NOM | |
| | ‘Buy the/a book!’ | |
| c. | Minu-n täytyy osta-a kirja | GEN-NOM |
| | I-GEN need buy-INF/TA book.NOM | |
| | ‘I have to buy the/a book’ | |

¹⁸ The status of accusative case in Finnish is somewhat contentious. Under traditional analyses, accusative comprises three forms: one homophonous with genitive, one homophonous with nominative, and one distinct form for human pronouns. Kiparsky (2001) argues that only the form for human pronouns is a genuine accusative case (see also e.g. Penttilä 1963; Timberlake 1975; Milsark 1985; Taraldsen 1986; Mitchell 1991; Maling 1993; Toivainen 1993; Vainikka 1993; Nelson 1998). This paper assumes a simplified picture: the genitive-homophonous accusatives are referred to as “accusative”, the nominative-homophonous accusatives are referred to as “nominative”, and the pronouns are set aside. This is in line with what Kiparsky (2001) argues, but with the terminology shifted to parallel the standard nominative-accusative pattern. This choice has no bearing on the claims made in this paper; the dependent case in Finnish is marked with *-n* regardless of whether one calls that form “accusative” or “genitive”.

¹⁹ Some notes on the Finnish data: Unless indicated otherwise, Finnish judgments are due to my informants. Glossing conventions have been unified across sources. To simplify examples, I do not gloss verbal agreement and have removed any instances of *pro*-drop (*pro* behaves just like an overt DP for the purposes of case assignment). The Finnish case patterns in this paper are all invariant (modulo that some of the objects can be partitive), e.g. if a DP is nominative in an example, it must be nominative in that position and cannot be accusative.

²⁰ For imperatives, this is only true if they are first and second person. See Nelson (1998:95-97) and Kiparsky (2001) for arguments that these imperatives do not have syntactically active subjects.

The case patterns exemplified in (26) and (27) receive a straightforward explanation under DCT. In (26), the subject licenses dependent case (= accusative) on the object; then, because there is no other DP that *c*-commands the subject, the subject remains unvalued for case throughout the derivation and is realized as having unmarked case (= nominative) at PF. In (27a) and (27b), there is no other DP that *c*-commands the object; as such, no dependent-case relationship is established, and the object is realized as having unmarked case. In (27c), although there is another DP that *c*-commands the object, it bears lexical genitive case. Recall from section 2 that only DPs unvalued for case factor into the calculus of dependent case. Lexically case-marked DPs are thus invisible to dependent-case assignment because their case will already have been assigned locally. Accordingly, because no other DP with *unvalued* case *c*-commands the object in (27c), the object remains unvalued for case and is realized as having unmarked case at PF. This analysis is summarized in (28).²¹

(28) *Finnish structural case*

Nominative is unmarked case ([CASE: □] ↔ NOM), and accusative is dependent case ([CASE: DEP] ↔ ACC).

The data in (26) and (27) could alternatively be analyzed in FHCT: the variants of v^0 in (27) would lack the ability to assign accusative case, so that T^0 could assign nominative case to the object (e.g. [Vainikka and Brattico 2014](#), though the identity of the heads differs on their account). Such an analysis would amount to a standard implementation of Burzio’s Generalization. Evidence that such an FHCT analysis is insufficient comes from adjuncts. In Finnish, there is a special class of adjuncts that are structurally case-marked, akin to subjects and objects ([Tuomikoski 1978](#); [Maling 1993](#)). These adjuncts include durational adjuncts (*for an hour*), spatial-measure adjuncts (*a kilometer*), and multiplicative adjuncts (*two times*). In DCT terminology, these adjuncts factor into the calculus of dependent case—i.e. they can license and be assigned dependent case—, and they are realized with unmarked case if their case remains unvalued in the derivation. To illustrate, in an intransitive clause with one of these adjuncts, the subject is nominative and the adjunct is accusative (29a). When the intransitive predicate is passivized (as some kind of impersonal passive), the adjunct becomes nominative (29b), the same case alternation that is observed for objects in passives (27a).²²

- (29) a.  **Minä** opiskelin [**vuode-n**]_{ADJUNCT} NOM-ACC
 I.NOM studied year-ACC
 ‘I studied for a year’

²¹ The notion that DPs compete for nominative case also underlies the analyses of Finnish case in [Maling \(1993\)](#), [Anttila and Kim \(2011, 2017\)](#), and [Poole \(2015a\)](#), though the implementations differ considerably as a result of using different frameworks. The core insights about Finnish and improper case in this paper could in principle be expressed using any of these analyses.

²² Further evidence that this class of adjuncts is structurally case-marked comes from the fact that they must be partitive when under the scope of negation, like subjects and objects (e.g. [Heinämäki 1984](#); [Kiparsky 2001](#)).

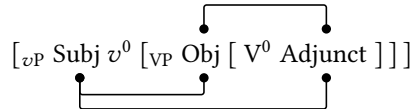
- b. Opiskel-tiin [vuosi]_{ADJUNCT} NOM
 study-PASS.PAST year.NOM
 ‘People studied for a year’ [Kiparsky 2001:323]

With a transitive predicate, where the object does not bear lexical case, structurally case-marked adjuncts are always accusative (30).

- (30) a. Liisa muisti matka-n [vuode-n]_{ADJUNCT} NOM-ACC-ACC
 Liisa.NOM remembered trip-ACC year-ACC
 ‘Liisa remembered the trip for a year’
- b. Liisa-n täytyy muista-a matka [vuode-n]_{ADJUNCT} GEN-NOM-ACC
 Liisa-GEN need remember-INF/TA trip.NOM year-ACC
 ‘Liisa has to remember the trip for a year’ [Maling 1993:57]

Following Lechner (2003), I will assume that the *vP* is right-branching, where adjuncts are *c*-commanded by the object (see also Csirmaz 2005:90–98, who also argues for such an analysis for Finnish). This is schematized in (31), where the possible dependent-case relationships are indicated. Accordingly, an object that is not assigned lexical case by the verb will invariably license dependent case on an adjunct, thereby accounting for the pattern in (30).²³

(31) *Right-branching vP structure in Finnish*



Crucially, clauses like (29b), where the adjunct is nominative, may contain *multiple* structurally case-marked adjuncts. In such configurations, the DCT analysis and the FHCT analysis make different predictions. The DCT analysis predicts that the highest adjunct is nominative and all the other adjuncts are accusative. The FHCT analysis, on the other hand, predicts that all of the adjuncts are nominative, because the functional head responsible for assigning accusative case is absent in clauses where the subject is absent; this is what accounted for the data in (27) under an FHCT analysis. The data bear out the prediction of the DCT analysis. This is shown in (32) with two structurally case-marked adjuncts and the verb *luottaa* ‘trust’, which assigns lexical illative case to its object, thereby removing it from the calculus of dependent case. When the subject is present, both of the adjuncts are accusative (32a). When the subject is absent, here

²³ Given the *vP*-structure in (31), more needs to be said about how *V*⁰ assigns lexical case to the object. Under Bare Phrase Structure, a head *X* reprojects and is a sibling with its specifier, so the locality conditions of lexical-case assignment laid out in section 2 are still satisfied with the structure in (31). Nevertheless, there needs to be something preventing *V*⁰ from assigning the lexical case to the adjunct instead of the object. As our concern in this paper is dependent case, I leave this problem for future research.

in a passive, the higher adjunct is nominative and the lower adjunct is accusative (32b).²⁴ Finally, when the first adjunct is dropped, the only remaining adjunct becomes nominative (32c).

- (32) a. *Subject* → *NOM*, *Durational* → *ACC*, *Multiplicative* → *ACC*

Minä luotin [Kekkose-en]_{LEX} [yhde-n vuode-n] [kolmanne-n kerra-n]
 I.NOM trusted Kekkonen-ILL one-ACC year-ACC third-ACC time-ACC
 ‘I trusted Kekkonen for a year for a third time’

- b. *Durational* → *NOM*, *Multiplicative* → *ACC*

[Kekkose-en]_{LEX} luote-ttiin [yksi vuosi] [kolmanne-n kerra-n]
 Kekkonen-ILL trust-PASS.PAST one.NOM year.NOM third-ACC time-ACC
 ‘Kekkonen was trusted for a year for a third time’

- c. *Multiplicative* → *NOM*

[Kekkose-en]_{LEX} luote-ttiin [kolmas kerta]
 Kekkonen-ILL trust-PASS.PAST third.NOM time.NOM
 ‘Kekkonen was trusted for a third time’

[Maling 1993:59]

The pattern in (32) follows in the DCT analysis without further ado. For example, in (32b), the first adjunct licenses dependent case on the second adjunct; then, because no relevant DP c-commands the first adjunct, it remains unvalued for case in the derivation and is realized with unmarked case at PF. The FHCT analysis, on the other hand, would need to make additional stipulations to account for (32), in particular to deal with (32b), in which accusative would have to be assigned in a passive, where the functional head responsible for accusative would not occur (similarly in (30b)). As far as I am aware, there is no FHCT analysis of Finnish case that extends to the case pattern in (32).²⁵ I take the fact that this adjunct pattern is entirely regular and productive in Finnish to indicate that Finnish *requires* the notion of dependent case in order to capture the distribution of accusative (Poole 2015a; see also Maling 1993; Anttila and Kim

²⁴ (32) raises the question of the hierarchical positions of structurally case-marked adjuncts with respect to one another. As might be expected, the linear order is unrevealing: the order of the two adjuncts in (32) can be reversed, and still the durational adjunct will be nominative and the multiplicative adjunct accusative. Maling (1993) reports the following preference amongst the adjuncts for being nominative: spatial-measure > durational > multiplicative. She captures this preference in terms of a grammatical-function hierarchy, but it can also be captured in terms of structure: multiplicative adjuncts are merged before durational adjuncts, which are merged before spatial-measure adjuncts. Because of Earliness (17), the base-generated order of the adjuncts will correctly dictate their case assignment—all happening inside the *vP*—and any subsequent reordering of these adjuncts (via movement to *vP*-external positions) will not change the case assignment amongst the adjuncts (or with respect to the subject and object, which are introduced in *vP* in positions that c-command the adjuncts; see (31)).

²⁵ This problem is true of the analyses in Vainikka (1989), Brattico (2012), and Vainikka and Brattico (2014). Nelson (1998) is able to account for the basic pattern in (32); she proposes that nominative case *must* be assigned in every finite clause, essentially the Inverse Case Filter of Bošković (1997, 2002). However, her analysis does not explain the possibility of impersonal passives of intransitive predicates, which have no arguments that could receive nominative case, e.g. *Tanssittiin* ‘There was danced’. Her account also does not extend to nonfinite clauses; see section 5.3 for discussion. Space limitations unfortunately prevent giving an exposé of these alternative accounts.

2011, 2017). I will thus adopt such an account in what follows. Against this backdrop, let us now consider case assignment in nonfinite clauses.

4.2 Case in nonfinite clauses

Finnish has a number of nonfinite constructions (Vainikka 1989, 1995; Toivonen 1995; Koskinen 1998; also Hakulinen et al. 2004:§490). The nonfinite construction of interest in this paper is the MA-infinitive (traditionally called the “third” infinitive). The reason that MA-infinitives are interesting is because when they function as clausal complements, case assignment within the nonfinite clause interacts with the makeup of the clause of the embedding verb (e.g. Vainikka 1989). That is, the matrix (= embedding) and embedded clauses constitute a single coextensive domain for the purposes of dependent-case assignment.

The MA-infinitive requires the verb to bear an inner locative case marker (inessive, elative, or illative) after the infinitival morpheme *-mA* (33).²⁶ The case marker matches what a DP would bear in that same position, with the same “directional” meaning (33). In this sense, the verb in a MA-infinitive is nominal-like, but unlike a genuine nominal, it cannot be modified by nominal modifiers, only verbal modifiers (34).

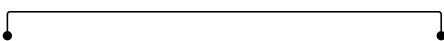
(33) *Occurs with a locative case marker*

Minä autoin Jukka-a { [TP kirjoitta-ma-an Marja-lle] / **bussi-in** }
 I.NOM helped Jukka-PTV write-INF/MA-ILL Marja-ALL bus-ILL
 ‘I helped Jukka { to write to Marja / onto the bus }’ [based on Koskinen 1998:329]

(34) *Can only occur with verbal modifiers*

Minä autoin Jukka-a [TP asettu-ma-an { mukavasti /* mukava }
 I.NOM helped Jukka-PTV settle-INF/MA-ILL comfortably comfortable
 päivätorkui-lle aurinko-on]
 afternoon.naps-ALL sun-ILL
 ‘I helped Jukka to sleep comfortably in the sun’ [based on Koskinen 1998:325]

When the matrix clause has an ordinary nominative subject, the embedded object is marked with accusative (35a). Then, when the matrix subject is absent or bears lexical case, the embedded object becomes nominative (35b).²⁷

(35) a.  Hän kävi [TP avaa-ma-ssa **ove-n**] NOM-ACC
 s/he.NOM went open-INF/MA-INE door-ACC
 ‘S/he went to open the door’

²⁶ MA-infinitives can also occur with the essive case marker, but not when they function as clausal complements.

²⁷ In this section, of the constructions that remove the subject from the dependent-case calculus, I only show imperatives, but all of the data can be replicated for passives and necessives.

- b. Käy [TP avaa-ma-ssa **ovi**]! NOM
 go.IMP open-INF/MA-INE door.NOM
 ‘Go open the door!’

This resembles the same pattern from monoclausal sentences in section 4.1. In (35a), the embedded object is c-commanded by another DP unvalued for case, i.e. the matrix subject, and thus is assigned dependent case (= accusative). In (35b), there is no other DP unvalued for case that c-commands the embedded object and thus it surfaces with unmarked case (= nominative) at PF.

Accordingly, the pattern in (35) can be accounted for under DCT by considering (i) the CP to be the relevant domain for dependent case and (ii) MA-infinitives to be projections smaller than CP, so that the domain over which dependent case is calculated includes both the matrix and embedded clauses. Following Koskinen (1998), I assume that MA-infinitives are TPs.²⁸

(35b) also reveals that PRO is either absent from these constructions or inert for the purposes of dependent-case assignment. Otherwise, there would be no principled way to explain why the embedded object’s case is contingent on the presence of an argument in the matrix clause. Another DP like PRO inside the embedded clause that c-commands the object and is unvalued for case (for some portion of the derivation) would invariably license dependent case on the object, thereby negating any effect that the matrix clause could ever have. While either analysis (i.e. no PRO or inert PRO) would in principle account for the case pattern in (35b), I will adopt the first analysis that MA-infinitives lack a PRO.²⁹ This choice is largely for the sake of simplicity, but there are two arguments in its favor. First, if PRO can only occur in CPs, as Landau (2000) argues, this absence would follow from MA-infinitives being smaller than CPs. Second, this analysis also allows for a uniform treatment of PRO crosslinguistically as a dependent-case licenser, rather than parametrizing its ability to license dependent case on a language-by-language basis. On this analysis, then, PRO has no effect on dependent case in MA-infinitives because it is not there.

The crucial pattern emerges when the embedding predicate has its own object.³⁰ As shown in (36), when the matrix subject is present, the matrix subject is nominative, the matrix object is accusative, and the embedded object is accusative; this is the pattern expected, given what we have seen so far.

²⁸ There is reason to believe that *-mA* corresponds to a v^0 head: (i) it cannot cooccur with verbal inflection, such as passivization, and (ii) *-mA* is the morpheme used to form agentive participles, which is in line with the argument-structure role of v^0 . Thus, MA-infinitives are at least as big as v Ps. Though I follow Koskinen (1998) in assuming that they are TPs, the analysis in section 5.3 is compatible with MA-infinitives being v Ps as well; see fn. 39. I assume that the case morphology that appears on the verb is assigned directly to the nonfinite clause, with no intervening nominal projections (roughly in line with Vainikka 1995), though this choice is relatively inconsequential. The choice is based on the fact that MA-infinitives do not allow nominal modification (34) and cannot occur with possessive suffixes, the latter of which is a hallmark of nominals in Finnish and is possible with other nonfinite clause types.

²⁹ A reviewer raises the question of how the subject position gets saturated in MA-infinitives if they do not contain a PRO. Standardly, in analyses of the semantics of control, the subject is taken to be *unsaturated* and the embedded control clause to denote a property of individuals (e.g. Montague 1973; Bach 1979; Dowty 1985; Chierchia 1989). On Chierchia’s (1989) analysis, this is produced by a λ -binder abstracting over PRO (for a recent instantiation of this analysis, see Pearson 2013). The absence of PRO produces the same semantic object, namely a property of individuals.

³⁰ Some of these predicates include *pakottaa* ‘force’, *taivuttaa* ‘persuade’, *pyytää* ‘ask’, and *kieltää* ‘deny’. For a comprehensive list of predicates that embed MA-infinitives, see Vainikka (1989:330).

- (36) a. **Hän** pakotti **lapse-n** [TP avaa-ma-an **ove-n**] NOM-ACC-ACC
 s/he.NOM forced child-ACC open-INF/MA-ILL door-ACC
 ‘S/he forced the child to open the door’ [Nelson 1998:238]
- b. **Maija** pyysi **Jukka-n** [TP luke-ma-an **kirja-n**] NOM-ACC-ACC
 Maija.NOM asked Jukka-ACC read-INF/MA-ILL book-ACC
 ‘Maija asked Jukka to read the book’ [Vainikka 1989:267]

Under DCT, this pattern could in principle be modelled in one of two ways: (i) a covariance derivation, where the matrix subject licenses dependent case on both objects (37), or (ii) a daisy-chain derivation (38), where the matrix object licenses dependent case on the embedded object and then the matrix subject licenses dependent case on the matrix object.

- (37) *Covariance derivation* [Subj V⁰ Obj [TP V⁰-MA Obj]]
- (38) *Daisy-chain derivation* [Subj V⁰ Obj [TP V⁰-MA Obj]]

However, in the absence of a matrix subject, *both* the matrix object and the embedded object surface with nominative case, as shown in (39). This rules out the daisy-chain derivation for MA-infinitives in (38). Rather, the case of the matrix and embedded objects covaries with the presence of the matrix subject, as predicted by the analysis in (37).

- (39) *Both objects must be nominative when the matrix subject is absent*
- a. Pakota {**lapsi** /***lapse-n**} [TP avaa-ma-an {**ovi** /***ove-n**}]!
 force.IMP child.NOM child-ACC open-INF/MA-ILL door.NOM door-ACC
 ‘Force the child to open the door!’ [Nelson 1998:238]
- b. Pyydä {**Jukka** /***Jukka-n**} [TP luke-ma-an {**kirja** /***kirja-n**}]!
 ask.IMP Jukka.NOM Jukka-ACC read-INF/MA-ILL book.NOM book-ACC
 ‘Ask Jukka to read the book!’ [Vainikka 1989:268]

Binding reveals that the matrix object nevertheless c-commands the embedded object. Finnish third-person possessive suffixes are subject to Condition A, as illustrated in (40a). Crucially, a third-person possessive suffix on the embedded object can be bound by the matrix object (in addition to the matrix subject), as shown in (40b). This shows that the matrix object does indeed c-command the embedded object. All else equal, the matrix object *should* then license dependent case on the embedded object. The fact that it does not thus needs to be explained.³¹

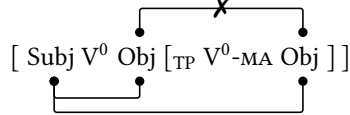
³¹ A reviewer raises the possibility of analyzing this asymmetry in terms of extraposition: the MA-infinitive extraposes before dependent-case assignment and then reconstructs for binding at LF. There are several arguments against such an analysis. First, the extraposition would be string-vacuous, so there is no independent evidence for extraposition. Second, unlike canonical extraposition, it would have to be obligatory. Third, it would require delaying dependent-

(40) *Matrix object c-commands the embedded object*

- a. **Poika**₁ myi marsu-**nsa**_{1/*2}
 boy.NOM sold guinea.pig.ACC-3.POSS
 ‘The boy₁ sold his_{1/*2} guinea pig’ [Nelson 1998:187]
- b. **Maija**₁ pyysi **Peka-n**₂ [TP tuo-ma-an levy-**nsä**_{1,2,*3}]
 Maija.NOM asked Pekka-ACC bring-INF/MA-ILL record.ACC-3.POSS
 ‘Maija₁ asked Pekka₂ to bring her/his_{1,2,*3} record’ [Vainikka 1989:270]

What (39) and (40) reveal is that a matrix subject, but not a matrix object can license dependent case across an embedded TP boundary into a MA-infinitive, as schematized in (41).

(41) *Case assignment in MA-infinitives*



Structurally case-marked adjuncts in the matrix clause are also unable to license dependent case across an embedded TP, and thus they pattern with matrix objects. This is shown in (42a), where the multiplicative adjunct has matrix scope and still both objects must be nominative. (42a) additionally shows that the matrix object has the *ability* to license dependent case, as it does so on the adjunct, making its inability to do so on the embedded object all the more striking. When the adjunct does have embedded scope, the embedded object licenses dependent case on the adjunct in an ordinary local configuration (42b).³²

(42) *Adjuncts do not affect the case of the objects*

- a. Pyydä **Jukka** [kolmanne-n kerra-n] [TP luke-ma-an **kirja**]
 ask.IMP Jukka.NOM third-ACC time-ACC read-INF/MA-ILL book.NOM
 ‘Ask Jukka for the third time to read the book!’ [Maling 1993:69]
- b. Pyydä **Jukka** [TP luke-ma-an **kirja** [kolmanne-n kerra-n]]
 ask.IMP Jukka.NOM read-INF/MA-ILL book.NOM third-ACC time-ACC
 ‘Ask Jukka to read the book for the third time!’ [Maling 1993:66]

case assignment; in section 3.2, I argued that dependent-case relationships are established as soon as the licensor is merged into the structure (see also section 7.1). Fourth, MA-infinitives are transparent for extraction (Toivonen 1995; Huhmarniemi 2012). This suggests that even if MA-infinitives were to string-vacuously extrapose, they would independently need to be accessible to syntactic operations prior to extraposition, and it would be unclear why this would preclude case assignment.

³² The adjunct in (42a) also has an embedded reading, which is presumably derived from (42b) via movement; see fn. 24.

The overarching pattern to emerge from Finnish MA-infinitives is summarized in (43).^{33,34}

(43) **FINNISH CASE GENERALIZATION**

In Finnish, a matrix subject can license dependent case across an embedded TP boundary, but a matrix object and a matrix adjunct cannot.

One might wonder why the Finnish pattern in (43) is not found in languages like English. There are two reasons. First, in languages like English, control infinitives are CPs (Landau 2000), and CPs are domains for dependent-case assignment. Second, as control infinitives in languages like English contain PRO, PRO will always locally license dependent case on the object. Thus, in languages like English, case assignment in control infinitives is always determined locally; it is never contingent on elements in the matrix clause.³⁵ Finnish MA-infinitives (and TA-infinitives), on the other hand, are smaller than CP and contain no PRO (or, alternatively, they contain a PRO inert for dependent case), which causes the embedded DPs to interact for dependent-case assignment with matrix DPs. (35) is the crucial datapoint showing this property. The prediction then is that in languages with a pattern like (35), the same generalization from Finnish in (43) should emerge. I leave exploring this prediction to future research.

Crucially, the Finnish Case Generalization in (43) does not involve movement, which will prove important in the next two sections. Like the Movement–Case Generalization from section 3, it also does not fall under the purview of standard notions of locality, e.g. phases, where a domain is either opaque to all operations or transparent to all operations. Under these standard, binary notions of locality, it is unexpected for a domain (here, a TP) to be penetrable by a DP in one position (matrix-subject position), but not another position (matrix-object position, which is arguably more local than the matrix subject). As such, the Finnish Case Generalization must be the result of some other kind of locality, namely one that is *nonbinary*. In the next section, I will argue that this nonbinary notion of locality is the Williams Cycle.

5 Improper case

In this section, I propose that dependent-case assignment is constrained by the *Ban on Improper Case* in (44). This constraint rules out dependent-case assignment configurations like (45).

³³ Something like the Finnish Case Generalization in (43) would presumably need to hold under an FHCT analysis as well, because whatever conditions assigning accusative case to the embedded object can only be triggered by a matrix subject. This is notwithstanding the problem that structurally case-marked adjuncts pose for an FHCT analysis in the first place; see section 4.1.

³⁴ (43) also captures the other canonical nonfinite clause type in Finnish, namely TA-infinitives (traditionally called the “first infinitive”). TA-infinitives behave identically to MA-infinitives for the purposes of case assignment. However, the predicates that embed TA-infinitives never have their own objects. Thus, while TA-infinitives exhibit the same basic pattern as (35), the more complex pattern involving matrix objects in (36) and (39) happens not to arise for them.

³⁵ For the same reason, we do not expect to find the Finnish pattern in ECM constructions, assuming that ECM involves movement of the embedded subject (though see section 7.3). In ECM constructions, the embedded subject will locally license dependent case on the embedded object, so that the embedded object’s case is never contingent on elements in the matrix clause.

(44) **BAN ON IMPROPER CASE**

A DP in [Spec, XP] cannot establish a dependent-case relationship with a lower DP across YP, where Y is higher than X in the functional sequence.

$$(45) \quad [_{YP} Y^0 [_{XP} DP_1 X^0 \dots [_{YP} \dots DP_2 \dots]]] \quad \text{where } Y > X$$

The Ban on Improper Case is a constraint in the spirit of the Williams Cycle (WC) (Williams 1974, 2003, 2013; van Riemsdijk and Williams 1981), which in its original form is only a constraint on movement dependencies. In section 6, I will propose that the WC be generalized to encompass case, movement, *and* agreement and then take up how to derive this generalized WC.

I begin in subsection 5.1 by introducing the WC in its instantiation for movement, known as the Generalized Ban on Improper Movement. Subsection 5.2 proposes the Ban on Improper Case, an extension of the WC particularized to case. In subsections 5.3 and 5.4, I then apply the proposal to the Finnish Case Generalization and the Movement–Case Generalization respectively. Subsection 5.5 briefly discusses the treatment of case and movement in Baker (2015).

5.1 The Williams Cycle

The Williams Cycle (WC) is a size-based locality constraint on (movement) dependencies spanning two clauses, going back to Williams (1974) and van Riemsdijk and Williams (1981). The basic idea behind the WC is that movement from a specific domain in an embedded clause may move to the *same kind* of domain or a *higher* domain in the matrix clause. In Williams (2003), the WC is formulated as the *Generalized Ban on Improper Movement* (GBOIM) in (46), where domains are defined in terms of the functional sequence (*fseq*).³⁶ I will notate X being higher in *fseq* than Y as $X > Y$, and, for concreteness, I will assume the simple functional sequence in (48).

(46) **GENERALIZED BAN ON IMPROPER MOVEMENT (GBOIM)**

Movement to [Spec, XP] cannot proceed from [Spec, YP] or across YP, where Y is higher than X in the functional sequence. [based on Williams 2003]

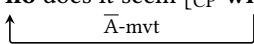
(47) A dependency relating α and β occurs **ACROSS** XP iff XP dominates β but not α .

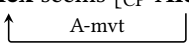
(48) $fseq = \langle C > T > v > V \rangle$

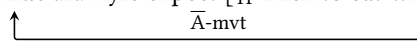
As its name suggests, the GBOIM is intended to subsume the traditional ban on improper movement (Chomsky 1973, 1981; May 1979). Thus, to illustrate the GBOIM, let us consider how it handles the classical instance of improper movement, namely the ungrammaticality

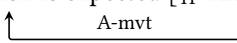
³⁶ The formulation of the GBOIM given in Williams (2003:72) does not represent the full generality of what Williams's analysis of the GBOIM actually derives (see section 6). All else being equal, that formulation allows movement *across* projections higher in the functional sequence than the launching site of movement because it is stated only in terms of the landing site. I have reformulated the GBOIM in (46) to avoid this problem.

of hyperraising: A-movement out of a finite clause. While \bar{A} -movement may leave a finite clause (49a), A-movement may not (49b). This contrast does not extend to nonfinite clauses, which allow both \bar{A} -movement (50a) and A-movement (50b) out of them alike.^{37,38}

(49) a. **Who** does it seem [_{CP} **who** ate the nattoo]?


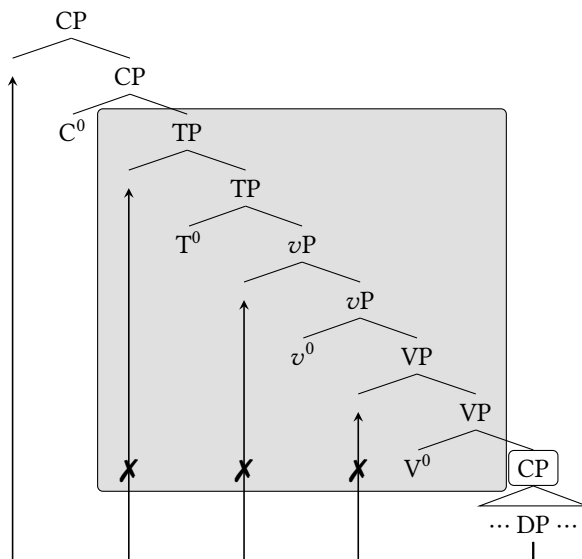
b. ***Alex** seems [_{CP} **Alex** ate the nattoo].


(50) a. **What** did Kyle expect [_{TP} Alex to eat **what**]?


b. **Alex** is expected [_{TP} **Alex** to eat the nattoo].


According to the GBOIM, the relative heights of the launching and landing sites determine whether extraction is possible. Because finite clauses are CPs, movement out of a finite clause can land no lower than [Spec, CP] in the next highest clause, as schematized in (51).

(51) *Movement from CP cannot land lower than CP*

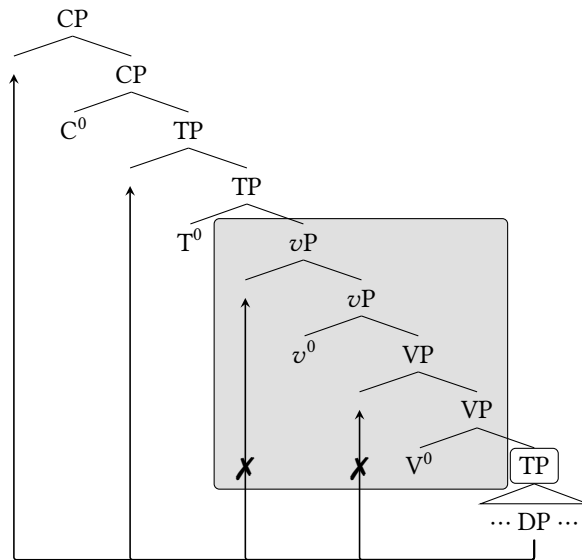


³⁷ For the sake of simplicity, I equate “finite” with CP and “nonfinite” with TP. English, of course, also has nonfinite CPs, which prohibit A-movement out of them, like finite clauses. Thus, the improper-movement pattern is genuinely about clause size, as the GBOIM encapsulates.

³⁸ In (49), I do not depict movement through [Spec, CP], but this would not change the movement derivations that are ruled out by the GBOIM. Note though that the traditional ban on improper movement (Chomsky 1973, 1981; May 1979) does require intermediate movement through [Spec, CP] to block hyperraising.

As depicted in (51), CP is a barrier for movement to [Spec, TP] because $C > T$ in *fseq*, but CP is not a barrier for movement to [Spec, CP] because $C \not> C$. Thus, \bar{A} -movement, but not A-movement out of a finite clause is grammatical. On the other hand, because nonfinite clauses are TPs, movement out of a nonfinite clause may land in either [Spec, TP] or [Spec, CP] because $T \not> T$ and $T \not> C$ respectively. Thus, both A-movement and \bar{A} -movement are possible out of a nonfinite clause, unlike finite clauses, as schematized in (52).

(52) *Movement from TP cannot land lower than TP*



Under the GBOIM, size matters. A smaller clause is permeable to more movement types than a larger clause, because the maximal projection of a smaller clause will be lower on *fseq* than the maximal projection of a larger clause. Constraining movement in terms of clause size extends beyond the distinction between A-movement and \bar{A} -movement. Here are several examples (taken from [Keine 2016](#)): (i) Infinitival clauses are opaque to extraposition, but not regular A-movement and \bar{A} -movement ([Ross 1967](#); [Baltin 1978](#)). (ii) Embedded questions are opaque to *wh*-movement, but not topicalization and relativization ([Williams 2013](#)). (iii) In Hindi-Urdu, finite clauses are opaque to A-scrambling, but not \bar{A} -scrambling ([Mahajan 1990](#)). In German, (iv) embedded V2 clauses are opaque for movement into a verb-final clause, but not movement into a V2 clause ([Haider 1984](#)); (v) finite clauses are opaque to scrambling and relativization, but not *wh*-movement or topicalization ([Bierwisch 1963](#); [Ross 1967](#); [Bayer and Salzmann 2013](#); [Müller 2014b](#)); and (vi) incoherent infinitives are opaque to scrambling, but not *wh*-movement and relativization ([Bech 1955/1957](#); [Wurmbrand 2001](#)). What these asymmetries share is involving a domain that is permeable to one movement type, but not another movement type (what [Keine](#) terms SELECTIVE OPACITY). The GBOIM derives these asymmetries as “generalized” improper movement configurations, i.e. in terms of clause size. For more discussion, see [Williams \(1974](#),

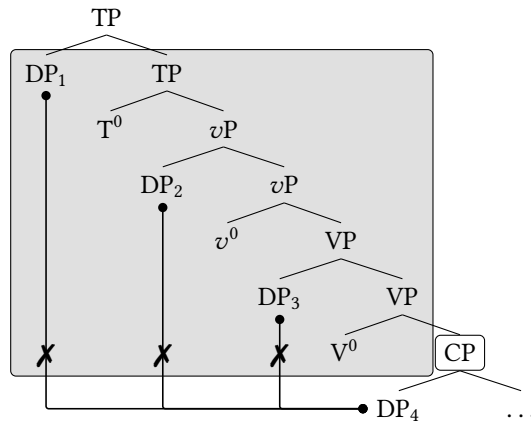
2003, 2013), Müller and Sternefeld (1993, 1996), Abels (2007, 2009, 2012a,b), Neeleman and van de Koot (2010), Müller (2014a,b), and Keine (2016, 2019, 2020).

5.2 Proposal

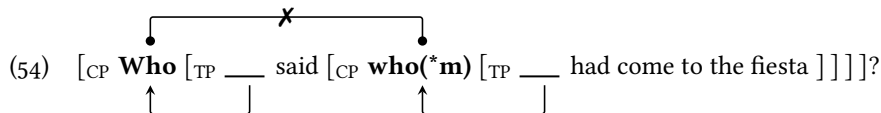
There are crucially parallels between the locality problems from sections 3 and 4 and the kinds of movement configurations ruled out by the Generalized Ban on Improper Movement. To see these parallels, let us consider the two locality problems in turn.

With respect to the Movement–Case Generalization, recall the Lower-DP Problem, according to which a DP cannot be the lower DP in a dependent-case pair when in an intermediate landing site. (I will return to the Higher-DP Problem in section 5.4.) This characterization can be recast in terms of the WC, viz. clause size and the functional sequence: a DP_α in [Spec, CP] cannot enter into a dependent-case relationship with a DP_β in a higher clause— DP_α being the lower in the pair—if DP_β is in [Spec, TP], [Spec, vP], or [Spec, VP], because $C > T$, $C > v$, and $C > V$ in *fseq*. This is schematized in (53).

(53) Lower-DP Problem



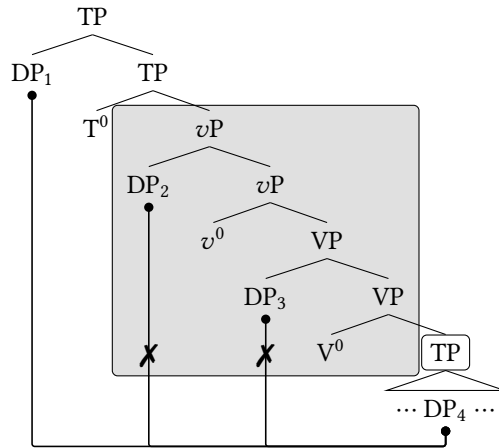
Note that a dependent-case relationship between two [Spec, CP] positions also needs to be ruled out (54). This configuration does not fall under the characterization of the Lower-DP Problem—or from the Ban on Improper Case, to be proposed below—because $C \not> C$ in *fseq*.



However, for the higher DP in (54) to be in [Spec, CP], it will have undergone \bar{A} -movement to that position. Thus, the impossibility of this particular configuration falls under the Higher-DP Problem (i.e. that an \bar{A} -moved element cannot be the higher DP in a dependent-case pair) and will follow from the analysis of the Higher-DP Problem in section 5.4.

The same parallels apply to the Finnish Case Generalization, according to which a matrix subject can license dependent case across an embedded TP clause boundary, but a matrix object and a matrix adjunct cannot. In terms of the WC: a DP in [Spec, TP] can license dependent case on another DP across a TP, because $T \not> T$, but a DP in [Spec, *v*P] or [Spec, VP] cannot do so, because $T > v$ and $T > V$. This is schematized in (55).

(55) *Finnish-Case Generalization*

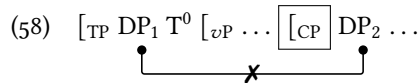
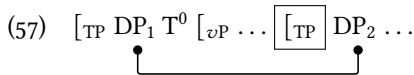


These parallels in (53) and (55) are the motivation for extending the WC to dependent-case assignment. I propose that dependent-case assignment is subject to the *Ban on Improper Case* in (56), a direct extension of the WC to case.

(56) **BAN ON IMPROPER CASE**

A DP in [Spec, XP] cannot establish a dependent-case relationship with a lower DP across YP, where Y is higher than X in the functional sequence.

The Ban on Improper Case states barrierhood for dependent-case assignment relative to the *fseq*-position of the higher DP in the dependent-case pair. For example, a DP in [Spec, TP] can license a dependent-case relationship with another DP past TP, *v*P, and VP, because none of these projections are higher than T in *fseq* (57). However, a DP in [Spec, TP] cannot license dependent case past CP, because $C > T$ so that CP is a barrier to dependent-case licensing from TP (58). Note that CP's barrierhood extends to all projections lower than T in *fseq* as well.



Notice that the Ban on Improper Case makes no reference to movement or clause types. It is more general than the empirical data that motivated it. The remainder of this section shows how the Ban on Improper Case applies to our two very different generalizations: the Finnish Case Generalization in subsection 5.3 and the Movement-Case Generalization in subsection 5.4.

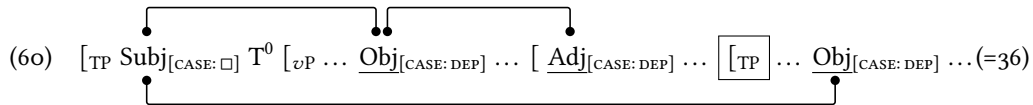
5.3 Application to Finnish

The Finnish Case Generalization is repeated below in (59).

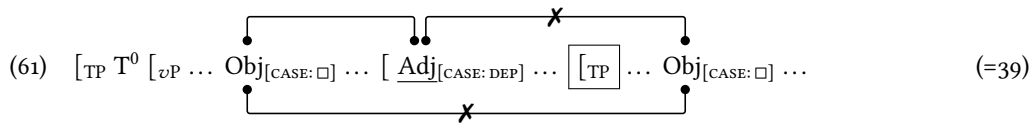
(59) **FINNISH CASE GENERALIZATION**

In Finnish, a matrix subject can license dependent case across an embedded TP boundary, but a matrix object and a matrix adjunct cannot.

Under the Ban on Improper Case, the matrix subject is able to license dependent case across the embedded TP boundary because the matrix subject is located in [Spec, TP] and $T \not> T$ in *fseq*. Thus, it licenses dependent case on the matrix object (within the same clause) and on the embedded object (across the clause boundary). This is schematized in (60).³⁹



The matrix object occupies a *vP*-internal position—the precise position is inconsequential, but somewhere below *v*. From its *vP*-internal position, the matrix object is unable to license dependent case across the embedded TP boundary because $T > v$ in *fseq*, thereby making TP a barrier for dependent-case licensing from *vP*-internal DPs, in particular from DPs in [Spec, *vP*] and any position lower in *fseq*. The same barrierhood applies for matrix adjuncts as well, which are generated in *vP*-internal positions too. As such, in the absence of a matrix subject, the [CASE: □] features on the matrix and embedded objects both remain unvalued throughout the derivation and are realized as unmarked case at PF. This is schematized in (61).



Under this analysis, there is nothing special about case in MA-infinitives. The same general case mechanism, namely dependent case, applies everywhere in the language as syntactic structure is built up, following section 2—but this mechanism is constrained by the Ban on Improper Case.

Previous analyses of MA-infinitives are all broadly based on the idea that when the matrix subject is absent or bears lexical case, i.e. the environments in Finnish with nominative objects, the ability to assign accusative case is gone altogether (Vainikka 1989; Nelson 1998; Vainikka

³⁹ An assumption of this analysis is that the subject undergoes A-movement to [Spec, TP], from where it is then able to penetrate the embedded TP to license dependent case. However, if we were to analyze MA-infinitives as being *vP*s, rather than TP's (contra Koskinen 1998), then the matrix subject would be able to penetrate the embedded *vP* from its base-generated position in [Spec, *vP*]. Note that the matrix object and adjuncts would not be able to penetrate an embedded *vP* because they are in positions below *vP*.

and Brattico 2014).⁴⁰ However, we saw in section 4.2 that structurally case-marked adjuncts are still accusative in configurations like (61); the relevant datapoint is repeated in (62).

- (62) Pyydä Jukka [kolmanne-n kerra-n] [TP luke-ma-an kirja] (=42a)
 ask.IMP Jukka.NOM third-ACC time-ACC read-INF/MA-ILL book.NOM
 ‘Ask Jukka for the third time to read the book!’ [Maling 1993:69]

If the ability to assign accusative case is absent in configurations like (61), as previous analyses assume, then there would be no source of accusative case for the adjunct in (62). However, (62) follows without further ado on the DCT analysis developed in this paper: the matrix object licenses dependent case on the adjunct, but the matrix and embedded objects cannot enter into a dependent-case relationship without violating the Ban on Improper Case.

5.4 Application to movement

The Movement–Case Generalization is repeated below in (63).

- (63) **MOVEMENT–CASE GENERALIZATION**
 A-movement can feed dependent case, but \bar{A} -movement cannot.

Let us begin with \bar{A} -movement. Recall that the locality problem with \bar{A} -movement is that an \bar{A} -moved element cannot enter into dependent-case relationships from its intermediate and final landing sites. Thus, we must consider (i) when an \bar{A} -moved element is the *lower* DP in a potential dependent-case pair (see (21)) and (ii) when it is the *higher* one (see (22)). These are the Lower-DP Problem and the Higher-DP Problem respectively. For the sake of clarity, I will label the higher and lower DPs in a dependent-case pair as DP_α and DP_β respectively, unless it is an \bar{A} -moved element, for which I will reserve the label DP_μ . It should be emphasized that this labelling is for expository purposes only, and the Ban on Improper Case does not (need to) take into account whether the relevant DPs have undergone movement.

According to the Ban on Improper Case, a DP_α in [Spec, TP], [Spec, ν P], or [Spec, VP] cannot enter into a dependent-case relationship with a DP_μ in [Spec, CP] because these projections are all lower than C in *fseq*. That is, CP is a barrier for dependent-case licensing from TP and all projections lower in *fseq* (64). This barrierhood accounts for why an \bar{A} -moved element may not have its case altered at its intermediate and final landing sites, i.e. the Lower-DP Problem.

- (64) [TP DP_α ... [ν P DP_α ... [CP DP_μ ...]] C > T, C > ν
-

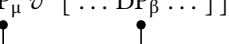
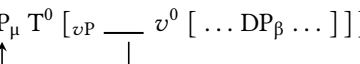
⁴⁰ Maling’s (1993) case-in-tiers analysis is an exception, but only because it simply stipulates that unless one of the objects bears a lexical case (or partitive case, whose status is contested), the matrix and embedded objects must bear the same case, a fact that my analysis derives.

The Ban on Improper Case, however, does not prohibit a DP_μ in [Spec, CP] from establishing a dependent-case relationship with a DP_β lower in the same clause, i.e. the Higher-DP Problem, as C is higher than these projections in *fseq*. I propose that the reason why \bar{A} -moved DPs cannot themselves license dependent case is because they are encased in a QP, i.e. Q-particle Phrase (in the sense of Cable 2007, 2010).⁴¹ Because only DPs may establish a dependent-case relationship, a DP inside a QP cannot be the *higher* DP in a dependent-case pair because it does not c-command out of the QP and hence never c-commands other DPs in the clause (65a). On the other hand, a DP inside a QP can be the *lower* DP in the pair because other DPs can still c-command into the QP (65b).

- (65) a. $[_{QP} Q^0 DP_\mu] [\dots DP_\beta \dots]$

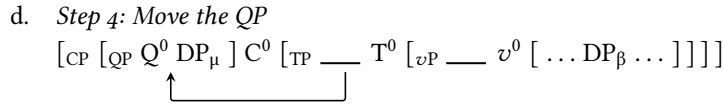
- b. $DP_\alpha [\dots [_{QP} Q^0 DP_\mu] \dots]$


However, a DP that undergoes \bar{A} -movement should still be able to enter into dependent-case relationships from the A-positions that it occupied prior to \bar{A} -movement, which (65a) does not permit. To solve this problem, I adopt Safir's (2019) independently-motivated proposal that the QP is countercyclically merged onto the DP immediately before it \bar{A} -moves (see also Rezac 2003; Stanton 2016).⁴² To illustrate how this applies to dependent case, consider the derivation of a simple *wh*-subject question in (66): (i) the subject is base-merged in [Spec, *v*P], from where it licenses dependent case on the object (66a); (ii) the subject A-moves to [Spec, TP] (66b); (iii) the QP is then merged on top of the subject (66c); and finally (iv) the QP moves to [Spec, CP] (66d). The \bar{A} -moving DP will always be encased in the QP before it reaches any intermediate or final [Spec, CP] landing sites, thereby preventing it from licensing dependent case on other DPs from those derived positions, i.e. the Higher-DP Problem.

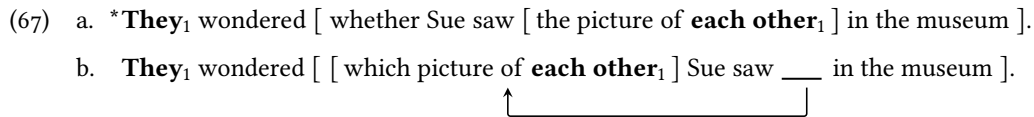
- (66) a. *Step 1: Merge in the subject, assign dependent case*
 $[_{vP} DP_\mu v^0 [\dots DP_\beta \dots]]$

- b. *Step 2: A-move the subject*
 $[_{TP} DP_\mu T^0 [_{vP} \text{---} v^0 [\dots DP_\beta \dots]]]$

- c. *Step 3: Build the QP on the DP*
 $[_{TP} [_{QP} Q^0 DP_\mu] T^0 [_{vP} \text{---} v^0 [\dots DP_\beta \dots]]]$

⁴¹ While Cable's (2007, 2010) QP-system is designed primarily for *wh*-movement, I follow Cable in assuming that a QP-analysis extends to other kinds of \bar{A} -movement as well; see Cable (2007:369–375).

⁴² A few notes are in order: First, Safir (2019) does not assume that the shell insulating an \bar{A} -moved element is always a QP, as I do here, although nothing critical in this paper rests on that assumption. Second, the reader is referred to Safir (2019) for discussion of how other facets of the A/ \bar{A} -distinction can be captured under a QP-shell analysis. Third, non-countercyclic implementations of the QP-shell analysis are conceivable. What I have in mind is the multidominance analysis of QP-movement in Johnson (2012) and Poole (2017), where the DP merges with its base position *and* with the Q-particle, the resulting QP then being merged in the landing site of movement; this is effectively sideways movement of the DP into the QP.

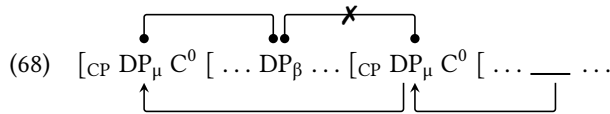


Note that the addition of the QP layer does not handle the Lower-DP Problem because other DPs can nonetheless c-command a DP encased in a QP; this problem still requires the Ban on Improper Case, as was schematized above in (64). This point, however, raises an alternative analysis where QPs are themselves opaque to case assignment, so that once they are formed on a DP, that DP no longer interacts with case assignment. Such an analysis faces the dilemma that the opacity for case assignment would have to come from its own source and not apply to other dependencies, because c-command into a QP for other dependencies, e.g. binding, is indeed possible. For example, consider (67), in which an anaphor in a moved *wh*-element—on the analysis here, a QP—can be bound from its landing site (Barss 1986; Lebeaux 1988).



Under this analysis, the behavior of \bar{A} -movement with respect to dependent case follows from two components: a QP-shell and the Ban on Improper Case. The former handles the Higher-DP Problem, and the latter handles the Lower-DP Problem.

Because the QP-shell and the Ban on Improper Case are independent from each other, this account predicts that if the ‘higher/lower’ symmetry breaks down, it should crucially do so in one direction. Namely, if an \bar{A} -moving DP is *not* encased in a QP-shell, then it should be able to be the higher DP in a dependent-case pair in its intermediate and final landing sites, but not the lower DP. In other words, it should exhibit the behavior of the Lower-DP Problem, but not the Higher-DP Problem. Empirically, this would be a movement type that is just like *wh*-movement—targeting a position high in *fseq*, like [Spec, CP]—, except it does not involve a QP-shell.⁴³ This prediction is schematized in (68).



This prediction appears to be borne out in Koryak, as described by Abramovitz (2020). Abramovitz shows that in Koryak, long *wh*-movement of an embedded nominative DP results in dependent ergative case on the matrix subject, as shown in (69).⁴⁴ (For arguments that ergative in Koryak is dependent case, see Abramovitz 2020:6–12.)

⁴³ Under Cable’s (2007, 2010) QP-system, such a movement type should also lack pied-piping. I do not know whether Koryak *wh*-movement—which I introduce shortly below as an instance of such a movement type—allows pied-piping or not, but none of the examples in Abramovitz (2020) involve pied-piping.

⁴⁴ For readability, I have not glossed the verbal morphology in the Koryak data.

(69) *Koryak long wh-movement*

- a. **ʏəmmo** təvalomək [CP əno ʔewŋəto-na-k jətɕimawnin **koŋŋ-o**]
 1SG.NOM hear that Hewngyto-OBL.SG-ERG break cup-NOM.PL
 ‘I heard that Hewngyto broke cups’
- b. **jeŋ-u₁** {**ʏə-nan** / *ʏətɕɕɕi } valomnaw [CP əno ʔewŋəto-na-k
 what-NOM.PL 2SG-ERG 2SG.NOM hear that Hewngyto-OBL.SG-ERG
 jətɕimawnin ___₁]?
 break
 ‘What all did you hear that Hewngyto broke?’ [Abramovitz 2020:17]

(69) does not reveal *where* the *wh*-element enters into the dependent-case relationship with the matrix subject: matrix [Spec, CP] (the higher position) or embedded [Spec, CP] (the lower position). For (68) to be true, it must be the higher position. Crucially, embedded questions show that it is indeed the higher position.⁴⁵ In embedded questions, moving a nominative *wh*-element to embedded [Spec, CP] does *not* trigger dependent ergative on the matrix subject, as shown in (70). This is precisely what (68) predicts. By contrast, if (69) involved the *wh*-element establishing the dependent-case relationship from the lower position, then we would expect the matrix subject to be ergative in (70) as well, contrary to fact.

(70) *Koryak embedded question*

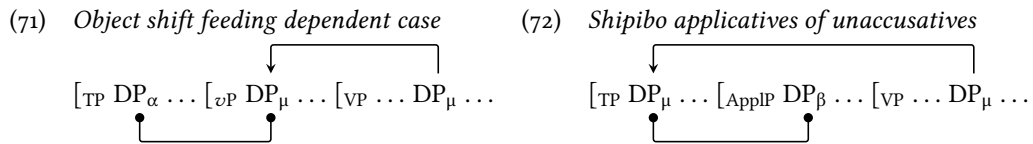
- {*ʏəm-nan / **ʏəmmo**} təkutɕetkejuŋəŋ [CP **jeq-qevi-jətɕ?-u** məjəlnew
 1SG-ERG 1SG.NOM think which-gift-contents-NOM.PL give
 ʔewŋəto-na-ŋ ənək-eto-ʔəlwəje-ŋ]
 Hewngyto-OBL.SG-DAT 3SG.POSS-birth-day-DAT
 ‘I am wondering what gifts I should give Hewngyto for his birthday’ [Abramovitz 2020:28]

Therefore, Koryak seems to confirm the prediction in (68) of the two-component analysis being proposed here. Whether this pattern can be found more widely, I leave for future research. Note that this analysis of Koryak requires the unorthodox assumption that ergative is dependent case assigned to the lower DP in a dependent-case pair, rather than the higher DP. Exploring this point is outside the scope of this paper, but see Yuan (2018, 2020) for independent arguments that “ergative” (i.e. dependent case on transitive subjects) can be assigned downwards.⁴⁶

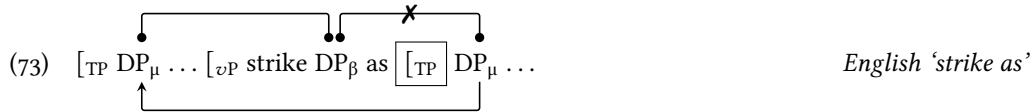
⁴⁵ Abramovitz (2020) proposes that the dependent-case relationship with the matrix subject is established when the moving *wh*-element is in matrix [Spec, vP]. He places the matrix subject in [Spec, TP], and so the directionality of dependent case appears to be upwards, as is standard for “ergative”. However, once we recognize that the subject’s base position is the inner specifier of vP, then on Abramovitz’s analysis, the directionality of dependent case would in fact need to be downwards, as in the analysis presented in the main text, because the *wh*-element would move through the outer specifier position of vP, which is above the matrix subject. Thus, as far as I can tell, my (re)analysis of the Koryak data is not substantively different from Abramovitz’s analysis, though technically a different claim.

⁴⁶ Yuan (2018, 2020) argues that in Inuit, the object must raise over the subject in order to license dependent ergative case on the subject. Abramovitz (2020:27–30) in fact provides data that can be taken as evidence in favor of extending this kind of analysis to Koryak. He shows that nominative objects raise to the edge of vP, which could plausibly be

Turning now to A-movement, recall the three examples of movement feeding dependent case from section 3.1: (some instances of) object shift, Shipibo applicatives of unaccusatives, and the English raising predicate *strike as*. All three of these examples obey the Ban on Improper Case. For object shift, the subject’s position is higher in *fseq* than the raised object’s position—presumably [Spec, TP] and [Spec, *v*P] respectively—, as schematized in (71). Similarly for Shipibo applicatives, the raised theme’s position is higher in *fseq* than the applicative’s position (72). Generally, if a DP moves clause-internally (i.e. within the same extended projection), it will be able to establish dependent-case relationships with other DPs in that same clause, because the higher DP in the pair will always be higher in *fseq* than the lower DP.



For English *strike as*, the movement crosses a clause boundary (the movement itself obeying the GBOIM; see section 5.1), but the position in which the moved DP lands is higher in *fseq* than the position of the internal argument of *strike*. Therefore, according to the Ban on Improper Case, the two DPs can establish a dependent-case relationship, as schematized in (73).



The Ban on Improper Case also explains why the internal argument of *strike* does not license dependent case on the subject in its embedded position before it moves: the *v*P-internal position of the internal argument is lower in *fseq* than T. Therefore, TP is a barrier to it licensing a dependent-case relationship with the embedded subject.

The discussion thus far has focused on dependent case, but it should be noted that this analysis does not preclude *lexical case* from being assigned to a moved position. First, the Ban on Improper Case is not formulated to encompass lexical-case assignment. But, because lexical case is assigned in a siblinghood relation, it is out of the purview of the Ban on Improper Case regardless. Assuming Bare Phrase Structure, where what projects is the head itself (Chomsky 1995a), it is then possible for a DP to move to a specifier position and be assigned a lexical case under siblinghood, in what would traditionally be a specifier–head relation (à la Rezac 2003) (see also fn. 5). To illustrate, consider dative–accusative constructions in Faroese (74), which are historically related to the more familiar Icelandic dative–nominative constructions.⁴⁷

analyzed as a position above the subject’s base position, so that dependent ergative is assigned downwards. Strictly speaking, such an analysis would require relaxing Earliness; see section 7.1 for discussion.
⁴⁷ Thanks to Ellen Woolford for bringing the Faroese data to my attention.

(74) *Faroese dative–accusative constructions*

Mær líkar **hana** væl
 I.DAT likes her.ACC well
 ‘I like her a lot’

[Thráinsson et al. 2004:255]

These constructions can be analyzed as the following: (i) the subject is base-merged in [Spec, *v*P], from where it licenses dependent case (= accusative) on the object (75a); (ii) the subject moves to a higher projection in the clause, e.g. *Exp*⁰ (75b); and (iii) the head of this projection assigns the subject lexical dative case (75c). The difference between Faroese and Icelandic is that in Icelandic, the subject is assigned dative case in its base-generated position, thus bleeding dependent-case assignment and yielding nominative objects.⁴⁸

(75) a. *Step 1: Merge in the subject, assign dependent case*

$$[{}_{vP} \text{Subj}_{[CASE: \square]} v^0 [\dots \text{Obj}_{[CASE: DEP]} \dots]]$$

b. *Step 2: Move the subject*

$$[{}_{ExpP} \text{Subj}_{[CASE: \square]} \text{Exp}^0 [\dots [{}_{vP} \text{---} v^0 [\dots \text{Obj}_{[CASE: DEP]} \dots]]]]$$

c. *Step 3: Assign the subject lexical dative case*

$$[{}_{ExpP} \text{Subj}_{[CASE: DAT]} \text{Exp}^0 [\dots [{}_{vP} \text{---} v^0 [\dots \text{Obj}_{[CASE: DEP]} \dots]]]]$$

There are several other examples that, to my knowledge, might instantiate this kind of derivation with movement to a lexical-case position: ergative subjects in what Woolford (2015) terms *Active Ergative* languages (where ergative case is associated with external arguments),⁴⁹ the “marked nominative” construction in Dinka (van Urk 2015), and differential object marking in Hindi-Urdu (Bhatt and Anagnostopoulou 1996). There are likely many other such instances, but these exemplify when such a derivation might be reasonably invoked.

In sum, the Ban on Improper Case accounts for the interactions between movement and dependent case: roughly, A-movement, but not \bar{A} -movement may feed dependent-case assignment. Importantly, the analysis does not invoke separate operational primitives for A-movement and \bar{A} -movement. Rather, the analysis derives from the positions targeted by different movement types. Moreover, if Safir (2019) is correct that the QP-shells in \bar{A} -movement can be derived from independent factors, then the analysis presented here captures the A/ \bar{A} -distinction in this (narrow) domain purely as an epiphenomenon. This thinking is in line with minimalist syntax, where all structure building is the result of the operation MERGE. The foundations of the analysis

⁴⁸ It should be noted that the derivation in (75) goes against Marantz’s (1991) original dependent-case algorithm in (7), but it is admissible under the present analysis, based on Preminger’s (2011, 2014) syntactic implementation.

⁴⁹ As such, in Active Ergative languages, ergative case is not dependent case, but rather lexical case—or more accurately, inherent case (e.g. Woolford 1997, 2006; Legate 2002, 2008). This is a nice example of why the descriptive label given to a case should not be taken to entail a particular case-assignment mechanism, as discussed in section 2.

were also independently motivated from the Finnish Case Problem, which crucially does not involve movement.

5.5 Remarks on Baker (2015)

As the most comprehensive dependent-case system to date, it is instructive to consider how Baker's (2015) system fares on the data considered in this paper. First, Baker does not investigate anything comparable to Finnish MA-infinitives, and hence nothing in his system handles the Finnish Case Generalization. Second, his treatment of the A-movement examples from section 3.1 is—more or less—in line with what I propose for them in section 5.4. Therefore, let us set aside these two issues and focus on the Higher-DP Problem and the Lower-DP Problem, where a comparison is more fruitful.

Regarding the interaction of case and movement, Baker proposes that (i) dependent case is assigned at phasal Spellout and that (ii) it is calculated within the phase complement, crucially excluding the phase edge.⁵⁰ Consider the *wh*-question in (76) at the point when the CP phase is spelled out. As the phase complement of C, TP is the domain of dependent case. The higher copy of *who* in [Spec, CP] is not within the phase complement and hence does not factor into the dependent-case calculus. Thus, the only dependent-case relationship established at the CP phase is between *she* and the lower copy of *who*. The higher copy of *who* in [Spec, CP] is spelled out in the next phase (or by whatever procedure spells out the edge of the highest phase).

$$(76) \quad [_{CP} \text{ who } C^0 \underbrace{[_{TP} \text{ she see who }]}_{\text{phase complement}}]$$

This analysis accounts for the Higher-DP Problem: in its intermediate and final landing sites, where it could be the higher DP of a dependent-case pair, an \bar{A} -moved element is not included in the calculus of dependent case for that phase.

However, there are two drawbacks. First, this treatment of the Higher-DP Problem does not extend to the Koryak data, where an \bar{A} -moved element in [Spec, CP] does in fact establish a dependent-case relationship with a DP lower in that clause. Baker's analysis categorically rules out such configurations, which appears to be too strong. Second, it leaves the Lower-DP Problem unresolved. As laid out in section 3.2, from embedded [Spec, CP] positions, a DP unvalued for case should be eligible to be the lower DP in a dependent-case pair, but crucially it is not. There is nothing in Baker's analysis to rule out such pairs. In general, appealing to the locality afforded by phases cannot account for the Lower-DP Problem (see (23)).

Therefore, Baker's (2015) analysis of the interaction between case and movement does not extend to the full range of facts considered in this paper. However, this should not be construed

⁵⁰ For the sake of simplicity, I am leaving out the role of 'soft' phases and case inheritance in chains in Baker's (2015) system, as they do not directly affect the points being made here.

as an argument against Baker’s overall dependent-case system, since it is otherwise compatible with the Ban on Improper Case.

6 Deriving the Williams Cycle

While the Ban on Improper Case derives the range of facts presented in this paper, the fact that analogous restrictions have been observed for movement (e.g. Williams 1974, 2003, 2013; Müller and Sternefeld 1993, 1996; Abels 2007, 2009, 2012a,b; Neeleman and van de Koot 2010; Müller 2014a,b) and agreement (Keine 2016, 2019, 2020) strongly suggests that these “WC effects” have a *unified source*. Here, there are two interconnected issues: (i) how to formulate the WC so as to encompass case, movement, and agreement and (ii) how to derive the WC in the grammar.

The existing analyses of WC effects—other than Williams’s own—analyze the WC as the result of a constraint on either MERGE (Abels 2007, 2009; Müller 2014a,b) or AGREE (Keine 2016, 2019, 2020). Examining the specific details of these proposals would take us too far afield. What is important for present purposes is that they are *operation-specific*. For these proposals to extend to the case facts presented here, i.e. the Ban on Improper Case, dependent-case assignment would need to involve AGREE (or, in principle, MERGE). However, a dependent-case relation does not resemble an AGREE-relation, in that it does not involve any obvious valuation (or checking) of syntactic features between the two DPs. That is, it is not clear how the two DPs valuing features on each other (in either direction) would result in one of them being assigned dependent case. Thus, if case, movement, and agreement all exhibit WC effects and (dependent) case assignment does not involve AGREE, then WC effects must be the result of a more general (non-operation-specific) constraint in the grammar (as Keine 2020:332 also acknowledges).

The line of thinking that I advance in this section is that WC effects can be uniformly derived in an operation-general way if they are analyzed as the result of how clausal embedding works, as Williams (2003) originally proposed. The challenge for this approach is that it enforces a very strict locality. While this strict locality appears to be appropriate for case (the focus of this paper), it has been argued that it is empirically too restrictive for movement (Abels 2007, 2009); I will return to this issue in section 7.3.

Broadly construed, the WC is the notion that one and the same node can be a barrier to some dependencies, but not to other dependencies. I propose adopting the particularly strong formulation of the WC in (77).

(77) **WILLIAMS CYCLE** (*strong version*)

Within the current XP, a syntactic operation may not target an element across YP, where Y is higher than X in the functional sequence.

The formulation in (77) is operation-general; it does not reference specific operations and thereby covers case, movement, and agreement alike. It also encodes the strict locality of the Generalized Ban on Improper Movement (see section 5.1) and thus is commensurate with the formulation in Williams (2003, 2013). Accordingly, the Ban on Improper Case is a subset of (77).

An additional upshot of the formulation in (77) is that it is compatible with the various syntactic implementations of dependent-case assignment: a distinct syntactic operation (as assumed in this paper; see Preminger 2011, 2014; Poole 2015a), parasitic on cyclic linearization (Baker 2015), or binding relations (Pesetsky 2011). In other words, the WC does not require a particular analysis of dependent case, only that it occurs in the syntax (see section 7.1).

To derive the WC as formulated in (77), I suggest returning to Williams's (2003, 2013) own analysis of the WC, which derives the WC from the syntax of embedding (for an overview of this theory, see Hornstein and Nevins 2005). The core idea of the analysis is that embedding is constrained by the *Level Embedding Conjecture* (LEC) in (78).⁵¹

(78) **LEVEL EMBEDDING CONJECTURE (LEC)**

An XP can only be embedded in a structure that is also built up to an XP.

The basic idea behind the LEC is that clauses are built up in parallel. Embedding may take place at any point, but once a clause has been embedded, it no longer increases in size.⁵² The different points in the derivation at which embedding may take place correspond to the functional sequence. Williams calls this notion the *derivational clock* (or 'F-clock'). To illustrate, consider *that*-clause embedding in (79) (ignoring the *vP*). First, both clauses are built up to the VP-level; here, the embedding verb *think* merges with a *placeholder* for a CP-clause.⁵³ Second, both clauses are built up to the TP-level (79b). Third, both clauses are built up to the CP-level (79c). Last, embedding occurs: the embedded CP is substituted for the placeholder in the matrix clause (79d).

- (79) a. *Step 1: Build up to the VP-level*
 [_{VP} thinks CP], [_{VP} saw Sue]
- b. *Step 2: Build up to the TP-level*
 [_{TP} Mary thinks CP], [_{TP} John saw Sue]
- c. *Step 3: Build the CP-level*
 [_{CP} Mary thinks CP], [_{CP} that John saw Sue]
- d. *Step 4: Embed the CP in the other CP*
 [_{CP} Mary thinks [_{CP} that John saw Sue]]

Under the LEC, embedding is a substitution operation (though Williams does not explicitly call it such), analogous to Chomsky's (1955, 1957) theory of generalized transformations (for a reprise,

⁵¹ The formulation of the LEC in (78) adapts Williams's (2003) formulation into more standard parlance. His formulation is in terms of his Representation Theory, which organizes the grammar differently from standard minimalist syntax.

⁵² The LEC makes embedding countercyclic, under the standard minimalist incremental version of the cycle. See Williams (2003:70–71) for some discussion of the ramifications.

⁵³ For the sake of illustration, the placeholder for the embedded CP is depicted in (79) as a childless CP node, in the style of TAG-substitution. For Williams (2003:175–185), the placeholder is the head of the embedded clause, e.g. *that* in the case of (79). In addition to marking the point of substitution, the placeholder also serves to satisfy the selectional requirements of the embedding verb. Moreover, the placeholder itself could move within the matrix clause prior to embedding (though see Williams 2013:104–107), thereby allowing for short movement of clauses within *vP* (as Moulton 2015 and Bruening 2018 recently argue for). Thanks to two reviewers for raising these points.

(84) *Applied to dependent-case assignment*



As such, all syntactic dependencies are subject to the WC, regardless of whether or not they share the same operational core. All of the WC effects are thus uniformly derived from the timing of embedding.

Before concluding this section, it is worth briefly considering what counts as a ‘syntactic dependency’ in the context of the WC and the LEC. Recall from section 4.2 that in Finnish MA-infinitives, the matrix and embedded objects cannot establish a dependent-case relationship with each other (see (39)), but they can establish a binding relationship (see (40)). The former is predicted by the LEC, but the latter is not; the asymmetry between the two thus needs to be explained. I take a ‘syntactic dependency’ to be a dependency that exists in the narrow syntax, thereby being interspersed with structure building. By this definition, case, movement, and agreement are all syntactic dependencies. However, I contend that LF relations are not syntactic in this relevant, narrow sense. Because the LEC only constrains the narrow syntax, LF relations therefore do not exhibit WC effects. The justification for this claim is that under the LEC, semantic interpretation can only proceed after all embedding has taken place—because the embedded clause’s denotation is needed to compute the VP’s denotation, and so forth; thus, it is independently the case that LF must be computed on the basis of the *final* output of the narrow syntax. Crucially, the core of a binding dependency, i.e. the λ -operator and the variable that it binds, is a relation that only needs to hold at LF (Lebeaux 2009).⁵⁵ Returning to Finnish, the asymmetry between dependent case and binding in MA-infinitives, then, is because the dependent-case relationship would need to be established in the narrow syntax, which is not possible because of the LEC, while the binding relationship is established at LF and hence is unaffected by the LEC. Verification of this hypothesis would require examining other LF relations, such as scope and focus, the problem being that many (perhaps most) of these relations are clause-bounded and thus are not the kind of dependencies that the LEC would affect. I leave pursuing this topic to future research.⁵⁶

7 Discussion

This paper has argued that case assignment is subject to the *Ban on Improper Case* in (85). This constraint is an extension of the *Williams Cycle* (WC) particularized to case. The motivation for

⁵⁵ This is not to say that binding dependencies never involve a syntactic component, e.g. an AGREE-relation. However, this is not the case in Finnish: third-person possessive suffixes (the anaphors) do not reflect the ϕ -features of their binder; namely, there is no number distinction. First- and second-person possessive suffixes, on the other hand, do reflect number and are not anaphors, at least in the sense of being subject to Condition A (e.g. Vainikka 1989).

⁵⁶ Some of these issues are preliminarily explored in Williams (2003, 2011, 2013, 2014). In short, LF relations do not generally seem to obey the WC, but there are some crossclausal interactions with LF relations that are *reminiscent* of the WC. The overall picture, though, is unclear and deserves more attention.

improper case came from two disparate empirical domains: the interaction between case and movement and crossclausal case assignment in Finnish. Both of these locality problems were shown not to fall under the purview of standard notions of locality, e.g. phases, but rather they follow from the Ban on Improper Case.

(85) **BAN ON IMPROPER CASE**

A DP in [Spec, XP] cannot establish a dependent-case relationship with a lower DP across YP, where Y is higher than X in the functional sequence.

It was then shown that [Williams's \(2003\)](#) analysis of the WC in terms of clausal embedding uniformly captures WC effects for case, movement, and agreement. Thus, it crucially derives the Ban on Improper Case.

The remainder of this paper is devoted to discussing some of the issues that emerge from the Ban on Improper Case. Subsections 7.1 and 7.2 discuss two broader ramifications: the timing of case assignment and the WC's relation to phases, respectively. Finally, the issue of potential counterexamples to the strict locality enforced by the (strong) WC is taken up in subsection 7.3.

7.1 Timing of case assignment

An overarching question in the case literature is at what point in the derivation case assignment happens. In [Marantz's \(1991\)](#) original implementation of DCT, case assignment is situated at PF, i.e. in the postsyntactic morphological component. This line of thinking prevailed in the early work on DCT (e.g. [McFadden 2004](#); [Bobaljik 2008](#); an exception being [Bittner and Hale 1996](#)). It was also often considered a key difference between DCT and the more standard FHCT, which situates case assignment in the narrow syntax (see e.g. [Legate 2008](#)). However, more recent work on DCT has argued that, even under DCT, case assignment must be in the narrow syntax and not at PF ([Baker and Vinokurova 2010](#); [Preminger 2011, 2014](#); [Baker 2015](#)).

The Ban on Improper Case lends further support to the argument that case assignment must be in the narrow syntax. First, as movement is subject to the WC and movement occurs in the narrow syntax, the WC itself must be due to a constraint in the syntax in order for it to restrict movement so. It stands to reason then that anything else that is subject to the WC must be in the syntax as well. As such, because case assignment is subject to the WC, it too must be in the syntax. Second, the information required for the WC in the first place is fundamentally syntactic in nature, and replicating it at PF just so that it could apply to case assignment would be redundant.⁵⁷ Third, it was argued in section 3.2 that dependent-case assignment is interspersed with structure building, which would not follow if case assignment were at PF.

When in the narrow-syntactic derivation does case assignment happen then? In section 3.2, I proposed that dependent case is assigned as early as possible, a principle that I called Earliness. This notion of earliness also trivially extends to lexical case under DCT, where it is always

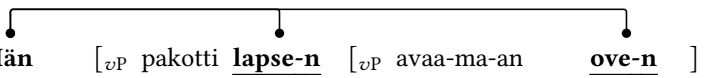
⁵⁷ [Legate \(2008\)](#) raises the same point about the dependent-case calculus in general, i.e. that it requires fundamentally syntactic information, though she takes this as an argument against DCT and in favor of FHCT.

assigned locally (see section 2). Earliness is the strongest hypothesis about the timing of case assignment, and it is fully consistent with the data presented in this paper. Another possibility, though, is that case assignment is delayed until phasal Spellout (as in Baker 2015). The effects of such an analysis largely depend on what the phases are. If vP is a phase, such an analysis is in principle possible, as long as case assignment precedes any movement to the phase edge, in order to avoid the problems solved by Earliness. However, such an analysis would have to grapple with the fact that vP does not seem to erect a locality domain for case in the same way as CP does; see the next section. If only CP is a phase, though, then delaying case assignment until phasal Spellout will require saying something special about A-scrambling, which would be phase-internal (e.g. to [Spec, TP]) on such an analysis, but does not affect case assignment. Fully exploring these issues is beyond the scope of this paper, so I leave them for future research. However, it should be pointed out that whether XP is a domain at which some operation applies and whether XP is a locality domain that blocks that same operation are in principle distinct questions, even if they are typically conflated in phase theory.

7.2 Phases

The WC and phases—the more standard notion of locality—are not mutually exclusive. They may coexist as independent constraints on syntactic operations. For instance, the WC does not force successive-cyclic movement through [Spec, CP]; this is still a consequence of phases.

It is standardly assumed that CP and vP are phases, and consequently that successive-cyclic movement targets [Spec, CP] and [Spec, vP] (Chomsky 2000, 2001, 2008). Throughout this paper though, I have tacitly assumed that only CP is a phase, because in Finnish, a dependent-case relationship can span an arbitrary number of intervening vP s, as illustrated in (86).⁵⁸

- (86)  (86) Hän [vP pakotti **lapse-n** [vP avaa-ma-an **ove-n**]] (=36a)
s/he.NOM forced child-ACC open-INF/MA-ILL door-ACC
'S/he forced the child to open the door' [Nelson 1998:238]

(86) shows, minimally, that dependent-case assignment is not subject to the Phase Impenetrability Condition (PIC) at the vP -level. There are two potential conservative explanations for this status, both of which are compatible with the Ban on Improper Case. The first is that dependent-case assignment is simply not subject to the PIC, as Bošković (2007) has argued about AGREE. The second is that the vP -phase does not intervene in the same way for dependent-case assignment as the CP-phase does, as Baker (2015) proposes with his 'soft'- and 'hard'-phase distinction.

There is also the more radical explanation that vP is not a phase. vP -phasehood, in fact, conflicts with the WC more generally. First, according to the WC, movement from [Spec, CP] to [Spec, vP] is barred because $C > v$ in *fseq*. Second, if such movement were permitted, it would

⁵⁸ For other problems that Finnish case poses for phase theory, see Brattico (2009).

obscure crucial distinctions needed to account for generalized improper movement. For example, consider hyperraising: at the point at which movement to [Spec, TP] occurs, the moving DP would be in [Spec, *v*P], so it would be necessary to backtrack into the previous phase to see whether it moved out of a CP or a TP. If the movement to [Spec, TP] proceeds directly from the CP/TP (see section 5.1), then such backtracking is unnecessary. For more discussion of this particular problem, see Müller (2014a,b). Based on (i) these kinds of considerations involving the WC and (ii) long-distance agreement configurations parallel to (86), Keine (2016, 2017, 2020) argues that *v*P should not be considered a phase (see also Keine and Zeijlstra 2020)—which would also solve the problems that the WC poses for *v*P-phasehood.

7.3 Potential exceptions to the Williams Cycle

As shown in section 6, the *Level Embedding Conjecture* (LEC) successfully derives the strong formulation of the WC, repeated in (87), thereby providing a uniform analysis of WC effects for case, movement, and agreement. I will refer to the formulation in (87) as the ‘strong’ WC.

(87) **WILLIAMS CYCLE** (*strong version*)

Within the current XP, a syntactic operation may not target an element across YP, where Y is higher than X in the functional sequence.

Abels (2007, 2009), however, has argued that the strong WC is empirically too restrictive because it rules out several purported movement dependencies. This criticism extends to the LEC, since it derives the strong WC. The recent, operation-specific analyses of WC effects have taken these purported exceptions at face value and gone on to develop analyses that derive weaker versions of the WC. As discussed in section 6, the dilemma with these analyses is that they do not extend to dependent case. Let us focus on Keine’s (2016, 2019, 2020) AGREE-based analysis, setting aside the MERGE-based analyses of Abels (2007, 2009) and Müller (2014a,b). Dependent-case relations do not resemble canonical AGREE-relations, and thus it is not immediately evident that dependent-case assignment involves AGREE. We are therefore at an impasse between two options: (i) develop a fully AGREE-based implementation of DCT, thereby allowing us to (in principle) extend Keine’s analysis to case, or (ii) revisit and reanalyze the purported exceptions to the strong WC, thereby allowing us to maintain the LEC.

I argue that the purported exceptions to the strong WC should be revisited and reanalyzed. My argument against the first option is twofold. First, no AGREE-based implementations of DCT have been proposed in the literature. Thus, given the current state-of-the-art in DCT, it is not presently possible to directly extend Keine’s analysis to case. Second, Keine’s analysis handles the exceptions largely through a stipulation: in a nutshell, some AGREE-probes are not subject to the WC (in his terms, they do not have a ‘horizon’).⁵⁹ In light of these two points, I contend that it is not at all certain that abandoning the strong WC—and by extension the


⁵⁹ This characterization of Keine’s system is too simplistic, missing some of the nuance in how it handles some of the purported exceptions to the strong WC, but space limitations prevent a more detailed discussion and comparison.

LEC—is warranted based on a set of limited exceptions, especially given the importance of the strong WC’s operation-generality. At the very least, the introduction of improper case into the empirical landscape warrants subjecting the purported exceptions to closer scrutiny.

Fully reconciling this issue is beyond the scope of this paper, but there are—I believe—promising directions towards reanalyzing the exceptions. In what follows, I briefly discuss each exception, the first two of which are from [Abels \(2007\)](#), and sketch how they might be reanalyzed in ways compatible with the LEC. For fair comparison, see the appendix where I sketch a basic AGREE-based implementation of DCT, and argue that the resulting system is baroque and undesirable.


ECM: In ECM infinitives, it is commonly assumed that the embedded subject moves from inside the embedded TP to a *vP*-internal position in the matrix clause, as schematized in (88) ([Postal 1974](#)). However, according to the strong WC, TP should be a barrier for such movement because $T > v$ in *fseq*. As such, ECM infinitives appear to pose a challenge for the strong WC. Note that under the WC, the matrix subject can establish a dependent-case relationship with matrix [Spec, *vP*] or embedded [Spec, TP], so the actual case in ECM is unproblematic.

(88) Alex believes [_{*vP*} **Taylor** (with all her heart) [_{TP} ___ to be guilty]].



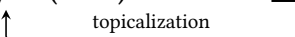
The classical evidence cited in favor of this analysis is that matrix adverbs and particles may intervene between the embedded subject and the embedded predicate, e.g. *with all her heart* in (88). However, recent work by [Neeleman and Payne \(2020\)](#) has reevaluated this argument. On the basis of scope-freezing effects and adverb order, they argue that an ECM infinitive does not actually involve moving the embedded subject, as in (88), but rather extraposing part of the embedded clause rightwards, as in (89). If [Neeleman and Payne](#)’s analysis is on the right track, then ECM infinitives do not pose a problem for the strong WC after all.

(89) Alex believes [_{TP} **Taylor** ___] (with all her heart) [to be guilty].

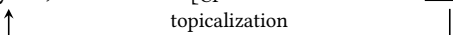


Movement over complementizers: In some languages, movement that lands below a complementizer is then able to cross that complementizer to move to a higher clause. To illustrate, consider English topicalization. In an embedded clause, topicalization lands in a position below the complementizer (90a), from which it could be concluded that $C > \text{Top}$ in *fseq* (where TopP represents whatever position topicalization targets). Topicalization can, however, cross an embedded finite clause boundary, moving over a complementizer (90b). If topicalization targets TopP and $C > \text{Top}$, the strong WC incorrectly predicts that CP should be a barrier for movement to TopP, thereby prohibiting topicalization over a complementizer.

(90) a. Alex thinks [_{CP} (**that**) **Taylor** (***that**) no one likes ___]



b. **Taylor**, Alex thinks [_{CP} **that** no one likes ___].



This class of exceptions would disappear if complementizers in these languages are analyzed as edge markers that uniformly linearize at the clause boundary, along the lines of [Manetta's \(2006, 2011\)](#) proposal for Hindi-Urdu *ki*. Under such an analysis, a moved element appearing to the right of a complementizer, like in (90a), would not entail that the complementizer corresponds to a projection higher than the landing site of movement and therefore would not constitute a violation of the strong WC if that movement can also cross the complementizer.

Hyperraising: Several languages have been claimed to allow hyperraising. This phenomenon has been most thoroughly investigated for Bantu languages (e.g. [Carstens 2011](#); [Diercks 2012](#); [Carstens and Diercks 2013](#); [Halpert 2015, 2019](#)), so I will center the discussion around them. A representative example of Bantu hyperraising from Lubukusu is given in (91).

- (91) **Babaandu**₁ ba-lolekhana [(mbo) ___₁ ba-kwa] [Lubukusu]
 2.people 2SA-seem that 2SA.PAST-fall
 'The people seem like they fell' [Carstens and Diercks 2013:100]

Because the WC expressly prohibits hyperraising, if (91) is indeed hyperraising, it is problematic for the WC. However, [Carstens and Diercks](#) observe a crucial interaction between hyperraising and complementizers, which suggests that this picture is too simplistic. They report on three Bantu languages: Digo, Lubukusu, and Lusaamia. Digo and Lusaamia crucially do *not* allow hyperraising over complementizers. On the other hand, some Lubukusu speakers allow hyperraising over complementizers, but only the complementizer *mbo* and not the agreeing complementizer *-li*. They analyze this pattern as follows: (i) CPs are generally barriers to hyperraising because they are phases; (ii) finite clauses without complementizers are TPs in Bantu, not CPs; and (ii) *mbo* in Lubukusu is special in that it is not a phase head, thereby projecting a nonphasal CP that is not a barrier to hyperraising.

Under the WC, TP is not a barrier for movement to [Spec, TP], since $T \not> T$ in *fseq*, irrespective of whether the TP is considered finite or nonfinite. Therefore, on [Carstens and Diercks's](#) analysis, hyperraising out of complementizer-less clauses is in fact compatible with the strong WC. This leaves *mbo*-clauses in Lubukusu. Rather than analyzing *mbo* as a special nonphasal complementizer, I suggest that *mbo* be analyzed as an edge marker, along the lines discussed above. Like their complementizer-less counterparts, *mbo*-clauses would then be TPs, and thus A-movement out of them would not violate the WC. Similar considerations can, I believe, be applied to the other purported cases of hyperraising, such as Brazilian Portuguese ([Nunes 2008](#)), Greek ([Alexiadou and Anagnostopoulou 2002](#)), and Zulu ([Halpert 2015, 2019](#)).

Long-distance agreement: There are several languages that have been reported to allow agreement between a matrix verb and a DP at the edge of an embedded finite clause, e.g. Innu-aimûn ([Branigan and MacKenzie 2002](#)), Passamaquoddy ([Bruening 2001](#)), and Tsez ([Polinsky and Potsdam 2001](#)). This is problematic for the WC because CP should be a barrier to a ϕ -probe on T^0 , because $C > T$ in *fseq*. However, these instances of long-distance agreement can be reanalyzed in a way compatible with the strong WC: (i) the embedded DP (i.e. the agreement controller) moves

to embedded [Spec, CP], (ii) the DP's features percolate up to CP via Spec–Head agreement, and (iii) matrix T^0 agrees with the CP. This analysis is similar in spirit to [Koopman's \(2006\)](#) analysis of Tsez long-distance agreement, in that there is no direct crossclausal agreement. In terms of the LEC, matrix T^0 would agree with the CP before the full CP has been embedded (= substituted in); thus, upon embedding the CP, the CP's features must be shared along (or match) its existing AGREE-relations. Similar analyses can, I believe, be extended to *wh*-agreement in Chamorro and Palauan (e.g. [Chung 1982, 1994](#); [Chung and Georgopoulos 1988](#)) and to crossclausal object agreement in Nez Perce (see [Deal 2017a](#), who analyses it in terms of covert movement, however).

Sakha accusative subjects: In Sakha, an embedded subject can be assigned dependent case (= accusative) iff the matrix clause has another DP ([Baker and Vinokurova 2010](#)). [Baker and Vinokurova](#) analyze this pattern in terms of raising: the embedded subject is eligible to move to embedded [Spec, CP], where it may then enter into dependent-case relationships with DPs in the matrix clause. This analysis is problematic for the strong WC (and the Ban on Improper Case) because it involves a DP in [Spec, CP] being the lower DP in a dependent-case pair, which should be impossible (see section 5.4).

- (92) min **ehigi(-ni)**₁ [бүгүн ___₁ kyaj-yax-xyt dien] erem-mit-im [Sakha]
 I.NOM you -ACC today win-FUT-2PL.SA that hope-PAST-1SG.SA
 'I hoped that you would win today' [[Baker and Vinokurova 2010:615](#)]

I argue that so-called accusative subjects in Sakha are actually proleptic arguments: they are base-generated as an argument of the matrix clause and are indirectly linked to an embedded gap via resumption (in the spirit of [Den Dikken 2017, 2018](#)). As an argument of the matrix clause, it participates in the dependent-case calculus in the matrix clause, and thus is sensitive to the DPs there. [Baker and Vinokurova](#) themselves consider a prolepsis analysis of accusative subjects. They claim that while some instances of accusative subjects are indeed prolepsis, there are at least some instances that are not. To support this claim, they show that an accusative subject can be an NPI that would only be licensed in the embedded clause. They argue that this constitutes evidence against adopting a prolepsis analysis across the board for accusative subjects. However, [Den Dikken \(2017, 2018\)](#) explicitly argues that prolepsis does in fact allow NPI licensing. (Technically for him, such constructions are *complex predicates* with no crossclausal syntactic dependencies, the same analysis that [Den Dikken 2009](#) gives for Hungarian long focus movement; see section 3.3.) Thus, the NPI facts are in fact compatible with the prolepsis analysis. Evidence ruling out prolepsis would have to come from other reconstruction data, which are not available in the literature. Crucially, under a prolepsis analysis, Sakha accusative subjects are not problematic for the strong WC—as no crossclausal syntactic dependencies are involved.

As shown with the above discussion, it is not at all clear that these phenomena constitute evidence against the strong WC and the LEC. At the very least, in light of improper case, they deserve more attention and careful scrutiny. If the reanalyses sketched above can be sustained, then the LEC can be maintained in its full strength.

Abbreviations

ACC	accusative	ILL	illative	PAST	past
ADE	adessive	IMP	imperative	PCL	particle
ALL	allative	INDEF	indefinite	PFV	perfective
APPL	applicative	INE	inessive	PL	plural
DAT	dative	INF/MA	MA-infinitive	POSS	possessive
DEF	definite	INF/TA	TA-infinitive	PTV	partitive
ERG	ergative	NOM	nominative	SG	singular
FUT	future	OBL	oblique	SA	subject agreement
GEN	genitive	PASS	passive		

Acknowledgements

Thanks to Arto Anttila, Rajesh Bhatt, Claire Halpert, Kyle Johnson, Stefan Keine, Paul Kiparsky, Julie Anne Legate, Travis Major, Mark Norris, Ellen Woolford, Michelle Yuan, and audiences at GLOW 39, WCCFL 34, the University of Massachusetts Amherst, UCLA, McGill, and the University of Leipzig for helpful discussion and feedback. Special thanks to three anonymous reviewers and Amy Rose Deal for very helpful comments on an earlier draft of this paper.

References

- Abels, Klaus. 2007. Towards a restrictive theory of (remnant) movement. In *Linguistic Variation Yearbook 7*, eds. Jeroen van Craenenbroeck and Johan Rooryck, 53–120. Amsterdam: John Benjamins.
- Abels, Klaus. 2009. Some implications of improper movement for cartography. In *Alternatives to Cartography*, ed. Jeroen van Craenenbroeck, 325–359. Berlin: de Gruyter.
- Abels, Klaus. 2012a. The Italian left periphery: A view from locality. *Linguistic Inquiry* 43:229–254.
- Abels, Klaus. 2012b. *Phases: An Essay on Cyclicity in Syntax*. Berlin: de Gruyter.
- Abramovitz, Rafael. 2020. Successive-cyclic *wh*-movement feeds case competition in Koryak. Ms., MIT.
- Alexiadou, Artemis, and Elena Anagnostopoulou. 2002. Raising without infinitives and the nature of agreement. In *Dimensions of Movement: From Features to Remnants*, eds. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers, and Hans-Martin Gärtner, 17–30. Amsterdam: Benjamins.
- Anttila, Arto, and Jong-Bok Kim. 2011. On structural case in Finnish and Korean. *Lingua* 121:100–127.

- Anttila, Arto, and Jong-Bok Kim. 2017. Locality and variation in Finnish structural case. *Natural Language and Linguistic Theory* 35:581–634.
- Bach, Emmon. 1979. Control in Montague Grammar. *Linguistic Inquiry* 10:515–531.
- Baker, Mark. 2014. On dependent ergative case (in Shipibo) and its derivation by phase. *Linguistic Inquiry* 45:341–379.
- Baker, Mark. 2015. *Case*. Cambridge, UK: Cambridge University Press.
- Baker, Mark, and Jonathan Bobaljik. 2017. On inherent and dependent theories of ergative case. In *The Oxford Handbook of Ergativity*, eds. Jessica Coon, Diane Massam, and Lisa Travis, 111–134. New York: Oxford University Press.
- Baker, Mark, and Nadya Vinokurova. 2010. Two modalities of case assignment: Case in Sakha. *Natural Language and Linguistic Theory* 28:593–642.
- Baltin, Mark. 1978. Toward a theory of movement rules. Ph.D. dissertation, MIT, Cambridge, MA.
- Barss, Andrew. 1986. Chains and anaphoric dependence. Ph.D. dissertation, MIT, Cambridge, MA.
- Bayer, Josef, and Martin Salzmann. 2013. *That*-trace effects and resumption: How improper movement can be repaired. In *Repairs: The Added Value of Being Wrong*, eds. Patrick Brandt and Eric Fuß, 275–333. Berlin: de Gruyter.
- Bech, Gunnar. 1955/1957. *Studien über das deutsche Verbum Infinitum*. Copenhagen: Munksgaard. Two volumes. Reprint Niemeyer, Tübingen in 1983.
- Bhatt, Rajesh, and Elena Anagnostopoulou. 1996. Object shift and specificity: Evidence from *ko*-phrases in Hindi. In *Proceedings of the 32nd Annual Meeting of the Chicago Linguistic Society (CLS 32)*, eds. Lise M. Dobrin, Kora Singer, and Lisa McNair, 11–22. Chicago, IL: Chicago Linguistics Society.
- Bierwisch, Manfred. 1963. *Grammatik des deutschen Verbs*. Berlin: Akademie Verlag.
- Bittner, Maria, and Kenneth Hale. 1996. The structural determination of case and agreement. *Linguistic Inquiry* 27:1–68.
- Bobaljik, Jonathan. 2008. Where's phi? Agreement as a post-syntactic operation. In *Phi-theory: Phi features across interfaces and modules*, eds. Daniel Harbour, David Adger, and Susana Béjar, 295–328. Oxford: Oxford University Press.
- Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, MA: MIT Press.
- Bošković, Željko. 2002. A-movement and the EPP. *Syntax* 5:167–218.
- Bošković, Željko. 2007. Agree, phases, and intervention effects. *Linguistic Analysis* 54–96.
- Branigan, Phil, and Marguerita MacKenzie. 2002. Altruism, A'-movement and object agreement in Innu-aimûn. *Linguistic Inquiry* 33:385–407.
- Brattico, Pauli. 2009. Long-distance case assignment in Finnish and the theory of phases. *Biolinguistica Fennica Working Papers* 1:79–104.
- Brattico, Pauli. 2012. Long distance case assignment in Finnish. *Studia Linguistica* 66:245–284.
- Bruening, Benjamin. 2001. Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy. Ph.D. dissertation, MIT, Cambridge, MA.

- Bruening, Benjamin. 2018. CPs move rightward, not leftward. *Syntax* 21:362–401.
- Cable, Seth. 2007. The grammar of Q: Q-particles and the nature of *wh*-fronting, as revealed by the *wh*-questions of Tlingit. Ph.D. dissertation, MIT, Cambridge, MA.
- Cable, Seth. 2010. *The Grammar of Q: Q-Particles, Wh-Movement and Pied-Piping*. Oxford: Oxford University Press.
- Carstens, Vicki. 2011. Hyperactivity and hyperagreement in Bantu. *Lingua* 121:721–741.
- Carstens, Vicki, and Michael Diercks. 2013. Parameterizing case and activity: Hyper-raising in Bantu. In *Proceedings of the 40th Meeting of the North East Linguistic Society (NELS 40)*, eds. Seda Kan, Claire Moore-Cantwell, and Robert Staubs, 99–118. Amherst, MA: GLSA.
- Chierchia, Gennaro. 1989. Anaphora and attitudes *de se*. In *Language in context*, eds. Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, 1–31. Dordrecht: Foris.
- Chomsky, Noam. 1955. The logical structure of linguistic theory. Ms., Harvard & MIT. Published in part by Plenum, New York, 1975.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1973. Conditions on transformations. In *A Festschrift for Morris Halle*, eds. Stephen Anderson and Paul Kiparsky, 232–286. New York: Academic Press.
- Chomsky, Noam. 1977. On *wh*-movement. In *Formal syntax*, eds. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995a. Bare phrase structure. In *Government and Binding Theory and the Minimalist Program*, ed. Gert Webelhuth, 383–439. Cambridge, MA: Blackwell.
- Chomsky, Noam. 1995b. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, eds. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, eds. Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.
- Chung, Sandra. 1982. Unbounded dependencies in Chamorro grammar. *Linguistic Inquiry* 13:39–77.
- Chung, Sandra. 1994. *Wh*-agreement and “referentiality” in Chamorro. *Linguistic Inquiry* 25:1–44.
- Chung, Sandra, and Carol Georgopoulos. 1988. Agreement with gaps in Chamorro and Palauan. In *Agreement in Natural Language: Approaches, theories, description*, eds. Michael Barlow and Charles A. Ferguson, 251–267. Stanford, CA: CSLI Publications.
- Csirmaz, Aniko. 2005. Semantics and phonology in syntax. Ph.D. dissertation, MIT, Cambridge, MA.
- Deal, Amy Rose. 2015. Ergativity. In *Syntax – Theory and Analysis. An International Handbook*, eds. Artemis Alexiadou and Tibor Kiss, volume 1, 654–707. Berlin: Mouton de Gruyter.

- Deal, Amy Rose. 2017a. Covert hyperraising to object. In *Proceedings of the 47th Meeting of the North East Linguistic Society (NELS 47)*, eds. Andrew Lamont and Katie Tetzloff, 257–270. Amherst, MA: GLSA.
- Deal, Amy Rose. 2017b. Syntactic ergativity as case discrimination. In *Proceedings of the 34th West Coast Conference on Formal Linguistics (WCCFL 34)*, eds. Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel, and Edward J. Rubin, 141–150. Somerville, MA: Cascadilla Proceedings Project.
- Deal, Amy Rose. 2019. Raising to ergative: Remarks on applicatives of unaccusatives. *Linguistic Inquiry* 50:388–415.
- Diercks, Michael. 2012. Parameterizing case: Evidence from Bantu. *Syntax* 15:253–286.
- den Dikken, Marcel. 2009. On the nature and distribution of successive cyclicity. Ms., CUNY Graduate Center.
- den Dikken, Marcel. 2017. Predication in the syntax of hyperraising and copy raising. *Acta Linguistica Academica* 64:3–43.
- den Dikken, Marcel. 2018. Secondary predication and the distribution of raising to object. *Acta Linguistica Academica* 65:87–117.
- den Dikken, Marcel. to appear. On the strategies for forming long A'-dependencies: Evidence from Hungarian. In *Minimalist Approaches to Syntactic Locality*, ed. Balázs Surányi. Cambridge, UK: Cambridge University Press.
- Dowty, David. 1985. On recent analyses of the semantics of control. *Linguistics and Philosophy* 8:291–331.
- É. Kiss, Katalin. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.
- Gervain, Judit. 2009. Resumption in focus(-raising). *Lingua* 119:687–707.
- Haider, Hubert. 1984. Topic, focus, and V-second. *Groninger Arbeiten zur Germanistischen Linguistik* 25:72–120.
- Hakulinen, Auli, Maria Vilkuna, Riita Korhonen, Vesa Koivisto, Tarja Riita Heinonen, and Irja Alho. 2004. *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Halpert, Claire. 2015. *Argument licensing and agreement*. Oxford: Oxford University Press.
- Halpert, Claire. 2019. Raising, unphased. *Natural Language and Linguistic Theory* 37:123–165.
- Heinämaa, Orvokki. 1984. Aspect in Finnish. In *Aspect bound: A voyage into the realm of Germanic, Slavonic and Finno-Ugrian aspectology*, 153–177. Dordrecht: Foris.
- Hornstein, Norbert, and Andrew Nevins. 2005. Review of “Representation Theory” by Edwin Williams. *Language* 81:757–761.
- Huhmarniemi, Saara. 2012. Finnish A'-movement: Edges and islands. Ph.D. dissertation, University of Helsinki.
- Jánosi, Adrienn. 2013. Long split focus constructions in Hungarian with a view on speaker variation. Ph.D. dissertation, KU Leuven/HU Brussels.
- Jánosi, Adrienn, Jeroen van Craenenbroeck, and Guido Vanden Wyngaerd. 2014. Long split focalization in Hungarian and the typology of A'-dependencies. *Lingua* 150:117–136.
- Jenks, Peter, and Hannah Sande. 2017. Dependent accusative case and caselessness in Moro. In *Proceedings of the 47th Meeting of the North East Linguistic Society (NELS 47)*, eds. Andrew

- Lamont and Katie Tetzloff, volume 2, 109–119. Amherst, MA: GLSA.
- Johnson, Kyle. 2012. Towards deriving differences in how *wh*-movement and QR are pronounced. *Lingua* 122:529–553.
- Joshi, Aravind K., Leon S. Levy, and Masako Takahashi. 1975. Tree adjunct grammars. *Journal of Computer and System Sciences* 10:136–163.
- Kayne, Richard. 1981. On certain differences between French and English. *Linguistic Inquiry* 12:349–371.
- Kayne, Richard. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.
- Keine, Stefan. 2016. Probes and their horizons. Ph.D. dissertation, University of Massachusetts, Amherst, MA.
- Keine, Stefan. 2017. Agreement and vP phases. In *A Schrift to Fest Kyle Johnson*, eds. Nicholas LaCara, Keir Moulton, and Anne-Michelle Tessier, 177–185. Amherst, MA: Linguistics Open Access Publications.
- Keine, Stefan. 2019. Selective opacity. *Linguistic Inquiry* 50:13–62.
- Keine, Stefan. 2020. *Probes and Their Horizons*. Cambridge, MA: MIT Press.
- Keine, Stefan, and Hedde Zeijlstra. 2020. Morphology of extraction: Reappraising vP phases. Handout from talk presented at Generative Linguistics in the Old World 43 (GLOW 43).
- Kiparsky, Paul. 2001. Structural case in Finnish. *Lingua* 111:315–376.
- Koopman, Hilda. 2006. Agreement configurations: In defense of “Spec head”. In *Agreement systems*, ed. Cedric Boeckx, 159–199. Amsterdam: John Benjamins.
- Koopman, Hilda, and Dominique Sportiche. 2014. The *que/qui* alternation: New analytical directions. In *Functional Structure From Top to Toe*, ed. Peter Svenonius, 46–96. Oxford: Oxford University Press.
- Kornfilt, Jaklin, and Omer Preminger. 2015. Nominative as no case at all: An argument from raising-to-ACC in Sakha. In *Proceedings of the 9th Workshop on Altaic Formal Linguistics (WAFL 9)*, eds. Andrew Joseph and Esra Predolac, 109–120. Cambridge, MA: MITWPL.
- Koskinen, Päivi. 1998. Features and categories: Non-finite constructions in Finnish. Ph.D. dissertation, University of Toronto.
- Kroch, Anthony, and Aravind K. Joshi. 1985. The linguistic relevance of Tree Adjoining Grammar. Technical Report MS-CIS-85-16, University of Pennsylvania.
- Landau, Idan. 2000. *Elements of control: Structure and meaning in infinitival constructions*. Dordrecht: Kluwer.
- Lasnik, Howard, and Nicholas Sobin. 2000. The *who/whom* puzzle: On the preservation of an archaic feature. *Natural Language and Linguistic Theory* 18:343–371.
- Lebeaux, David. 1988. Language acquisition and the form of the grammar. Ph.D. dissertation, University of Massachusetts, Amherst, MA.
- Lebeaux, David. 2009. *Where does Binding Theory apply?*. Cambridge, MA: MIT Press.
- Lechner, Winfried. 2003. Phrase structure paradoxes, movement and ellipsis. In *The Interfaces: Deriving and interpreting omitted structures*, eds. Kerstin Schwabe and Susanne Winkler, 177–203. Amsterdam: John Benjamins.
- Legate, Julie Anne. 2002. Warlpiri: Theoretical implications. Ph.D. dissertation, MIT, Cambridge,

- MA.
- Legate, Julie Anne. 2008. Morphological and abstract case. *Linguistic Inquiry* 39:55–101.
- Mahajan, Anoop. 1990. The A/A-bar distinction and movement theory. Ph.D. dissertation, MIT, Cambridge, MA.
- Maling, Joan. 1993. Of nominative and accusative. In *Case and other functional categories in Finnish syntax*, eds. Anders Holmberg and Urpo Nikanne, 49–74. Berlin: Mouton de Gruyter.
- Manetta, Emily. 2006. Peripheries in Kashmiri and Hindi-Urdu. Ph.D. dissertation, University of California, Santa Cruz, CA.
- Manetta, Emily. 2011. *Peripheries in Kashmiri and Hindi-Urdu: The Syntax of Discourse-Driven Movement*. Amsterdam: John Benjamins.
- Marantz, Alec. 1991. Case and licensing. In *Proceedings of the 8th Eastern States Conference on Linguistics (ESCOL 8)*, eds. German Westphal, Benjamin Ao, and Hee-Rahk Chae, 234–253. Ithaca, NY: CLC Publications.
- Massam, Diane. 1985. Case theory and the Projection Principle. Ph.D. dissertation, MIT, Cambridge, MA.
- Massam, Diane. 2000. VSO and VOS: Aspects of Niuean word order. In *The syntax of verb initial languages*, eds. Andrew Carnie and Eithne Guilfoyle, 97–116. Oxford: Oxford University Press.
- Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19:153–197.
- May, Robert. 1979. Must comp-to-comp movement be stipulated? *Linguistic Inquiry* 10:719–725.
- McFadden, Thomas. 2004. The position of morphological case in the derivation: A study on the syntax-morphology interface. Ph.D. dissertation, University of Pennsylvania, Philadelphia, PA.
- Milsark, Gary. 1985. Case theory and the grammar of Finnish. In *Proceedings of the 15th Meeting of the North East Linguistic Society (NELS 15)*, eds. Stephen Berman, Jae-Woong Choe, and Joyce McDonough, 319–331. Amherst, MA: GLSA.
- Mitchell, Erika. 1991. Case and the Finnish object. *Cornell Working Papers in Linguistics* 9:193–228.
- Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In *Approaches to Natural Language*, eds. Jaako Hintikka, Julius Moravcsik, and Patrick Suppes, 221–242. Dordrecht: Dordrecht.
- Moulton, Keir. 2015. CPs: Copies and compositionality. *Linguistic Inquiry* 46:305–342.
- Müller, Gereon. 2011. *Constraints on displacement: A phase-based approach*. Amsterdam: John Benjamins.
- Müller, Gereon. 2014a. A local approach to the Williams Cycle. *Lingua* 140:117–136.
- Müller, Gereon. 2014b. *Syntactic buffers*, volume 91 of *Linguistische Arbeitsberichte*. Universität Leipzig: Institut für Linguistik.
- Müller, Gereon. 2017. Structure removal: An argument for feature-driven Merge. *Glossa* 2:1–35.
- Müller, Gereon, and Wolfgang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24:461–507.
- Müller, Gereon, and Wolfgang Sternefeld. 1996. A-bar chain formation and economy of derivation. *Linguistic Inquiry* 27:480–511.

- Neeleman, Ad, and Hans van de Koot. 2010. A local encoding of syntactic dependencies and its consequences for the theory of movement. *Syntax* 13:331–372.
- Neeleman, Ad, and Amanda Payne. 2020. On matrix-clause intervention in accusative-and-infinitive constructions. *Syntax* 23:1–41.
- Nelson, Diane. 1998. *Grammatical case assignment in Finnish*. New York: Garland.
- Nunes, Jairo. 2008. Inherent case as a licensing condition for A-movement: The case of hyper-raising constructions in Brazilian Portuguese. *Journal of Portuguese Linguistics* 7:83–108.
- Pearson, Hazel. 2013. The sense of self: Topics in the semantics of *de se* expressions. Ph.D. dissertation, Harvard, Cambridge, MA.
- Penttilä, Aarni. 1963. *Suomen kielioppi*. Helsinki: WSOY.
- Pesetsky, David. 2011. Dependent case as Binding Theory. Handout from talk presented at the 42nd Meeting of the North East Linguistic Society (NELS 42).
- Polinsky, Maria, and Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19:583–646.
- Poole, Ethan. 2015a. A configurational account of Finnish case. In *Proceedings of the Penn Linguistics Conference 38 (PLC 38)*, ed. David Wilson. Philadelphia, PA: University of Pennsylvania.
- Poole, Ethan. 2015b. Deconstructing quirky subjects. In *Proceedings of the 45th Meeting of the North East Linguistic Society (NELS 45)*, eds. Thuy Bui and Deniz Özyıldız, volume 2, 247–256. Amherst, MA: GLSA.
- Poole, Ethan. 2017. Movement and the semantic type of traces. Ph.D. dissertation, University of Massachusetts, Amherst, MA.
- Postal, Paul. 1974. *On Raising*. Cambridge, MA: MIT Press.
- Preminger, Omer. 2011. Agreement as a fallible operation. Ph.D. dissertation, MIT, Cambridge, MA.
- Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press.
- Preminger, Omer. 2017. No case for agreement (as a causer of case). Talk given at the 48th Meeting of the North East Linguistic Society (NELS 48).
- Rezac, Milan. 2003. The fine structure of Cyclic Agree. *Syntax* 6:156–182.
- van Riemsdijk, Henk, and Edwin Williams. 1981. NP structure. *The Linguistic Review* 1:171–218.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.
- Ross, John R. 1967. Constraints on variables in syntax. Ph.D. dissertation, MIT, Cambridge, MA.
- Safir, Ken. 2019. The A/ \bar{A} -distinction as an epiphenomenon. *Linguistic Inquiry* 50:285–336.
- Schütze, Carson T. 2001. On the nature of default case. *Syntax* 4:205–238.
- Stanton, Juliet. 2016. Wholesale Late Merger in A'-movement: Evidence from preposition stranding. *Linguistic Inquiry* 47:89–126.
- Taraldsen, Knut Tarald. 1986. On the distribution of nominative objects in Finnish. In *Features and projections*, eds. Pieter Muysken and Henk van Riemsdijk, 139–161. Dordrecht: Foris.
- Thráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan í Lon Jacobsen, and Zakaris Svabo Hansen. 2004. *Faroese: An overview and reference grammar*. Tórshavn: Føroya Fróðskaparfelag.
- Timberlake, Alan. 1975. The nominative object in Finnish. *Lingua* 35:205–230.
- Toivainen, Jorma. 1993. The nature of the accusative in Finnish. In *Case and other functional*

- categories in Finnish syntax*, eds. Anders Holmberg and Urpo Nikanne, 111–128. Berlin: Mouton de Gruyter.
- Toivonen, Ida. 1995. A study of Finnish infinitives. Master's thesis, Brandeis University, Waltham, MA.
- Tuomikoski, Risto. 1978. Objektinsijaisista määrän adverbiaalesita. *Virittäjä* 82:26–51.
- van Urk, Coppe. 2015. A uniform syntax for phrasal movement: A case study of Dinka Bor. Ph.D. dissertation, MIT, Cambridge, MA.
- Vainikka, Anne. 1989. Deriving syntactic representations in Finnish. Ph.D. dissertation, University of Massachusetts, Amherst, MA.
- Vainikka, Anne. 1993. The three structural cases in Finnish. In *Case and other functional categories in Finnish syntax*, eds. Anders Holmberg and Urpo Nikanne, 129–159. Berlin: Mouton de Gruyter.
- Vainikka, Anne. 1995. Functional projections in Finnish non-finite constructions. *University of Pennsylvania Working Papers in Linguistics* 2:141–146.
- Vainikka, Anne, and Pauli Brattico. 2014. The Finnish accusative: Long-distance case assignment under agreement. *Linguistics* 52:73–124.
- Williams, Edwin. 1974. Rule ordering in syntax. Ph.D. dissertation, MIT, Cambridge, MA.
- Williams, Edwin. 2003. *Representation theory*. Cambridge, MA: MIT Press.
- Williams, Edwin. 2011. *Regimes of Derivation in Syntax and Morphology*. New York: Routledge.
- Williams, Edwin. 2013. Generative semantics, generative morphosyntax. *Syntax* 16:77–108.
- Williams, Edwin. 2014. Compositionality and derivation. Ms., Princeton University.
- Woolford, Ellen. 1997. Four-way case systems: Ergative, nominative, objective, and accusative. *Natural Language and Linguistic Theory* 15:181–227.
- Woolford, Ellen. 2006. Lexical case, inherent case, and argument structure. *Linguistic Inquiry* 37:111–130.
- Woolford, Ellen. 2015. Ergativity and transitivity. *Linguistic Inquiry* 46:489–531.
- Wurmbrand, Susi. 2001. *Infinitives: Restructuring and Clause Structure*. New York: Mouton de Gruyter.
- Yip, Moira, Joan Maling, and Ray Jackendoff. 1987. Case in tiers. *Language* 63:217–250.
- Yuan, Michelle. 2018. Dimensions of ergativity in Inuit: Theory and microvariation. Ph.D. dissertation, MIT, Cambridge, MA.
- Yuan, Michelle. 2020. Ergativity and object shift across Inuit. Ms., UC San Diego.

Appendix: AGREE-based dependent-case assignment

The alternative to reanalyzing the exceptions to the strong WC and the LEC is to try to extend [Keine's \(2016, 2019, 2020\)](#) AGREE-based analysis of the WC to improper case (setting aside that he assumes a weaker version of the WC). As discussed above, the foremost challenge to this approach is that it would preliminarily require that AGREE underlie dependent-case assignment—and a standard AGREE-relation does not obviously resemble a dependent-case relation. Against this backdrop, I sketch below the contours of what an AGREE-based implementation of dependent-case assignment would (presumably) have to look like. This sketch should not be viewed as a bona fide proposal, but rather as a proof of concept. The AGREE-based dependent-case system rests on several technologies that have all been proposed independently in the literature:

- Features may be CASE-DISCRIMINATING, whereby they are relativized to positions on the following hierarchy: unmarked >> dependent >> lexical ([Bobaljik 2008](#); [Preminger 2011, 2014](#); [Poole 2015b](#); [Deal 2017b](#)). For example, a feature relativized to ‘dependent’ can establish an AGREE-relation with a DP marked with dependent or unmarked case, but not lexical case.
- A feature’s search space may be restricted to either just siblinghood or its entire c-command domain, which I notate as $[F]_s$ and $[F]_c$ respectively. (This is presumably needed independently for siblinghood properties, like selection).
- Following Bare Phrase Structure, the nodes labelled XP, \bar{X} , and X^0 share the same features, as what projects is the head itself ([Chomsky 1995a](#)), thereby allowing specifier–head agreement as siblinghood agreement ([Rezac 2003](#)) (see also fn. 5).
- If a feature can be valued, then it must be valued, but unvalued features do not crash the derivation ([Preminger 2011, 2014](#)).
- Features may be ordered on a head $\langle x_1, x_2, \dots, x_n \rangle$, where x_m is active only if x_{m-1} has first been satisfied ([Müller 2011](#)).

Dependent case is assigned via AGREE with the feature stack $\langle [UNM], [DEP] \rangle$, which is borne by every functional head. $[UNM]$ is a simple agreement probe, relativized to unmarked case such that it can only establish an AGREE-relation with a DP whose $[CASE: \square]$ feature is unvalued; it does not assign any case itself. $[DEP]$ assigns dependent case. $[UNM]$ encodes the dependency requirement, i.e. that another DP be present in order to assign dependent case, because $[DEP]$ only becomes active *after* $[UNM]$ has itself been satisfied. In other words, satisfying $[UNM]$ “unlocks” the ability to assign dependent case with $[DEP]$. Not satisfying the features, though, does not affect or crash the derivation. The two different case alignments are the result of parameterizing the search space of the two features in the stack. In a nominative–accusative alignment, the case-assignment stack is $\langle [UNM]_s, [DEP]_c \rangle$, where $[UNM]$ ’s search space is restricted to siblinghood and $[DEP]$ ’s search space is its entire c-command domain (93). In an ergative–absolutive alignment, it is the reverse: $\langle [UNM]_c, [DEP]_s \rangle$, where $[UNM]$ ’s search space is its entire c-command domain and $[DEP]$ ’s search space is restricted to siblinghood (94).

- (93) *Nominative–accusative alignment*
- $$[{}_{vP} DP_{\alpha} v^0 \langle [UNM]_s, [DEP]_c \rangle [VP V^0 DP_{\beta}]]$$
-
- (94) *Ergative–absolutive alignment*
- $$[{}_{vP} DP_{\alpha} v^0 \langle [UNM]_c, [DEP]_s \rangle [VP V^0 DP_{\beta}]]$$
-

In a nominative–accusative alignment, depicted in (93), $[UNM]_s$ is satisfied by DP_{α} as sibling to v^0 , this unlocks $[DEP]_c$, and then $[DEP]_c$ assigns dependent case to DP_{β} . In an ergative–absolutive alignment, depicted in (94), $[UNM]_c$ is satisfied by DP_{β} in its c–command domain, this unlocks $[DEP]_s$, and then $[DEP]_s$ assigns dependent case to DP_{α} , its sibling. Under this system, dependent-case assignment is reduced to AGREE and a particular ordered set of features, and its directionality results from different locality profiles for the features in that ordered set. It should be noted that this system is reminiscent of Marantz’s (1991) original implementation of dependent-case assignment, which requires both DPs to be governed by V+I.

What this system demonstrates is that dependent-case assignment could *in principle* be reduced to an AGREE-based procedure. Adopting such a system would in turn allow Keine’s (2016, 2019, 2020) AGREE-based analysis of the WC to be extended to case (though some details would need to be worked out). However, there is no denying that the system sketched above is baroque and, I would argue, undesirable. It cashes out dependent-case assignment as a seemingly accidental combination of features and locality profiles and thus misses the underlying generalization that two DPs are required for dependent case. We are also left to wonder why such complex feature stacks are not utilized elsewhere in the syntax (though perhaps they are). This baroque stands in stark contrast to the elegance of the analysis in terms of the LEC. As such, I contend that in light of improper case, the LEC is the preferable analysis of WC effects.