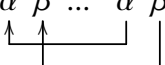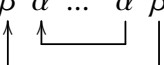# What Parasitic Gaps Reveal about Overlapping Paths and Multiple Specifier Formation[*]

Colin Davis / USC / colind@usc.edu

Abstract: This paper investigates the nature of sentences in which A′-movement paths overlap, using parasitic gaps as a central diagnostic. Primarily, I argue building from previous work on parasitic gaps for the following generalization: If two A′-moving phrases XP1 and XP2 form specifiers of a phase YP such that XP1 c-commands XP2, then XP1 c-commands XP2 in all subsequent positions these phrases occupy. I argue that this result emerges from a theory in which cross-phasal movement is constrained by cyclic spell-out and linearization preservation, with consequences for topics such as the path containment condition and the nature of multiple specifier formation.

1 Introduction

In this paper, I investigate certain properties of sentences in which (A′-)movement paths overlap. I argue that the behavior of parasitic gaps (PGs) such sentences clarifies their derivational history, and in doing so, leads to several consequences for syntactic theory. An overlapping movement configuration minimally involves two moving phrases, for which there are two possible outcomes in principle: either the two phrases' paths will cross, resulting in no change in their order (1a), or their paths will nest, reversing their order (1b).

(1)   a.   *Crossing paths*   b.   *Nesting paths*

      α β ... α β β α ... α β

      (Order-preserving)   (Order-reversing)

While it is conceivable that only one of these scenarios might be attested in natural language, in reality both are, as the following examples illustrate:

(2)   *Examples of crossing paths* (XP$_1$ YP$_2$ ... t_1 t_2)

        a.   *Bulgarian multiple wh-movement* (Richards 1997 p. 277 ex. 75a)

           Kogo$_1$ **kakvo**$_2$ e   pital   Ivan t_1 t_2?
           who.NOM what.ACC AUX asked Ivan

           'What did Ivan ask who?'

b. *Multiple wh-scrambling in Japanese* (Richards 1997 p. 63 ex. 12b)

Dare-ni$_1$ **nani-o**$_2$ John-ga [Takana-sensee-ga t_1 t_2 yomaseta to]
who-DAT what-ACC John-NOM [Takana-teacher-NOM read-CAUS that]
itta no?
said Q

'Who did John say professor Tanaka made read what?'

(3) *Examples of nesting paths* (YP$_2$ XP$_1$ ... t_1 t_2)

a. *Multiple topicalization in Mandarin* (Chen 2017, ex. 27a)

[**Shengfan**]$_2$, [**lubian xiao gou**]$_1$, wo wei guo t_1 t_2.
leftover-rice streetside small dog 1SG feed PERF

'Leftovers, I have fed them to small dogs on the street side'

b. *Topic/focus movement plus wh-movement in English*
[**This book**]$_2$, **who**$_1$ should we talk to t_1 about t_2?

The simplest way to achieve a configuration with overlapping A′-movement paths in English in a surface-evident way is to combine topic/focus fronting with *wh*-movement, as in (3b) above. Here we see that the landing site of the former movement is above that of the latter. We will see other overlapping movement configurations in English later in this paper, but for the moment, let's first consider what the derivational history of (3b) might look like.
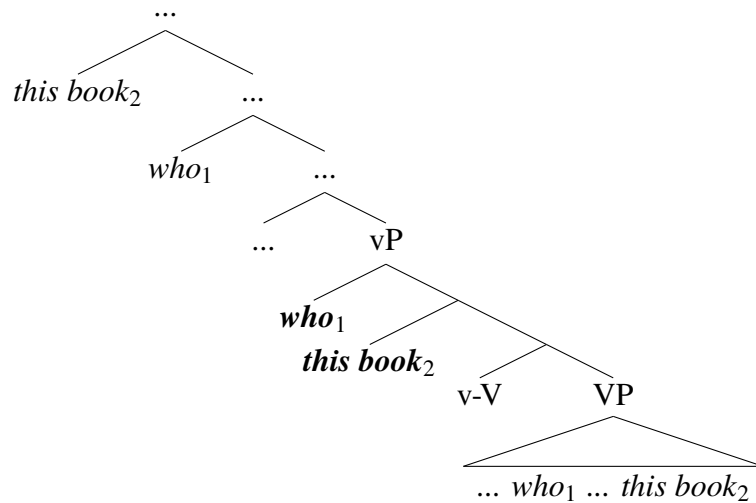
A central concern of this paper is *phase theory* (Chomsky 2000, 2001, 2008, and many following). This theory argues that (at least) vP and CP are *phases*, which phrasal movement from must pass through a specifier position of. As described in the next section, this prediction arises from a particular proposal concerning the nature of *spell-out*—an operation endemic to phases. Importantly, notice that if vP is a phase, then the two phrases undergoing A′-movement in (3b) above must both pass through specifier positions of the vP here. This consequence sits comfortably with the widely-held hypothesis that a phrase may have multiple specifiers (Chomsky 1995; Koizumi 1995; Ura 1996; Zwart 1997; Richards 1997, 1999; Doron and Heycock 1999; Chomsky 2000). However, since in (3b) above there are two moving phrases, there are two conceivable structures for the vP level of this derivation, illustrated in the simplified diagrams in (4) below: either the topic/focus phrase could form an outer spec-vP, with the *wh*-phrase forming an inner one (4a), or the reverse might occur (4b), prior to these phrases' further movement to the left periphery. In principle, either derivation would accomplish multiple movement from the vP as in (3b). However, here I will propose that the behavior of PGs in sentences like (3b) reveals that only the derivation in (4a) is correct, and argue for a theory that predicts this, among other results.

(4)   *Two plausible specifier orders for multiple A′-movement from vP in (3b)*

a.

this book$_2$ ... *who*$_1$ ... vP — **this book**$_2$ **who**$_1$ v-V VP ... *who*$_1$ ... *this book*$_2$

b.

this book$_2$ ... *who*$_1$ ... vP — **who**$_1$ **this book**$_2$ v-V VP ... *who*$_1$ ... *this book*$_2$
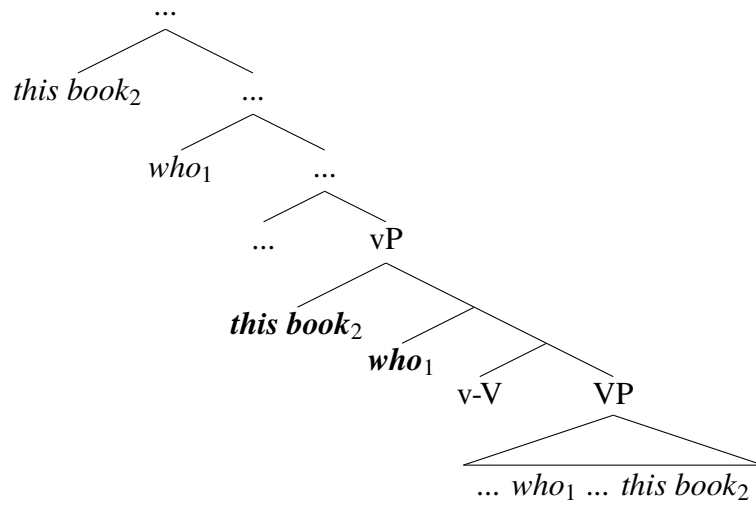
In particular, this paper advances a theory referred to as *Cyclic Linearization* (CL; Fox and Pesetsky 2005a,b; Ko 2007, 2011, 2014; Sabbagh 2007; Podobryaev 2009; Fanselow and Lenertová 2011; Jenks 2011, 2013a,b; Medeiros 2013; Overfelt 2015; Erlewine 2017; Davis 2019, 2020, a.o.), which attributes a great deal of explanatory power to the way in which phase-by-phase spell-out establishes word order, as described in section 2.

1.1   Main results of this paper

The first goal of this paper is to argue for the generalization in (5) below:

(5)   *Overlapping Chain Generalization (OCG)*
If two A′-moving phrases XP_1 and XP_2 form specifiers of an intermediate phase YP such that XP_1 c-commands and hence precedes XP_2, then XP_1 c-commands and hence precedes XP_2 in all subsequent positions these phrases occupy.

The derivation in (4a) above which I will argue is accurate for sentences like (3b) fits the OCG, since here the phrases *this book* and *who* have the same relative order both in their final landing sites, and in their intermediate landing sites in vP. In this paper, I argue that overlapping successive-cyclic movement is generally order-preserving in this way, and in doing so amass evidence for the CL theory, for which consistency of linear order is central.

The key evidence for the OCG comes from PGs. PGs have been used in much previous work to diagnose the properties of (A′-)movement (Engdahl 1983; Nissenbaum 1998, 2000; Legate 2003; Overfelt 2015; Branan 2017; Erlewine and Kotek 2018; Fox and Nissenbaum 2018, a.o.), and are central here as well, as a diagnostic for the derivational history of overlapping movement configurations. In particular, this paper will use a prediction from Nissenbaum (2000) about PG licensing in multiple movement scenarios, stated in (6) below:

(6)   *Multiple Specifier Single Parasitic Gap Prediction*
        In contexts where XP_1 is the highest specifier of vP and XP_2 is a low specifier of vP, only XP_1 can license a single parasitic gap in an adjunct to vP.[1]

The framework for PGs from which (6) arises is justified shortly. If the OCG is considered alongside this prediction, then we come to expect that when A′-movement chains overlap, only the moved phrase with the highest final landing site should be able to license a PG in an adjunct to a vP crossed by those movements. I argue that this is precisely what we find, as previewed in (7) below, which modifies (3b) to include a sentential adjunct with a PG:

(7)   *A PG is licensed only by the higher of two overlapped moved phases*
        [This book]$_2$, who$_1$ should we talk to t_1 about t_2...
        a.   ✓ ...before commenting on PG$_2$?
        b.   * ...before arranging a meeting with PG$_1$?

While this result clarifies the derivation of overlapping movement paths, it raises a question about multiple specifier formation. As described shortly, Richards (1997, 1999) argues that all but the first movement to a given head must "tuck-in" to a lower specifier of that head, yielding crossing paths. However, the derivation I attribute to configurations like (3b) and (7) involves nesting paths to the vP edge (4a), and thus a violation of Richards' hypothesis. This tension makes salient the question of how the distribution of crossing versus nesting multiple specifier formation is determined. Here I will propose that crossing paths are not always mandated by multiple movement to one head, but rather are only required for multiple movements to one head that are triggered by the same feature. In contrast, I propose that when a head bears multiple movement-triggering features, the goals respectively attracted by those features may form crossing or nesting paths.[2] I will argue that

---

[1]This statement of the prediction is adapted from Fox and Nissenbaum (2018), who consider it in the context of certain sentences involving pied-piping which are beyond this paper's scope.

[2]The concept that the presence of multiple movement triggering features on one head removes the need to tuck-in has precedent in several works (McGinnis 1998; Rackowski 2002; Doggett 2004; Rackowski and Richards 2005), many of which assume that tucking-in in fact cannot occur in this situation. Here I maintain that tucking-in is possible in this case, but not required.

this understanding accurately predicts the full range of possibilities for multiple specifiers.

A phenomenon closely related to the above concerns is the *path containment condition* (PCC, Pesetsky 1982), which describes the fact that in various languages, English among them, overlapping (overt) A′-movement generally requires the formation of nested dependencies. This is evident in English scenarios like the one in (3b) above, in which the original order of moved phrases must be reversed post-movement, as (8) shows:

(8)   *Forced order reversal for overlapping movement paths*
    a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2?   [2 1 ... 1 2]
    b.   * [This student]$_1$, what$_2$ should we talk to t_1 about t_2?   [1 2 ... 1 2]

I argue that such order reversal emerges from the considerations of CL when combined with the proposal that in languages like English, inner specifiers formed by A′-movement are sometimes covert (Richards 1997; Nissenbaum 2000, a.o.). I argue further that this account accurately predicts the possibility of PCC-violating configurations in the multiple *wh*-movement languages Bulgarian (as Richards 1997 observed) and Romanian, and fits into a more general theory of crossing versus nesting multiple specifier formation.
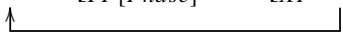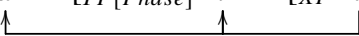
1.2   Contents of the paper

Next, section 2 introduces phase theory and CL. Section 3 explains the approach to multiple specifier formation in Richards (1997, 1999), its relationship to the theory of PGs in Nissenbaum (2000), and the prediction in (6) above. Section 4 shows that when (6) is considered in light of the PG asymmetry in contexts like (7), we find evidence for the OCG, which CL predicts. Section 5 argues for a CL-based account of the PCC, and is complemented by a revised account of multiple specifier formation in section 6. Section 7 discusses a correct prediction about overlapping movement and stranding, and section 8 shows that A-movement also obeys the OCG, as predicted. Section 9 concludes.

2   Phase theory and CL

A great deal of work has argued that movement paths are at least sometimes composed of multiple shorter paths connected by intermediate landing sites, in successive-cyclic fashion. See Chomsky (1973, 1977, 1986); Du Plessis (1977); Henry (1995); Cole and Hermon (2000); McCloskey (2000, 2001, 2002); Legate (2003); Sauerland (2003); Bruening (2001, 2006); Barbiers (2002); Torrence (2012); Abels (2003, 2012); Wiland (2010); Henry (2012); van Urk (2015); van Urk and Richards (2015); Korsah and Murphy (2019); Davis (2019, 2020), and many others. Chomsky (2000, 2001) and much following work argues that movement must pass successive-cyclically through the edges of phases (vP, CP) in order to escape spell-out—an operation endemic to phases. For this theory, spell-out applies to the complement of phase heads, transferring the content of the complement to the interfaces of PF and LF, which respectively assign linear ordering to and interpret that content, among

other processes. By hypothesis, spelled-out structure is inaccessible for the rest of the syntactic derivation, as defined by the Phase Impenetrability Condition. Thus, for this theory, a phrase moving from a phase's complement must reach the phase edge before spell-out, to avoid being prematurely rendered immobile by transfer to the interfaces (9):

(9) *Must exit phase complement via the phase edge*

    a. $*$ $[_{ZP}\ \alpha\ Z\ [_{YP[Phase]}\ Y\ [_{XP}\ t\ ]]]$

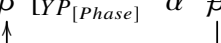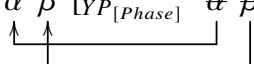   b. $\checkmark$ $[_{ZP}\ \alpha\ Z\ [_{YP[Phase]}\ t\ Y\ [_{XP}\ t\ ]]]$

In contrast, for the CL theory spell-out applies to entire phasal constituents, including their edge, as soon as they are done being built up by successive applications of Merge. This hypothesis necessitates the proposal that spelled-out material remains accessible for the rest of the derivation, since in this system all movement from a phase is necessarily of material that has undergone spell-out within that phase. As such, for CL, successive-cyclic movement through phase edges does not occur because edges are exempt from phase-level spell-out. Rather, under this approach successive-cyclic movement through phase edges is motivated by the information-preserving property of spell-out, *Order Preservation*:

(10)   *Order Preservation*   (Fox and Pesetsky, 2005a, p. 6)
           Information about linearization, once established at the end of a given Spell-Out domain, is never deleted in the course of a derivation.

If Order Preservation holds, it is not possible to revise established ordering information to save derivations for which phase-by-phase spell-out has generated contradictory linearization instructions. Therefore, to avoid a crash at PF, syntax must be able to form a structure with linearization information that is consistent for all phases in the derivation in question. As the works on CL cited above argue, exiting a phase via its linear edge serves to prevent that movement from yielding a linearization contradiction. By passing through the linear edge of each phase exited, phase-exiting phrases are determined by spell-out to precede the content of each phase in question. This is ultimately consistent with a final representation where the moved material surfaces preceding the content of all phases that it has exited.

    CL also predicts that there is a way to avoid a crash at PF if movement from a phase doesn't pass through the linear edge, hence crossing over some material in the phase on the way out. Moving that crossed-over material into the next phase to a position preceding what previously crossed it, thus restoring the original order of those elements, keeps linearization coherent. For instance, (11a) below is illicit if it remains as-is, due to α non-successive-cyclically crossing over β on the way out of the phase YP. However, the derivation won't fail if, as (11b) shows, β later moves to precede α in the next phase XP as it did in YP:
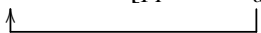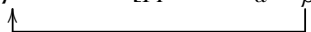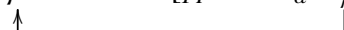
(11)   a.   *Illicit phase exit...*   b.   *...repaired by additional movement*

           $*$ $[_{XP[Phase]}\ \beta\ [_{YP[Phase]}\ \alpha\ \beta\ ]]$ $\checkmark$ $[_{XP[Phase]}\ \alpha\ \beta\ [_{YP[Phase]}\ \alpha\ \beta\ ]]$

In brief, as the works on CL cited in section 1 argue, cross-phasal movement is not constrained by a strict notion of phase impenetrability. Rather, because the linear order established for the content of each phase upon its completion by spell-out (which applies to the entire phase) cannot be contradicted later on, all cross-phasal movement must be order-preserving. In this paper, I argue that the behavior of overlapping movement reveals further evidence for this view of phases, which here manifests as the OCG, and simultaneously leads to consequences for various other topics in the study of syntactic movement.

3   Multiple specifiers, parasitic gaps, and their interaction

Richards (1997, 1999) argues that when two phrases are attracted to one head, the closest phrase is attracted first (12a), while the lower one which is attracted second "tucks-in" to a specifier below the one formed by the preceding movement (12b). The non-initial movement must tuck-in rather than move to a position above the phrase that moved first (12c) because of an economy constraint requiring movement paths to be as short as possible.

(12)   *Non-initial instances of movement to one head must tuck-in*

    a. $[_{XP}\ \alpha\ \text{X}\ ...\ [_{YP}\ ...\ t_\alpha\ \beta\ ]]$

    b. $\checkmark\ [_{XP}\ \alpha\ \beta\ \text{X}\ ...\ [_{YP}\ ...\ t_\alpha\ t_\beta\ ]]$

    c. $*\ [_{XP}\ \beta\ \alpha\ \text{X}\ ...\ [_{YP}\ ...\ t_\alpha\ t_\beta\ ]]$

Richards argues that tucking-in is responsible for sentences like those in (2) above, where multiple phrases move to adjacent positions in order-preserving fashion. For Richards, in these sentences the final order of moved phrases reflects the order in which they were attracted, since the lower of the two moved phrases, which moves second, must land below the one that moved first. Richards' theory accurately predicts that the sentences in (2) are unacceptable if the moved phrases reverse their order, as he shows to be the case.

Importantly, in this paper we will see evidence that a tucking-in strategy is not always required for multiple movements to one position. Thus, as previewed above, I will argue for an alternative theory of multiple specifier formation. This evidence will come from the behavior of PGs in a variety of English sentences like (3b) above, repeated in (13) below, in which multiple phrases overtly A′-move in overlapping fashion. As discussed above, the PCC applies to such sentences. While the PCC will be addressed later in this paper, for the meantime it will be useful to set aside the fact that the PCC holds and simply consider what the theory of Richards (1997, 1999) predicts about how such a sentence was derived.

(13)   *Topic/focus movement + wh-movement*
    [This book]$_2$, who$_1$ should we talk to t_1 about t_2?

Under Richards' hypothesis, the derivation of the vP in (13) should proceed as follows.

First, the originally higher A′-moved phrase who_1 should form an initial A′-specifier of vP:[3]

(14) *Originally higher moved phrase forms higher spec-vP*
$[_{vP}\ who_1\ \text{S}\ \text{v-V}\ [_{VP}\ t_1\ this\ book_2\ ]]$

Second, the initially lower phrase *this book*$_2$ should tuck-in to a spec-vP below who_1:

(15) *Originally lower moved phrase forms lower spec-vP*
$[_{vP}\ who_1\ this\ book_2\ \text{S}\ \text{v-V}\ [_{VP}\ t_1\ t_2\ ]]$

From their intermediate positions in the edge of the vP phase, these phrases would then continue on to their respective final landing sites in the clause's periphery:

(16) *Movement to left periphery after tucking-in at vP*

$[\ this\ book_2\ who_1\ \text{S}\ \text{T}\ [_{vP}\ who_1\ this\ book_2\ t_S\ \text{v-V}\ [_{VP}\ t_1\ t_2\ ]]]$

In contrast, this theory excludes a derivation in which the initially lower A′-moving phrase forms a specifier of vP above that of the phrase that moved first, resulting in these phrases' paths nesting within vP prior to their later movement onward:

(17) *Multiple A′-movement from vP without tucking-in (=nesting paths in vP)*

$[\ this\ book_2\ who_1\ \text{S}\ \text{T}\ [_{vP}\ this\ book_2\ who_1\ t_S\ \text{v-V}\ [_{VP}\ t_1\ t_2\ ]]]$

Building from the prediction about PGs in (6) above, I will argue that this derivation with nesting paths in vP is in fact correct. Notice that this derivation fits the OCG, which emerges directly from the order preservation requirement of CL: here *this book* precedes *who* both in these phrases' final landing sites, and in their intermediate vP landing sites.

Next, I will overview the components of Nissenbaum (2000) that justify the prediction in (6) and thus allow the evidence for the OCG to become apparent.

3.1 Parasitic gaps, successive-cyclicity, and specifier order

A PG is a gap which depends on A′-movement structurally crossing over the constituent that contains it, among other requirements (Culicover and Postal 2001). That a given gap is a PG is clearest when in an island, since in this case it is clear that the gap was not derived

---

[3]I diagram A′-movement of non-subjects as passing through a position in vP above the subject in situ, as in (14-17). This order of constituents has independently been assumed by a variety of works which take tucking-in below thematic specifiers to be banned or dis-preferred (McGinnis 1998; Nissenbaum 2000; Chomsky 2001; Rackowski 2002; Rackowski and Richards 2005), and is also predicted by CL, since a non-subject A′-moving from vP must precede the subject at the spell-out of vP just as it will in its final landing site in CP.

via typical extraction. For instance, in (18) below we see PGs in sentential adjuncts licensed by *wh*-movement external to those adjuncts, which are generally islands.[4]

(18)   *PGs in adjunct islands*
    a.   Who$_1$ did you forget about t_1 [because I didn't mention PG$_1$]?
    b.   Tell me what$_1$ I should reread t_1 [before giving you comments on PG$_1$].

Following Nissenbaum (2000), PGs in sentential adjuncts will be the focus of this paper, since these are the most straightforward PG-hosting constituents that interact with movement within the clause in the relevant way. PGs in subjects will be briefly addressed as well.

PG-licensing has several requirements. For instance, the PG-containing island must be adjoined to a position structurally crossed by the licensing movement chain (Kayne 1983; Longobardi 1984, a.o.) Additionally, this property must hold of the smallest island containing the PG. Separating the PG-containing island from the licensing movement chain by including a second island results in unacceptability (Kayne 1983; Chomsky 1986, a.o.):

(19)   *PG licensing across multiple islands fails*
    a.   *Relative clause island in adjunct island*
        * Tell me who$_1$ you talked to t_1 [after meeting a person [who likes PG$_1$]].
    b.   *Subject island in adjunct island*
        * Durian is a fruit [which$_1$ I ate t_1 for the first time [after [a fan of PG$_1$] visited me]].
    c.   *Adjunct island in adjunct island*
        * Guess who$_1$ I ironically ran into t_1 [after taking the other hallway [because I wanted to avoid PG$_1$]].

Nissenbaum (2000) argues that these and further facts are captured by a theory in which PGs are formed by movement of a null operator (OP) to the edge of the containing island (Chomsky 1986, a.o.),[5] as in (20) below, in combination with the proposal that for semantic reasons, the island adjoins to a landing site of (successive cyclic) A′-movement:

(20)   *PG as trace of null operator*
    Who$_1$ did you forget about t_1 [OP after talking to t_{OP}(=PG)]?

Nissenbaum argues that a sentence like (20) is formed and interpreted in the following way. Within the sentential adjunct island, operator movement to its edge triggers the semantic rule of Predicate Abstraction (Heim and Kratzer 1998), which changes the island

---

[4]We can also tell that the PGs in (18) are not actually omitted objects, since *mention* in (18a) is obligatorily transitive, and the preposition *on* in (18b), like all English prepositions, is as well:

(i)   a.   * Did you mention ___ yesterday?   b.   * Please give me comments on ___ soon.

[5]This is in contrast to "shared antecedent" theories of PGs, for which a null operator is not involved. See Nissenbaum (2000) and Nissenbaum and Schwarz (2011) for arguments against this theory.

into a predicate of type <e,t>, as in (21).[6]

(21)   *Null operator movement inside adjunct forms a derived predicate*

```
              ...
           ╱    ╲
    ...    AdjunctP
           <e,t>
        ╱         ╲
  ─────────────────────────
  OP₁ λ₁ after talking to t_OP₁(=PG)
```

OP_1 λ_1 after talking to t_{OP_1}(=PG)
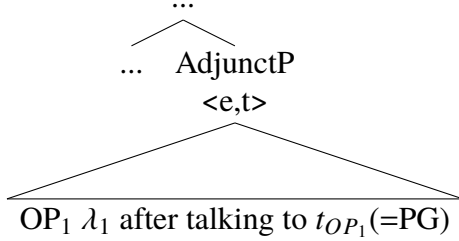
Additionally, PG licensing in (20) requires successive cyclic movement of the licensing phrase *who* through the edge of vP, which the phasehood of vP independently necessitates. This intermediate step of A′-movement triggers an application of Predicate Abstraction there as well, creating an <e,t> node in the vP edge. This is diagrammed in the partial structure in (22) below. Here we see that the resulting <e,t> node is sister to the type e trace left behind in the vP edge by successive cyclic movement of *who* from this domain:

(22)   *Successive-cyclic A′-movement creates an <e,t> node in vP*

```
            vP
            t
          ╱   ╲
      t_wh₂    <e,t>
        e    ╱      ╲
        ──────────────────
        λ₂ you forgot about t_wh₂
```

t_{wh_2} λ_2 you forgot about t_{wh_2}

The PG-containing adjunct in (21) above is type <e,t>, and as we've seen in (22), there is a node of the same type in the vP edge after successive-cyclic A′-movement from it. Thus this adjunct can be merged to this derived <e,t> position in the vP, yielding a structure that is successfully interpreted via Predicate Modification (Heim and Kratzer 1998), as in (23):
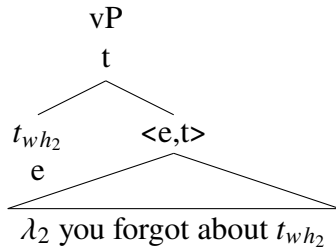
---

[6]I adopt from Nissenbaum the simplifying assumption that vPs and vP modifiers (such as the sentential adjuncts under consideration) are type t, modulo A′-movement within them triggering Predicate Abstraction. As Nissenbaum (2000, p. 47) notes, this is a simplification because it ignores the presence of temporal and event arguments, but enriching the semantic type of these constituents does not make any important difference for the account of PGs. In particular, given a more general version of the rule of Predicate Modification that allows constituents of the same semantic type to be combined (which is independently required to handle adjuncts containing multiple PGs, as we'll see shortly), the account described here functions the same regardless of whether the type of vPs and their adjuncts is t, or something more complex.

(23) *Predicate Modification of vP with PG-containing island*

vP
t

t_{wh_2} | <e,t> |
e

<e,t>   AdjunctP
                         <e,t>

λ_2 you forget about t_{wh_2}

OP$_1$ λ_1 after talking to t_{OP_1}(=PG)
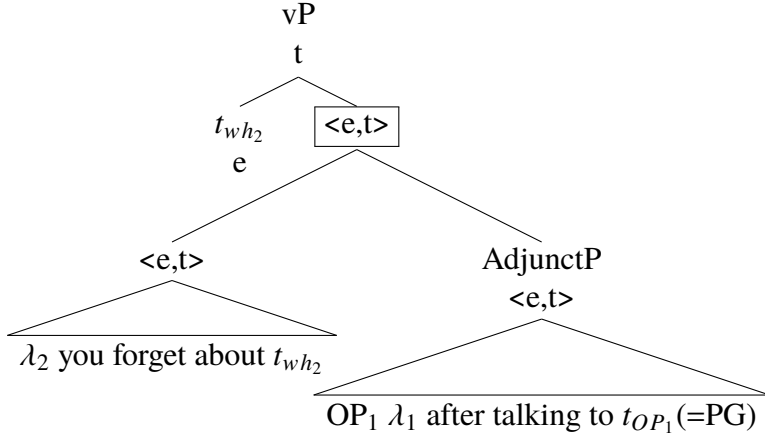
Nissenbaum provides arguments from phenomena including binding, NPI licensing, VP ellipsis, and VP fronting to establish that the PG-containing adjunct in such contexts is indeed adjoined at a clause-medial position, which he takes to be the vP edge, as (23) shows. In this structure, the (boxed) intermediate vP node created by merge of the adjunct to the site of successive-cyclic movement in vP denotes a function of type <e,t>, which is true of individuals that the addressee both talked to and forgot about. The intermediate type e trace of the A′-moved nominal phrase will saturate the individual argument of this function, "filling in" both the variable that corresponds to its trace in VP, and the trace of the null operator in the adjunct, resulting in a PG that is successfully interpreted.

This theory of PGs captures the fact that a PG can only be licensed if the smallest island that contains it is adjoined within the licenser's movement path. It is semantically necessary for the PG-forming operator to reach the edge of the island that will adjoin within the licenser's movement path, but if an embedded island intervenes, such movement fails. Furthermore, if the PG-containing constituent is semantically a predicate that must undergo Predicate Modification with a landing site of the licensing phrase, then it is straightforward why a PG cannot be interpreted in a position outside of the licenser's movement path.

Next we will see how Nissenbaum relates this theory to contexts with multiple specifiers of vP, and thus makes the prediction that this paper's results depend on.

3.1.1 Multiple specifiers and PG licensing

Nissenbaum's examination of PGs in multiple specifier structures focuses on multiple *wh*-questions. In many Slavic languages like Bulgarian, all *wh*-phrases in such contexts move to the interrogative CP, with lower *wh*-phrases landing in a position below the initially highest one, which Richards (1997, 1999) attributes to tucking-in at CP (24a). In English, such questions involve overt movement only of whichever *wh*-phrase is initially highest (24b), modulo D-linking. Nissenbaum follows Richards in arguing that in both types of languages, all *wh*-phrases move in multiple *wh*-questions, the difference being that in a language like English the movement responsible for forming tucked-in specifiers is covert:

(24)   a.   *Bulgarian multiple wh-question* ([Richards 1997](#) p. 277 ex. 75a)

Kogo$_1$ **kakvo**$_2$ e   pital   Ivan t_1 t_2?
who.NOM what.ACC AUX asked Ivan

   b.   *English multiple wh-question (with covert tucking-in of lower wh-phrase)*
Who$_1$ $\varnothing_2$ did Ivan ask t_1 **what**$_2$?

If seemingly in situ *wh*-phrases truly do move, we would expect them to be capable of PG licensing. As [Engdahl](#) ([1983](#)) observed, it appears that they are not:

(25)   *Wh-in-situ cannot license a PG* (Engdahl 1983, ex. 34)
        * I forget who filed [**which articles**]$_1$ [without reading **PG**$_1$].

However, Nissenbaum observed that when overt *wh*-movement licenses the first of two PGs in one island, an in situ *wh*-phrase can license the second PG:

(26)   *Wh-in-situ licenses second PG in the same island* (Nissenbaum 2000, p. 12 ex. 8)
        [Which senator]$_1$ did you persuade t_1 to borrow [**which car**]$_2$ [after getting an opponent of PG$_1$ to put a bomb in **PG**$_2$]?

Thus Nissenbaum argues that English in situ *wh*-phrases do move. Only the licensing of the second of two PGs allows such movement to become apparent in English, Nissenbaum argues, because of the semantic effect of multiple specifiers of A′-movement. Next I summarize Nissenbaum's theory of the interpretation of multiple specifiers, which will lead us directly to the diagnostic prediction that this paper will make central use of.

Nissenbaum argues that when multiple phrases A′-move to one head, the lowest saturated segment of that head's projection undergoes one instance of Predicate Abstraction for each moved phrase, as in (27) below. Thus when vP is targeted by two intermediate A′-movements in a multiple *wh*-question, the vP segment that was the root before movement (labeled β below) is interpreted as a two-place predicate <e,<e,t>>. In this context, the trace in the lower specifier of A′-movement saturates the first argument of this two-place predicate, yielding a one place predicate (labeled α below) which is in turn saturated by the higher specifier's trace of A′-movement, resulting in a type t for the root node of the vP:[7]

---

[7]Since the trace of movement through the inner spec-vP (t_{wh2}) in (27) is the sister of β, the two semantic argument positions of the function denoted by β must have an order that is the opposite of the two specifiers of vP that saturate it (t_{wh1} t_{wh2} ... λ_2 λ_1). This ensures that when t_{wh2} combines with β, it correctly binds its corresponding trace in VP, rather than that of the outer specifier. After this occurs the result is α, a one-place predicate whose saturation by the trace of the outer spec-vP (t_{wh1}) binds the corresponding trace in VP as well, resulting in the correct interpretation.

(27) *vP after two successive-cyclic movements*

vP
t

t_{wh_1} α
e   <e,t>

   t_{wh_2} β
   e   <e,<e,t>>

   λ_2 λ_1 (S) v-V t_{wh_1} t_{wh_2}
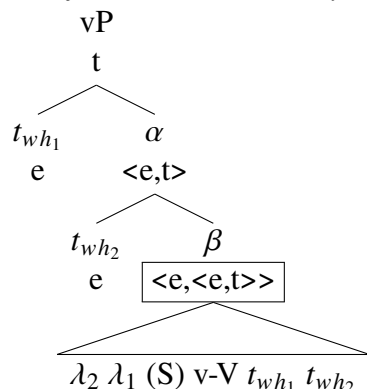
If an adjunct with one PG is type <e,t> as shown in (21) above, then such an adjunct can only combine via Predicate Modification with the node α in the multiple specifier structure in (27). This results in the configuration in (28) below. The node in (28) produced by merger of the PG-containing adjunct to α is a predicate that is saturated by the intermediate trace in the higher specifier of A′-movement (t_{wh_1}).

(28) *Licensing of single PG by (trace of) A′-movement through higher spec-vP*

vP
t

t_{wh_1} <e,t>
e

   α AdjunctP
   <e,t>   <e,t>

   t_{wh_2} β OP$_3$ λ_3 ... t_{OP_3}
   e   <e,<e,t>>

   λ_2 λ_1 (S) v-V t_{wh_1} t_{wh_2}
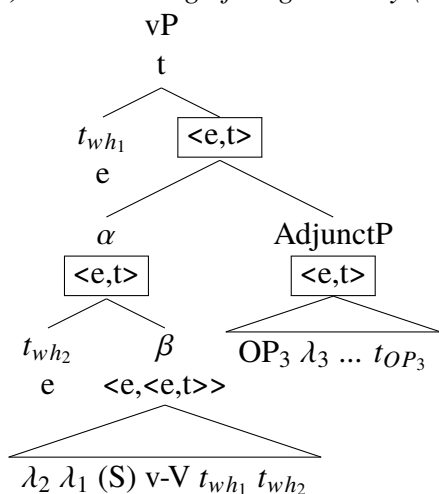
Consequently, in (28) the A′-moving phrase that formed the outer spec-vP licenses the PG. This derivation appropriately models a sentence where the initial *wh*-movement in an English multiple *wh*-question, which is overt, successfully licenses a PG:

(29) *Overt wh-movement in multiple wh-question can license PG*
      Who$_1$ ∅$_2$ did Bill ask t_1 about **what**$_2$ [in order to have a reason to talk to **PG**$_1$]?

In contrast, as we saw in (25) above, an in situ *wh*-phrase in a multiple *wh*-question cannot license a lone PG. If the in situ phrase covertly moves to an inner specifier of the vP (and later on the CP), what has been said above accurately predicts this restriction. In order for the inner specifier of A′-movement in (28) above to license the PG, the PG-containing

13

adjunct would have to merge with the node β, which is the sister of t_{wh_2}. If this were possible, t_{wh_2} would saturate the function output by Predicate Modification of the adjunct with β, and thus license the PG. However, in reality such a structure is not interpretable, as we see in (30) below. Since β is a segment of vP that is a two-place predicate due to the occurrence of two A′-movements, the one-place predicate instantiated by an adjunct with one PG cannot combine with β either by Predicate Modification, or by Functional Application. Thus the inner specifier of A′-movement cannot license the PG in this case:

(30)   *Type mismatch prevents licensing of single PG by lower spec-vP*
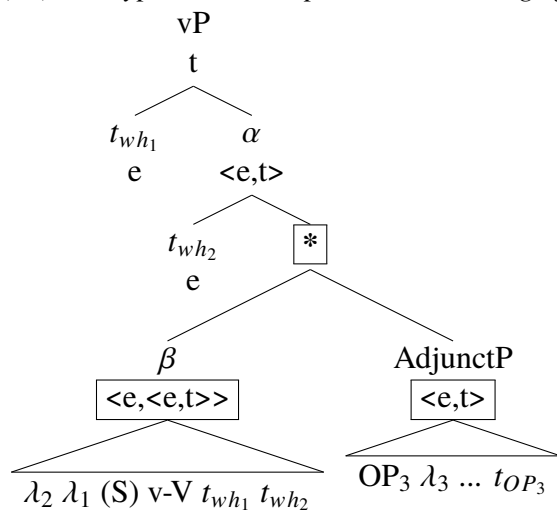
These concepts lead us to the correct prediction that when overt *wh*-movement licenses the first of two PGs in one island, a covertly moving *wh*-phrase can license the second PG, as (26) above showed. Assuming that Predicate Modification can apply to any two phrases of the same semantic type (in the spirit of Partee and Rooth 1983, see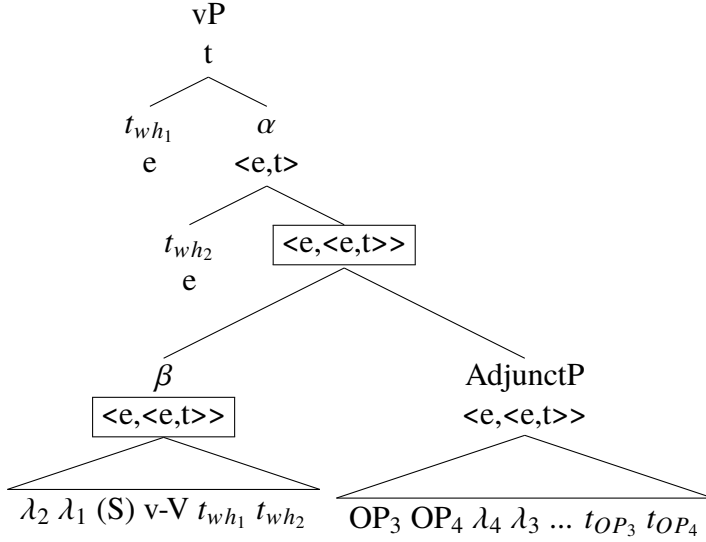 also Nissenbaum and Schwarz 2011), the (trace of the) inner specifier of A′-movement will be able to license a PG in a situation where the PG-containing adjunct is a two-place predicate, and thus can be interpreted by merging with β in (27). This is achieved if the adjunct contains two PG-forming operators, which each move to the edge of the adjunct and each trigger an instance of Predicate Abstraction, with the result that the adjunct is type <e,<e,t>> as needed, as in (31) below. Merge of this two PG adjunct to β in (31) allows both specifiers of A′-movement to license one PG each, as we have seen is possible in reality:[8]

---

[8]The concerns of footnote 7 extend to the operator movements in the adjunct in (31). If these operators must form crossing paths (as Nissenbaum proposes) and apply Predicate Abstraction in reverse order as footnote 7 describes, we predict Nissenbaum's observation that in a multi-PG adjunct like that in (26), the order of gaps and licensers must match, as we see by comparing (26) with (i). Since both the *wh*-phrases and operators in (26/31) form crossing paths and trigger Predicate Abstraction in the way just described, the first *wh*-phrase must license the first PG, while the second *wh*-phrase must license the second PG.

(i)   * [Which car]$_1$ did you lend t_1 to [which senator]$_2$ [after getting an opponent of PG$_2$ to put a bomb in PG$_1$]? (Nissenbaum 2000, p. 110 ex. 27a)

(31) *vP after two successive-cyclic movements with a two PG adjunct*

vP
t

t_{wh_1} — e

α — <e,t>

t_{wh_2} — e

<e,<e,t>>

β — <e,<e,t>>

AdjunctP — <e,<e,t>>

$\lambda_2 \lambda_1$ (S) v-V t_{wh_1} t_{wh_2}

OP$_3$ OP$_4$ $\lambda_4 \lambda_3$ ... t_{OP_3} t_{OP_4}

In summary, if covert *wh*-movement in English is a product of tucking-in to a lower specifier, the account described above predicts the fact that such movement cannot license a lone PG, but can license the second of two PGs contained in one island. Abstracting away from the fact that inner specifier formation happens to be covert in English multiple *wh*-questions, this understanding predicts that, in any multiple specifier configuration, the lower of two specifiers of A′-movement will be unable to license a lone PG. Consequently in any structure like (32) below in which α and β are both specifiers of successive-cyclic A′-movement from vP, we expect a single PG to be able to be licensed only by the higher specifier (here α), regardless of the phonological status of the movements involved:

(32) *Only the outermost A′-specifier of vP can license a lone PG*

vP

α_1
✓

β_2
X

... PG$_1$

v

VP

... *PG$_2$

V ... t_1 ... t_2 ...
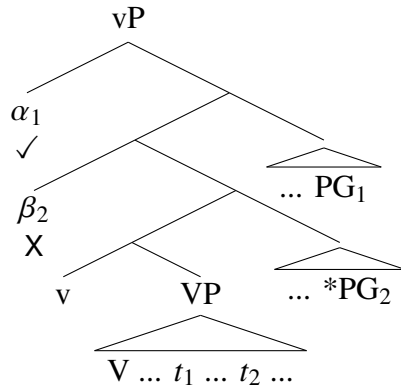
This is as described by the prediction previewed earlier in (6) above, repeated in (33) below:

(33) *Multiple Specifier Single Parasitic Gap Prediction*
In contexts where XP$_1$ is the highest specifier of vP and XP$_2$ is a low specifier of vP, only XP$_1$ can license a single parasitic gap in an adjunct to vP.

With this prediction about the relationship between PGs and multiple specifiers now justified, next I show how these concerns reveal that overlapping moved phrases have the same relative order in both their intermediate and final positions, as described by the OCG.

4 PGs reveal that final and intermediate orders match

To begin taking stock of the relevant facts, first consider once more an English configuration where topic/focus fronting and *wh*-movement are combined, such as (34a) below. In such examples the PCC holds (34b) and topic/focus movement must land above *wh*-movement as mentioned earlier, even if the PCC is obeyed (34c), but set these constraints aside for the moment in order to focus on how PG licensing behaves in such configurations.

(34)   *Topic/focus movement plus wh-movement in English*
   a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2?
   b.   * [This student]$_1$, what$_2$ should we talk to t_1 about t_2?
   c.   * What$_1$ [this person]$_2$ should we talk to t_2 about t_1?

We have seen in numerous preceding examples that *wh*-movement can license a PG. This is also possible for topic/focus movement, since PG licensing is available for all A′-movements, regardless of sub-type (Culicover and Postal 2001):

(35)   *Topic/focus movement can license a PG*
   a.   [This book]$_2$, we should talk about t_2 [before buying PG$_2$ for our class].
   b.   [These papers]$_2$, I asked the TA to grade t_2 [after making copies of PG$_2$].

Importantly, as previewed in section 1, combining these two A′-movements yields a restriction on PG licensing. This is shown again in (36) below. We see here that only the A′-moved phrase with the higher final landing site can license a single PG, even if the sentence would be semantically and pragmatically well-formed if the inner one were the licenser:

(36)   [This book]$_2$, who$_1$ should we talk to t_1 about t_2...
   a.   ✓ ...before commenting on PG$_2$?
   b.   *...before arranging a meeting with PG$_1$?

Given the prediction in (6/33) above, the fact that the outer moved phrase can license the PG here, but the inner one cannot, indicates that the latter formed an inner specifier of the vP when these two phrases underwent successive-cyclic movement from it. This is precisely what CL predicts: for CL, the relative order of these two moving phrases must be the same both in their final landing sites in the left periphery, and in their intermediate positions in the vP edge. The PG licensing asymmetry in (36) indicates that this is so: the inner of the two moved phrases behaves like the inner specifier of the vP.[9]

---

[9]Above, we saw evidence from Nissenbaum (2000) that, contra Engdahl (1983), covert *wh*-movement

While combining topic/focus movement and *wh*-movement in one clause is not the only way to yield overlapping overt A′-chains in English, it is the simplest one. The alternative is to use a multi-clausal derivation, which contains multiple CP edges that can be targeted for overt A′-movement. As Pesetsky (1982) notes, while such sentences have been taken to be ungrammatical by some works, many speakers do accept them provided that the PCC (which I address later on) is obeyed. Some such sentences are shown in (37-39) below, each of which includes a sentential adjunct with one PG. We see here that, as in (36), only the outermost of the two overtly A′-moved phrases can license the PG:

(37) a. [What car]$_2$ do you know [who$_1$ to persuade t_1 to buy t_2, [in order to get to borrow PG$_2$ for free]]?
      b. *[What car]$_2$ do you know [who$_1$ to persuade t_1 to buy t_2 [after having a few drinks with PG$_1$]]?

(38) a. Tell me [which paper]$_2$ she said [[which professor]$_1$ I should talk to t_1 about t_2 [before submitting PG$_2$ for publication]]?
      b. *Tell me [which paper]$_2$ she said [[which professor]$_1$ I should talk to t_1 about t_2 [after introducing myself to PG$_1$]]?

(39) a. This is the book [which$_2$ I asked [who$_1$ you gave t_1 a summary of t_2 [after writing a review of PG$_2$]].
      b. *This is the book [which$_2$ I asked [who$_1$ you gave t_1 a summary of t_2 [before having a meeting with PG$_1$]].

This result indicates that the higher moved phrase on the surface in such sentences was also the higher spec-vP, exactly as CL predicts, and as the OCG describes.

The above examples have shown contexts where the two phrases competing to license the PG originate in the same clause. The same asymmetry in PG licensing holds even if the two moved phrases originate in different clauses, however, as shown by (40-42) below:

(40) a. What$_2$ did you mention [who$_1$ you told t_1 [that you eat t_2 every day] [while buying PG$_2$ at the store]]?
      b. *What$_2$ did you mention [who$_1$ you told t_1 [that you eat t_2 every day] [while talking to PG$_1$ at the store]]?

(41) a. Those are the beans [which$_2$ I know [[what students]$_1$ you informed t_1 [that you'll never buy t_2 again] [after eating an expired can of PG$_2$]]].
      b. *Those are the beans [which$_2$ I know [[what students]$_1$ you informed t_1 [that you'll never buy t_2 again] [after running into PG$_1$ in the hallway]]].

can license PGs: it simply fails to do so when there is only one PG, because such movement forms an inner specifier. Conversely, in (36), we see overt *wh*-movement which fails to license a lone PG when combined with topic/focus movement, despite overt *wh*-movement being able to license PGs in isolation. This result indicates that the overt/covert distinction in A′-movement does not have anything to do with PG licensing: what really matters is the structure involved, and the mechanisms available for interpreting it.

(42)   a.   [That book]$_2$, who$_1$ were you loudly telling t_1 [that you refuse to buy t_2] [while burning a copy of PG$_2$ yesterday]?

   b.   *[That book]$_2$, who$_1$ were you loudly telling t_1 [that you refuse to buy t_2] [while eating lunch with PG$_1$ yesterday]?

Given CL, in such examples we predict that the two A′-moving phrases will adopt an order matching what their final relative order will be as soon as the initially lower phrase moves into the vP in which the initially higher phrase originates. This is because once the two moving phrases occupy the same phase, they will be linearized with respect to each other, thus at that time and no later it is necessary for them to adopt whatever their final relative order will be. A simplified derivation illustrating this prediction is shown in (43):

(43)   *Phrases from separate clauses match final order once in same phase [=(42)]*[10]
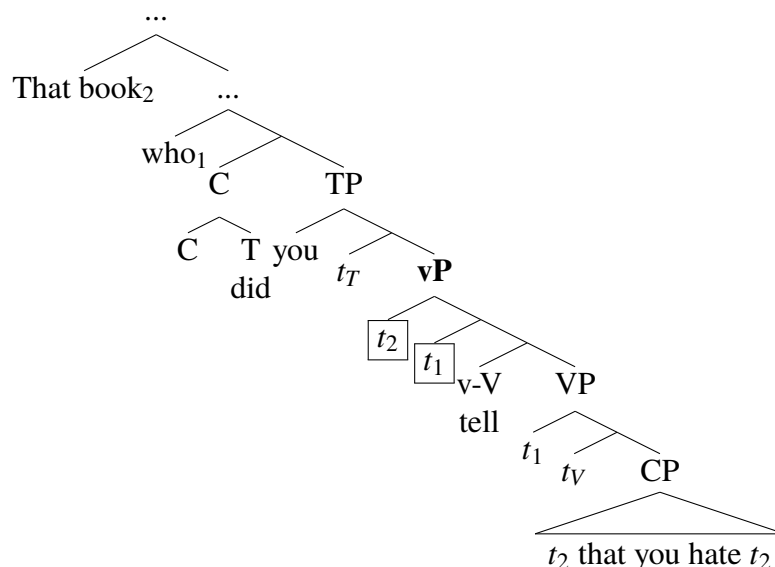
The fact that only the higher moved phrase on the surface in such configurations can license a lone PG, as (40-42) showed, indicates that this analysis is correct.

4.1   Confirming that the lower moved phrase forms an inner spec-vP

We saw above that while covert *wh*-movement in English cannot license a PG which is alone (25), such movement can license the second of two PGs contained in one island (26). As discussed, Nissenbaum argues that this fact emerges from the semantic nature of multiple specifier formation, which only allows an inner specifier to license the second of two PGs contained in one island. In (36-42) above, I provided an array of configurations with overlapping overt A′-movement, in which the moved phrase with the lower surface position

---

[10]I represent the indirect object in this sentence as originating in a specifier of the VP, since this accurately captures the word order of this construction, and because more fine-grained approaches to the structure of ditransitives would not contribute anything relevant to the analysis at hand.

is unable to license a lone PG. If as argued above the final relative order of these moved phrases reflects their intermediate relative order within vP, then the inner moved phrase in these examples should be an inner specifier of vP, which will reveal its PG licensing ability when an adjunct with two PGs is merged. This is indeed the case:

(44)   *Lower of two overtly A′-moved phrases can license a second PG*

    a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2, [before giving comments on PG$_2$ to a student of PG$_1$]?

    b.   This is the book which$_2$ I wondered who$_1$ we should send t_1 comments on t_2, [after giving a copy of PG$_2$ to a friend of PG$_1$].

    c.   Guess [which cookies]$_2$ I discovered who$_1$ you told t_1 that you are addicted to t_2, [while buying a package of PG$_2$ for a friend of PG$_1$]?

This fact confirms that these constructions really involve a multiple specifier configuration at the vP level, as opposed to separate movements through distinct projections within the verbal domain. If the later had occurred, there would not be a two-place predicate in vP that would facilitate the interpretation of a two-PG adjunct.

4.2   Supporting evidence from binding

Under CL, we expect the final relative order for overlapping moved phrases to have already been formed in the first phase that those phrases both occupied, even if many phases intervene between their origination points and final landing sites. This prediction is corroborated by (45) below, in which a PG-containing adjunct includes an anaphor bound by the subject of the embedded clause, ensuring that this adjunct is interpreted in the lower vP—the first phase of this multi-clausal derivation. Notice that the PG here can only be licensed by what is ultimately the phrase with the higher surface position, not the lower one. This shows that the relative c-command relation between these moved phrases that we see in their derived positions was indeed already established within the first phase of the derivation:

(45)   *Ordering established within the first vP as diagnosed by principle A*
    [**What book**]$_2$ did she wonder **who**$_1$ <u>you</u> should [$_{vP}$ t_2 t_1 ask t_1 about t_2 ...

    a.   [✓before making <u>yourself</u> a copy of PG$_2$]]?

    b.   [*after introducing <u>yourself</u> to PG$_1$]]?

    Example (46) below uses principle B to show the same result. Here merge of the PG-containing adjunct in the matrix clause is excluded by principle B, forcing it to attach in the embedded clause, and yielding the same asymmetry in PG licensing that we saw in (45).[11]

---

[11]Examples (45-46) rely on binding of a non-subject argument in the adjunct by the subject of the containing clause (presumably mediated by a co-indexed PRO in the adjunct), for which it is independently verifiable that principles A and B apply, in clause-bounded fashion as usual:

(i)   a.   You$_1$ went to the park [before taking pictures of yourself$_1$/*you$_1$].

(46) *Ordering established within the first vP as diagnosed by principle B*
[**What book**]$_2$ did you wonder **who**$_1$ we should [$_{vP}$ t_2 t_1 ask t_1 about t_2 ...

    a.   ✓ [before making you a copy of PG$_2$]]?

    b.   * [after introducing you to PG$_1$]]?

4.3 On PGs in subjects

So far this paper has focused on PGs in sentential adjuncts, which are straightforward for semantic composition. However, the important PG licensing asymmetry also holds for PGs in subjects. This fact was observed by Pesetsky (1982) by way of the following contrast:

(47) This volvo is one car which$_2$ I know who$_1$ to persuade [[owners of PG$_2$] / [friends of *PG$_1$] to talk to t_1 about t_2]. (Adapted from Pesetsky 1982, ch. 4, ex. 81-82)

This is precisely what we expect, if Nissenbaum (1998, 2000) is correct that the configuration that permits a PG in a subject is the same as that involved with a PG in a sentential adjunct: both involve operator-hosting phrases which merge to vP and take advantage of Predicate Abstraction triggered within it when there is successive-cyclic movement from it.[12]

    Notice that in (47), the PG-containing subject linearly precedes the gap left behind by the movement chain that licenses it. This is the opposite of what we see with PGs in sentential adjuncts, which linearly follow the licensing gap, since such adjuncts are linearized to the right of the clause they attach to. Importantly, it is apparently the case that the PG licensing asymmetry that is vital for this paper applies to both PGs in sentential adjuncts and in subjects, regardless of this difference in how they are ordered with respect to the licensing gap. This is expected if the PG licensing asymmetry stems from structural facts, rather than from something like a linear constraint on the order of gaps and their fillers resulting from limitations of sentence processing. A similar point is made by Pesetsky (1982).[13]

4.4 Interim summary

So far, this paper has used the prediction about PGs in (33) above to argue for the correctness of the OCG. This generalization states that overlapping moved phrases will have the same relative order in their final landing sites, and in the edge of any phase they both successive-cyclically move from. This is precisely as CL predicts. However, recall that all of the English sentences focused on in the previous section obey the PCC, and thus involve order-reversal. The PG evidence has shown us that this order-reversal is established in the first vP

---

    b.   You$_1$ said that we should go to the park [before taking pictures of you$_1$/*yourself$_1$].

[12]Though the PG-licensing configuration is the same, semantic composition is more complex for PGs in subjects. See Nissenbaum (1998, 2000) and Nissenbaum and Schwarz (2011) for discussion.

[13]Note that I do not argue that linear order constrains PG licensing. I do argue that facts about linear order correlate with PG licensing possibilities, but only because linearization constrains the set of syntactic derivations that PF can accept, and thus indirectly limits what structures will be available for a PG to inhabit.

in which both moving phrases are present. While CL leads us to expect this, CL by itself does not provide an obvious reason why order-reversal is necessary in such contexts in the first place. This early order-reversal in vP is even more puzzling given that, as mentioned above, Richards (1997, 1999) predicts that this is precisely the opposite of what should occur. These considerations make salient the need for a more refined understanding of when overlapping paths are permitted to either cross or nest. I consider this issue next, using the PCC as a starting point for an alternative account of crossing versus nesting.

5   On the PCC

In this section, I provide a CL-based account of the PCC, as a prelude to a more formalized proposal about the distribution of crossing and nesting paths. It will be useful to first discuss a rationale for the PCC, since examining the derivations in which it arises will necessarily bring into focus several important issues about when crossing and nesting occur. The primary goal of this section is to provide a linearization reason for why nesting paths will sometimes be required in English, while the following section will consider in detail the general syntactic mechanisms that allow crossing or nesting to occur in a given context.

The PCC (Pesetsky 1982) describes the fact that in many contexts involving overlapping A′-movement, it is necessary for the paths to nest, resulting in the base position and final position of the moved phrases in question reversing, as exemplified once more in (48):

(48)   *Forced order reversal as described by PCC*

    a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2?   [2 1 ... 1 2]

    b.   * [This student]$_1$, what$_2$ should we talk to t_1 about t_2?   *[1 2 ... 1 2]

As we've seen, in configurations like (48), topic/focus movement must form a final landing site above *wh*-movement, but even when this fact is held constant, as it is in (48), the PCC still holds. Indeed, Pesetsky shows that the PCC applies very generally to contexts with overlapping A′-movement in English, regardless of the sub-types of A′-movement involved.

However, if Richards (1997) and Nissenbaum (2000) are correct about the derivation of multiple *wh*-questions in English, then these violate the PCC since they involve crossing paths, as shown in (24) above. What distinguishes English multiple *wh*-questions from PCC-obeying contexts is that in the former, only one phrase can move overtly, while in the latter, all moving phrases are overt (or at least can be). Richards (1997) takes advantage of this fact, hypothesizing that the absence of the PCC in multiple *wh*-questions, and the arising of the PCC in situations with multiple overt A′-movements, is connected to the impossibility of CP hosting multiple overt specifiers in English and languages like it.[14]

---

[14]An explanation for the PCC that depends on the phonological status of the movements involved, which I will pursue later in this section, might lead us to expect that the PCC will not apply when one of the moved phrases involved is covert. This is not the case, however, as we see in (i), which includes a null relativizing operator. While this element can be overt, even if it is not the PCC applies to it, as we see:

For Richards, nothing prevents a second movement to the same head in English if covert, as in multiple *wh*-questions. However, since a second overt movement to the same CP is by hypothesis unavailable, the only possibility for a situation with overlapping overt A′-movements is the following: the first C merged will trigger overt movement of the closer phrase, with the C merged later on then attracting the lower phrase, as in (49) below. Richards follows Kitahara (1991, 1997) in hypothesizing that this derivation is responsible for the PCC, and posits that the somewhat degraded status of the relevant sentences stems from this assumed locality-violating attraction of the lower moving phrase:

(49)   *PCC configurations under Richards' analysis (following Kitahara)*
         $[_{CP2}$ XP_2 C2 ... $[_{CP1}$ XP_1 C1 ... t_1 t_2]]

In contrast, Richards argues that the possibility of overt tucking-in within CP in Bulgarian, as revealed by its multiple *wh*-questions (24a), correctly predicts the absence of the PCC in this language, as shown by the structure in (50) below. Richards' analysis of (50) is that since the first C merged can overtly attract both *wh*-phrases here, the lower *wh*-phrase need not remain in situ until the merge of a second C. Rather, both move with crossing paths to the first C, after which the second C attracts the higher of the two *wh*-phrases:

(50)   *An anti-PCC configuration in Bulgarian* (Richards 1997, p. 41, ex. 41)
         a. $[_{CP2}$ **Koj**$_1$ se   opitvat da razberat $[_{CP1}$ t_1 **kogo**$_2$ t_1 e   ubil t_2]]?
               who self try   to find.out   whom   AUX killed
         b.   * $[_{CP2}$ **Kogo**$_2$ se   opitvat da razberat $[_{CP1}$ **koj**$_1$ t_1 e   ubil t_2]]?
               whom self try   to find.out   who   AUX killed

This account for the difference between English and Bulgarian is not compatible with phase theory, or the results in the first half of this paper, both of which mandate against an account in which the initially lower A′-moving phrase is non-successive-cyclically attracted from its origination position in PCC derivations (49). In what follows, I will argue for an alternative account of the PCC that is compatible with this paper's results. I will first provide a general rationale for the PCC, building on Richards' proposal that the covert/overt distinction for multiple specifiers is the central factor, and then propose a more formalized account of the distribution of crossing versus nesting multiple specifier formation.

(i)   a.   This is the saxophone $\varnothing_2$ I know [precisely which songs]$_1$ to never play t_1 on t_2.
        b.   * This is the saxophone $\varnothing_2$ I know [precisely which songs]$_1$ to never use t_2 to play t_1.

For the purposes of this paper, it would suffice to hypothesize that in contexts like (i) the null moved element was originally overt and thus linearized, before the application of processes like Free Deletion in Comp (Chomsky and Lasnik 1977). Another possibility is that linearization is evaluated at a level prior to the stage of the PF derivation at which morpo-phonological form is assigned (Embick 2010; Arregi and Nevins 2012; Ostrove 2018, a.o.), and is thus insensitive to whether elements are ultimately overt or not.

5.1   A linearization understanding of the PCC

I argue that if English generally regards tucked-in specifiers of A′-movement as covert, as discussed above in reference to multiple *wh*-questions, then we come to a natural understanding of why the PCC must apply in contexts with overlapping overt A′-movement in this language. This analysis will require maintaining a distinction between inner specifiers derived in different ways, however.

   I hypothesize that when two phrases move to one head with crossing paths, spell-out of the containing phase linearizes the outer moved phrase at only the head of its chain, while the phrase that forms the inner specifier is linearized in only the tail of its chain, as in (51a). Further, I hypothesize that when two phrases move to one head with nesting paths, spell-out of the containing phase linearizes both phrases at the head of their chains, as in (51b), in which α moves overtly despite occupying an inner specifier after movement of β above it:[15]

(51)   *Multiple specifier linearization in English*

   a.   *Crossing paths → covert inner specifier*

   Linearization: $\alpha < X < \beta$
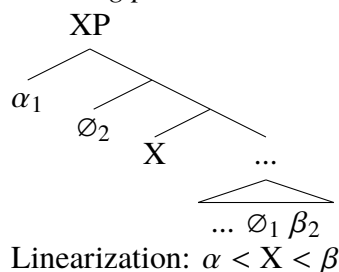
   b.   *Nesting paths → only overt movements*

   Linearization: $\beta < \alpha < X$
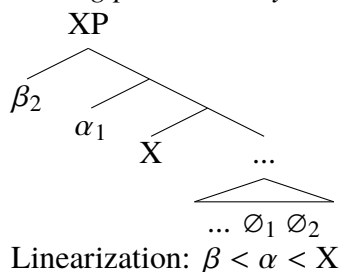
This is in contrast to a theory in which all inner specifiers, regardless of how they are formed, are treated as derived by covert movement.[16] While I have not yet explained what allows tucking-in to be avoided in order to derive a double overt movement configuration like (51b), here I will show that positing the possibility of such a movement pattern yields the correct

---
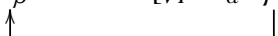
[15]The diagrams in (51) can be understood as illustrations of the structural descriptions under which each rule of chain linearization applies. The illustration in (51a) is analogous to what is proposed in Pesetsky (2000), who along with Bobaljik (1995); Brody (1995); Groat and O'Neil (1996) and others, I follow in assuming that the overtness or covertness of a given movement step is determined at spell-out. This is in contrast to the (inverted) Y-model (Chomsky and Lasnik 1977) under which covert movement occurs post-spell-out.

[16]For CL, such a theory would wrongly predict the impossibility of ever overtly moving more than one phrase from a phase in English, since a phase would only ever allow one phrase to be linearized at its edge.

linearization for PCC-obeying sentences. I will discuss the principles that determine the distribution of nesting versus crossing / tucking-in explicitly in the next section.

To see how the hypothesis in (51) facilitates an understanding of the PCC in English and languages like it, first consider the schema for a vP in (52) below, where the phrases α and β both undergo A′-movement to the vP edge. Here β tucks-in below α. If tucking-in movement is regarded by English PF as covert (51a), then β will be linearized at the tail of its movement chain, at the right edge of this vP:

(52)   *Covert tucking-in within vP* (Linearization: $\alpha < S < V < \beta$)
       $[_{vP}\quad \alpha\ \varnothing_\beta\ S\ \text{v-V}\ [_{VP}\ t_\alpha\ \beta\ ]]$

If later on α moves to an overt landing site external to the vP, as in (53) below, there will be no issue: since α occupied the left linear edge of the vP in (52), CL allows α to overtly move on leftward in this way. However, the same does not apply to β. Since β did not overtly move to the left edge of vP in (52), it cannot overtly exit the vP in this way without causing a linearization contradiction, as (53) shows. This is because β cannot both follow the content of vP as (52) established, and also precede it as the later movement in (53) requires:

(53)   *Later overt movement from vP* (Linearization: $\alpha < \beta < S < V$)
       $[\alpha\ ...\ \beta\ ...\ [_{TP}\ S\ [_{vP}\ ...\ \text{v-V}\ ...\ ]]]$
       Illicit if following (52) due to contradicting previous order $S < V < \beta$

This contradiction would not arise if in (52) β had moved above α in vP rather than tucking-in, as in (54) below. Following the rule in (51b) above, neither movement is treated as covert in the structure in (54). After (54), further overt movement of β and α is licit, since they both occupy the linear edge of this vP. Thus only a nesting derivation like (54), rather than one with crossing paths like (52), allows multiple overt A′-movements from the vP.

(54)   *Nesting paths in vP yields two overt movements* (Linearization: $\beta < \alpha < S < V$)
       $[_{vP}\ \beta\ \alpha\ S\ \text{v-V}\ [_{VP}\ t_\alpha\ \beta\ ]]$

Importantly, notice that by forming nesting paths in (54), these moving phrases reverse their order within vP. Provided that their subsequent landing sites preserve this reversed ordering, the result is a CL-satisfying derivation fitting what the PCC describes, in which the origination positions and final positions of the moved phrases are reversed. This is precisely the derivation that the PG facts examined earlier in this paper have shown us occurs in reality: overlapping moved phrases take on an order that matches their final relative order within the first vP they occupy, which in the case of a PCC-obeying sentence, involves immediate order-reversal followed by order-preserving movement onward.

5.2 Correct predictions and further puzzles

The above account of the PCC has, for convenience, assumed that crossing and nesting multiple specifier formation are both possible in principle, unless the concerns of linearization further restrict the derivation. As just argued, it is the influence of linearization that makes nesting the only possibility for PCC sentences. This understanding correctly allows the possibility of a crossing-paths derivation in English multiple *wh*-questions (24b): since the non-initial *wh*-movement in such contexts has a covert landing site, no linearization problem arises from a lower *wh*-phrase tucking-in within vP before doing the same later on in CP. The same considerations also predict that Bulgarian will be capable of violating the PCC, as in (50) above: since this language allows multiple overt specifiers in one phrase, it can simply allow overlapping overt A′-movements to overtly tuck-in in the vP phase before then moving on. The problem posed by the covertness of tucked-in specifiers in English thus does not arise in Bulgarian, which therefore allows overt nesting path derivations.

However, if crossing and nesting multiple specifier formation are in free variation, then we have no explanation for why in English and Bulgarian multiple *wh*-questions, crossing paths are not merely an option, but a requirement, as (55) shows:

(55)   *Unacceptable nesting multiple wh-questions (compare with (24) above)*
      a. *Bulgarian* (Richards 1997, p. 277 ex. 75b)

          * **Kakvo**$_2$ **kogo**$_1$ e   pital Ivan t_1 t_2?
          what.ACC who.NOM AUX asked Ivan

      b. *English*
          * **What**$_2$ $\varnothing_1$ did Ivan ask **who**$_1$ t_2?

The same question is relevant for numerous obligatory crossing path contexts shown by Richards (1997), such as *wh*-scrambling in Japanese, negative fronting in Bulgarian, multiple object shift in Germanic languages, and multiple clitic clustering in Tagalog and Serbo-Croatian.[17] In order to address this concern, it is necessary to develop a more precise account of when crossing and nesting paths are permitted.[18] I do this in the next section.

6 Formalizing the distribution of crossing and nesting

Recall that Richards argues that crossing paths occur when multiple movements target the same projection. In section 4, we saw evidence from PGs that in English sentences that obey the PCC, the reversed order that the moved phrases have in their final landing sites is mimicked by the configuration these phrases form upon successive-cyclically moving

---

[17]See Bruening (2001) for evidence that multiple instances of Quantifier Raising must also cross paths.

[18]Similarly, what I have just said accurately rules in the anti-PCC configuration in the Bulgarian (50a) above, but also rules in a PCC-style derivation for this Bulgarian configuration, which as we see in (50b), is unacceptable. I will show later in this paper that the expected freedom of ordering arises in other contexts in Bulgarian (as well as Romanian, another multiple *wh*-fronting language).

through the edge of the first vP in which they are both present. This indicates that crossing paths do not always occur in situations where multiple phrases move to the same projection.

In response to this finding, I propose that crossing paths / tucking-in is not a requirement of multiple movements to one phrase, but rather applies to movements that are triggered by one feature, as diagrammed in (56) below. Here the unvalued feature [uF1] on the head X attracts, with crossing paths, the goals α and β which both bear the feature [F1]:[19]

(56) *Proposal #1: One feature attracts multiple goals → tucking-in*
$$[_{XP} \ \alpha_{[F1]} \ \beta_{[F1]} \ X_{[uF1]} \ ... \ t_\alpha \ t_\beta \ ]$$
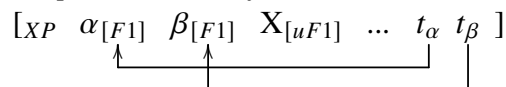
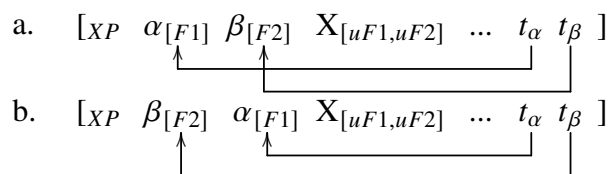Additionally, I propose that when a head bears multiple movement-triggering features, the goals respectively attracted by those features are free to cross or nest. I hypothesize that if the crossing requirement is limited to multiple movements triggered by one feature, then when multiple goals each move due to attraction by separate features, the grammar does not constrain the form of their paths.[20] This concept is illustrated in (57), where the head X bearing the features [uF1] and [uF2] can attract the goals $\alpha_{[F1]}$ and $\beta_{[F2]}$ in a way yielding either crossing (57a) or nesting (57b) paths to its edge. The freedom of ordering in such contexts will be vital to capturing the full range of multiple movement configurations.

(57) *Proposal #2: Attraction by separate features on one head → variability*
a. $$[_{XP} \ \alpha_{[F1]} \ \beta_{[F2]} \ X_{[uF1,uF2]} \ ... \ t_\alpha \ t_\beta \ ]$$
b. $$[_{XP} \ \beta_{[F2]} \ \alpha_{[F1]} \ X_{[uF1,uF2]} \ ... \ t_\alpha \ t_\beta \ ]$$

This hypothesis about the distribution of crossing and nesting can only be meaningful when combined with a proposal about how movement-triggering features are distributed. Following much contemporary work in syntax, I assume that whatever features are responsible for ultimately attracting a given phrase to its final position are also instantiated on the heads of intermediate phases which that phrase's movement must pass through. This is a straightforward expectation of Probe-Goal theories of syntactic dependencies under which

---

[19]I follow Chomsky (2000) and a great deal of following work in taking unvalued features (prefixed with "u" in (56-57) below) to be those which probe their c-command domain for a goal bearing the relevant feature, as a prerequisite for establishing an agreement or movement dependency.

[20]Alternatively, we might hypothesize that the freedom of specifier order in this situation stems from the possibility of the attracting features targeting their goals in either order. Since the attracting features in this case target goals with distinct features, it is plausible that α in (57) will not serve as an intervener for the purposes of the attraction of β by [uF2], if that feature happened to trigger movement before [uF1]. A similar possibility is raised by Doggett (2004). Such a derivation would create nesting paths without violating Richards' hypothesis about tucking-in, since the lower phrase would be attracted before the higher one. However, such a derivation would instead violate Richards' hypothesis that multiple specifier formation is constrained by a need to keep movement paths as short as possible, given that it involves attraction of the furthest goal first.
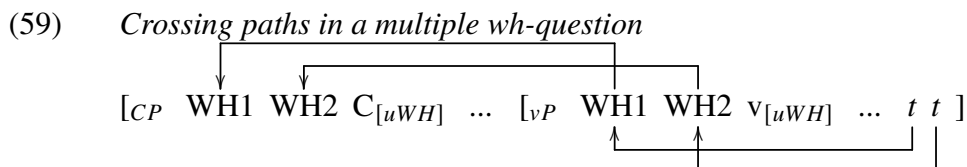
movement is feature driven. However, it is worth asking what mechanisms are responsible for enforcing the necessary distribution of features. In this paper, I will argue that the relevant facts about crossing and nesting emerge from the following two principles:[21]

(58)   a.   *Attractor minimization principle*
            A given head enters the derivation with at most one movement-triggering feature, unless additional features are motivated by the principle in (58b).
    b.   *Intermediate attractor principle*
            If a given head X bears a movement-triggering feature [uF], and a phase head Y structurally intervenes between X and the corresponding goal phrase bearing [F], then Y may also bear [uF].

The principle in (58a) prevents the addition of superfluous features in a way that we are about to see is necessary for the account of contexts with obligatory tucking-in. The additional principle in (58b) allows a head to bear additional movement-triggering features to the extent that they are necessary to facilitate successive cyclic movement.

6.1   Deriving the distribution of crossing and nesting

Consider a multiple *wh*-question, which as we've seen, involves multiple *wh*-phrases moving to the same CP with obligatory crossing paths. Given the principle in (58a), C in such a configuration must have a single [uWH] feature capable of attracting both *wh*-phrases. Given the principle in (58b), the v which intervenes between C and these *wh*-phrases also has [uWH]. First, [uWH] on v attracts these phrases, which being moved by the same feature form a crossing path configuration in vP, as (56) dictates. Precisely the same thing happens when these phrases are then targeted by the [uWH] feature on C, as (59) shows:

(59)   *Crossing paths in a multiple wh-question*

$$[_{CP}\ \ \text{WH1}\ \ \text{WH2}\ \ C_{[uWH]}\ \ ...\ \ [_{vP}\ \ \text{WH1}\ \ \text{WH2}\ \ v_{[uWH]}\ \ ...\ \ t\ \ t\ \ ]$$

If in this context it were possible for C to have two distinct [uWH] features (instantiated on

---

(i)   *Economy of Movement Metric* (Chomsky 1992; Reinhart 2006; Overfelt 2015)
        If a derivation D_1 of a spell-out domain α converges without some movement operation, then D_1 blocks a derivation D_2 of α that includes that movement operation.

In the context of a theory in which movement is feature-driven, a principle of this variety is reducible to an economy constraint not on movement operations themselves, but on the introduction of features responsible for triggering movement. Such a concept will appropriately rule in only derivations with the fewest number of movement-triggering features needed for a convergent result.

v as well) which separately attract the two *wh*-phrases, we would expect these phrases to be able to form a nesting path configuration within the vP, following the principle in (57). This is illustrated in (60), where these phrases nest within vP and then move on to CP in order-preserving fashion, both of which movements (57) permits in this situation:

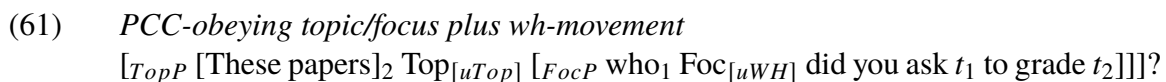(60)   *Hypothetical order reversal in a multiple wh-question*

$[_{CP}$ WH2 WH1 $C_{[uWH1,uWH2]}$ ... $[_{vP}$ WH2 WH1 $v_{[uWH1,uWH2]}$ ... t t]
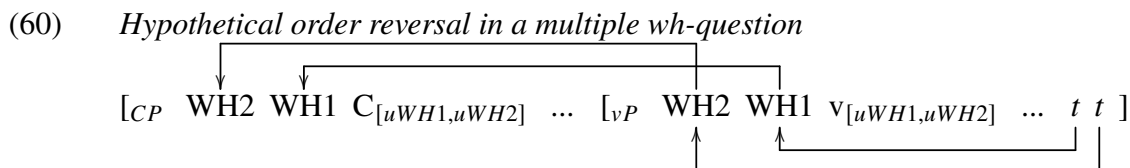
In reality, such a result is not possible for (non-D-linking) multiple *wh*-questions, as we saw in (24b) above. The hypothetical derivation in (60) is appropriately excluded if the principle in (58a) blocks the introduction of a C which superfluously bears two distinct tokens of [uWH], rather than just one, since one feature is all that is needed to attract both of the *wh*-phrases in this context. I hypothesize that all configurations with multiple movements to the same position that obligatorily form crossing paths involve those movements being driven by a single attracting feature, as in (59). The feature involved in triggering those movements will, of course, vary in accordance with the type of movement occurring.

   Next, lets examine a context in which the PCC applies. First I'll consider a sentence with topic/focus movement and *wh*-movement, as discussed extensively above. Since as mentioned the landing sites for these movements are rigidly ordered, I will assume that they target separate projections in the left periphery. Following Rizzi (1997), the respective projections could be termed TopP and FocP, as in (61) below. The respective movement triggering features on the heads of these projections are labeled [uTop] and [uWH]:

(61)   *PCC-obeying topic/focus plus wh-movement*
        $[_{TopP}$ [These papers]$_2$ $Top_{[uTop]}$ $[_{FocP}$ who$_1$ $Foc_{[uWH]}$ did you ask t_1 to grade t_2]]]?

In order to trigger vP-level movement here, v must bear attracting features corresponding to both landing sites in the left periphery as in (62), given the principle in (58b) above:

(62)   *Pre-movement structure of the vP in (61)*
        $[_{vP}$ $v_{[uWH,uTop]}$ ... who [these papers]]
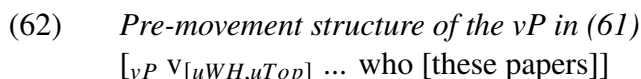
Since these A′-moving phrases are attracted by distinct features on v, then given the proposal in (57) above, we predict that they can either form crossing or nesting paths within this vP. If they form crossing paths here, the phrase that forms the inner spec-vP will move covertly, for the reasons laid out in section 5.1 above. As we saw in that section, since in such a context the inner moved phrase will move covertly in vP before overtly moving from it, such a derivation will result in a linearization contradiction. In section 5.1 we also saw that, in contrast, no linearization issue arises if overlapping phrases move in nesting fashion in vP. In this case both phrases move overtly, while reversing their order. After this initial reversal, these phrases can move on to their final landing sites in order preserving fashion. This is illustrated for (61) by (63) below. As we saw in section 5, such derivations yield the

structure described by the PCC, and also match the PG facts shown in section 4:

(63) *Nesting paths in vP permitted by distinct features on v*

$[_{TopP}$ [These papers]$_2$ Top$_{[uTop]}$ $[_{FocP}$ who$_1$ Foc$_{[uWH]}$ $[_{vP}$ t_2 t_1 v$_{[uTop,uWH]}$ V t_1 t_2]]]
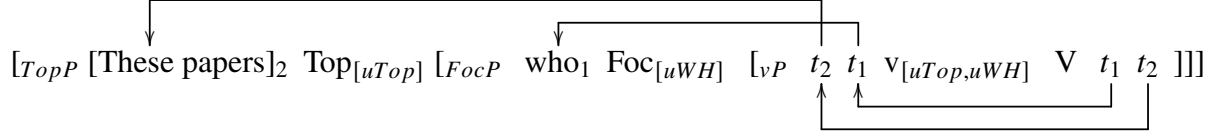
   This analysis extends to all other contexts with two phrases undergoing separate instances of overt overlapping A′-movement in English, all of which obey the PCC (Pesetsky 1982). For instance, consider a sentence with relativization and interrogative *wh*-movement (64):

(64) *Relativization plus interrogative wh-movement: PCC obeyed*
   a.   This is the book [which$_2$ I asked [who$_1$ you gave t_1 comments on t_2].
   b.   * This is the book [which$_1$ I asked [who$_2$ you sent a review of t_1 to a student of t_2].

We saw in (39) above that this configuration, like others of the same variety, displays the PG-licensing asymmetry which indicates that the final order of moved phrases here was established in the initial vP where those phrases originated. As in the context just examined above, the result that these moved phrases reverse order upon their movement within the initial vP is predicted, given (57), if these phrases' movements are driven by separate features. We could call these features [uREL] and [uWH] in (64), but the way these features are named is not important: however we label them, the features involved here are plausibly of distinct classes. This leads us to expect the possibility of an order-reversing, non-tucking-in derivation in (64)—a possibility that must be selected here, yielding the PCC, since a tucking-in derivation would pose a linearization problem as discussed in section 5.1.[22]

---

[22]The applicability of this account is less obvious for recursive embedded *wh*-questions. The PCC holds for such sentences, as (i) shows, despite the fact that the movements here are presumably featurally identical:

(i)   a.   Tell me $[_{CP}$ [what subject]$_1$ you know $[_{CP}$ who$_2$ to talk to t_2 about t_1]].
      b.   * Tell me $[_{CP}$ who$_2$ you know $[_{CP}$ [what subject]$_1$ to talk to t_2 about t_1]].

If both of the *wh*-phrases here were attracted by one *wh*-feature, we would expect such a sentence to have a form like that of a multiple *wh*-question, with overt movement of the initially highest *wh*-phrase, and covert tucking-in movement of the lower one. This result is impossible in this context, however, as (ii) shows:

(ii)   *Tell me $[_{CP}$ who$_1$ $\varnothing_2$ you know $[_{CP}$ t_1 t_2 to talk to t_1 about [what subject]$_2$]].

The actual result in (ia) above will accurately be predicted by positing that the two CP edges that serve as landing sites for *wh*-movement in this context bear distinct *wh*-sensitive features. This could be formalized by counter-indexing those features, for instance as [uWH1] and [uWH2]. If in (ia) the outermost embedded C bears [uWH1] and the inner embedded C bears [uWH2], both of which are also instantiated on the intermediate v that they both c-command, then the two *wh*-phrases here can be attracted separately within the vP and thus form a nesting configuration before moving onward, creating the PCC-obeying result we see in reality in (ia). As discussed extensively above, nesting rather than crossing (and thus a derivation culminating in order reversal) is required in the initial vP here, given that in English crossing paths within the vP would result in an ordering contradiction due to the covertness of the inner specifier.

6.1.1 A further constraint on overtness/covertness

A problem for the above analysis arises when we consider certain multi-clausal PCC-obeying sentences more carefully, such as (65) below. The analysis of the PCC that I have proposed entails order-reversal of moving phrases within the first vP where they are both present, followed by order-preserving movement of those phrases to their final landing sites. The traces placed in (65) are faithful to this proposal, showing the initial order reversal, followed by preservation of that reversed order in the intermediate CP and matrix vP that these phrases pass through on the way to their final positions in the matrix clause:

(65) *Initial reversal followed by order preserving cross-clausal movement*
      [This book]$_2$, who$_1$ did you [$_{vP}$ t_2 t_1 say [$_{CP}$ t_2 t_1 that we should [$_{vP}$ t_2 $t_1$talk to t_1
      about t_2]]]?

For these phrases to preserve the order reversal established in the initial vP, their subsequent movement steps must form crossing paths. This requires the lower moving phrase *who*$_1$ to tuck-in below *this book*$_2$ in the intermediate CP and matrix vP edges, as the traces in (65) show. However, if in English tucking-in is derived by covert movement as discussed above, it is unclear how movement of the lower phrase *who*$_1$ in (65) manages to be ultimately overt here, as opposed to becoming covert after tucking-in in the embedded CP edge.

  In order to avoid positing that *who*$_1$ in (65) covertly moves through lower specifiers of the intermediate CP and matrix vP before overtly moving to its final landing site (a derivation which would incur a linearization contradiction), I propose the following:

(66) *Chain Consistency Condition*[23]
      The overtness or covertness of the initial step of a movement chain driven by a given feature [uF] must be matched by all following movement steps in that chain driven by [uF], regardless of how they would otherwise have been linearized.

This principle allows tucking-in of *who*$_1$ in the edges of the embedded CP and matrix vP in (65) to be overt, contrary to usual circumstances, since those movement steps were preceded by overt non-tucking-in movement of *who*$_1$ within the initial vP.[24] Thus the moved phrases in (65) can overtly move on from the initial vP in order-preserving fashion as required.

6.2   A return to Bulgarian and extension to Romanian

Above, we saw a fact from Richards (1997) showing that Bulgarian lacks the PCC, repeated in (67). In this example, an embedded question within a matrix one involves two *wh*-phrases originating in the same clause that land in separate CP edges, without reversing their order.

---

[23]Fox and Nissenbaum (2018) precede me in positing a condition of this sort ("Early Determination") for covert movement. The formulation I offer here applies to overt movement as well.

[24]This principle also permits an element to undergo covert movement after landing in the terminal position of a given overt movement chain, though initiating overt movement after completing a covert movement chain will generally be ruled out by linearization for the reasons discussed in section 5.

(67)   *Anti-PCC in Bulgarian* ([Richards 1997](), p. 41, ex. 41)

    a.   **Koj**$_1$ se opitvat da razberat [$_{CP}$ t_1 **kogo**$_2$ t_1 e   ubil t_2]]?
        who self try   to find.out   whom   AUX killed

    b.   * **Kogo**$_2$ se opitvat da razberat [$_{CP}$ **koj**$_1$ t_1 e   ubil t_2]]?
        whom self try   to find.out   who   AUX killed

Order reversal is not only unnecessary here, as (67a) shows, but impossible (67b). To account for the PCC in English, I have proposed that both crossing and nesting multiple specifier formation are permitted when the movements are driven by separate features, but that in English overt overlapping A′-movements must end up nesting because the covert inner specifier of a crossing paths derivation would cause a linearization problem. However, since Bulgarian allows overt tucking-in as shown in its multiple *wh*-questions, nothing should prevent this language from allowing both of the orderings in (67). However, recall that in (56) above, I proposed that crossing paths are forced when a single feature on a given head attracts multiple goals. Given this proposal, the impossibility of order reversal in (67) would be expected if the two *wh*-phrases here are in fact attracted by one feature. This would be the case if in Bulgarian (unlike English) any feature sensitive to interrogative *wh*-phrases necessarily interacts with, and thus attracts, all such phrases in its c-command domain.[25] This would preclude the possibility of a nesting path derivation in (67).[26]

    If this hypothesis is correct, then the predicted freedom of crossing and nesting should arise in Bulgarian when two *wh*-phrases undergoing overlapping movement are more clearly featurally distinct—in particular, when at most one of them is interrogative. [Rudin]() ([1988]()) reports Bulgarian examples of interrogative *wh*-movement overlapping with relativization that fit this prediction (68). In (68a) below, we see an example where these two A′-movements form crossing paths, and in (68b), we see a situation where the same two movements nest instead, due to a corresponding reversal of their landing sites:[27]

(68)   *Freedom of ordering for distinct A′-movements in Bulgarian*

    a.   *Crossing paths* (Adapted from [Rudin 1988]() ex. 8c)

       čoveka [**kojto**$_1$ ne znaeš   [**kakvo**$_2$ kazvat [če e   kupil t_1 t_2]]]
       man   who.REL NEG know.2SG what   say.3PL that has bought

       'The man who you don't know what they say that he bought'

---

[25]This would instantiate what following [Deal]() ([2015a](),[b]()) we could classify as an *insatiable* feature.

[26]Another possibility is that an independent factor determines the order of *wh*-phrases in (67). Indeed, [Billings and Rudin]() ([1996]()) argue that various factors can interfere with the order of Bulgarian *wh*-phrases.

[27]Example (68b) shows relativization crossing not one but two interrogative *wh*-phrases. While the crossing of just one such phrase would be enough to make the point here, Rudin chose to include more than one to emphasize the unboundedness of relativization in this particular configuration.

b. *Nesting paths* (Adapted from Rudin 1988 ex. 19)

edna kniga [**kojato**$_1$ se čudja   [koj$_2$ t_2 znae   [**koj**$_3$ t_3 prodava t_1]]]
a   book which   wonder-1sG who   knows who   sells

'A book which I wonder who knows who sells (it)'

This freedom of ordering in Bulgarian is precisely what we expect to see at least some of the time, if moved phrases attracted by separate features may form multiple specifiers with crossing or nesting paths, provided that one of those options isn't independently excluded. In English PCC-obeying sentences, a linearization problem posed by the covertness of tucking-in bans crossing paths in those contexts. However, since Bulgarian allows overt inner specifiers formed by tucking-in, we accurately predict that it will in principle permit both overt crossing and overt nesting configurations, as we've seen.

6.2.1   Correct predictions for Romanian

Like Bulgarian, Romanian displays multiple *wh*-fronting (69). According to Rudin (1988), in Romanian all interrogative *wh*-phrases move to specifiers of CP just as in Bulgarian.

(69)   *Romanian multiple wh-fronting* (Adapted from Rudin 1988, ex. 9a)

Cine$_1$ **cui**$_2$ **ce**$_3$ ziceai   [că i-a   promis t_1 t_2 t_3]?
who   who.DAT what said.2sG that him.DAT-has promised

'Who did you say promised what to whom?'

If this is so, then Romanian also evidently allows overt tucking-in, and thus we expect Romanian to behave exactly like Bulgarian as far as the PCC is concerned. That is, we expect Romanian to show the same freedom of ordering that we saw in the Bulgarian examples in (68) above. This prediction is correct. In (70) below, we see that overlapping relativization and interrogative *wh*-movement in Romanian may either cross or nest:[28]

(70)   *Freedom of ordering for distinct A′-movements in Romanian*

a.   *Nesting paths* (Adapted from Comorovski 1986, ex. 4b)

Ion,   [[**pe care**]$_1$ am   uitat [**cine**$_2$ mi-ai   spus că t_2
John, whom   have.1sG forgot who   me.DAT-have.2sG told   that

ţi   l$_2$-a   prezentat t_1]]
you.DAT CLITIC-has introduced

'John, whom I forgot who you told me introduced to you'

---

[28]The *wh*-moved direct objects in these Romanian examples are clitic-doubled. Comorovski (1986) argues that this clitic is not something like a resumptive pronoun that occurs in the absence of movement, but rather that *wh*-movement is applying here as usual, with the doubled clitic being simply irrelevant.

b. *Crossing paths* (Donca Steriade, p.c.)

? Ion, [**care**$_1$ am   uitat   [[**pe cine**]$_2$ mi-ai   spus că t_1
  John, who   have.1SG forgotten whom   me.DAT-have.2SG told that

ţi   l$_2$-a   prezentat t_2]]
you.DAT CLITIC-has introduced

'John, who I have forgotten whom you told me introduced to you'

Example (70a) shows nesting paths, which Comorovski (1986) reports numerous other instances of in Romanian, while (70b) shows crossing paths, which we have also seen in the multiple *wh*-question in (69) above. Thus we have evidence that like Bulgarian, Romanian allows both overt crossing and overt nesting at least some of the time. This is importantly in contrast to English, which only allows overlapping overt movements to nest, as we've seen.

7 Stranding in edges with multiple moved phrases

Many works have argued for the existence of patterns where a successive-cyclically moving phrase can strand material in one of its intermediate landing sites (McCloskey 2000; Barbiers 2002; Wiland 2010; Torrence 2018; Davis 2020, a.o.). CL predicts that if two phrases XP and YP successive-cyclically move from a given phase ZP such that XP forms the outer spec-ZP and YP forms the inner spec-ZP, then stranding in the edge of ZP should only be possible by YP. This is because an element stranded by XP would be crossed over by the movement of YP, though material stranded by YP will not be crossed by anything:

(71)   *Prediction for intermediate stranding with two overlapping moved phrases*

a.   *Stranding by inner of two specifiers possible*
[XP YP(α) ... [$_{ZP_{[Phase]}}$ XP YP(α) Z ... XP YP(α)]]

b.   *Stranding by outer of two specifiers impossible*
[XP(α) YP ... [$_{ZP_{[Phase]}}$ XP(*α) YP Z ... XP(α) YP]]

This is a manifestation of the prediction of CL that movement from a phase must generally exit from its linear edge, as described in section 2 above. In (71a), stranding of α by YP in the edge of the phase ZP causes no problem, since throughout this derivation, the relative ordering XP < YP < α is maintained regardless of whether stranding occurs. In contrast, in (71b) stranding of α by the outermost moving phrase XP results in a linearization problem, since in this case α precedes YP at the completion and spell-out of the phase ZP, but would follow YP if left behind in the edge of ZP after XP and YP move on.[29]

---

[29]Note that re-arrangement of the relevant constituents within the phase edge prior to stranding cannot occur to ameliorate this problem, given the well-supported ban on edge-internal movement within the same phrase (Ko 2007, 2011, 2014), a concept also predicted by many versions of anti-locality (Ishii 1999; Grohmann

A number of works have argued that certain adjuncts of *wh*-phrases can be stranded at clause edges in English (Urban 1999; Stroik 2009; Davis 2020; Zyman To appear), as illustrated in (72) below. Such patterns can be used to test the prediction illustrated above.

(72)   *Adjunct stranding under wh-movement in English* (Davis 2020, ex. 10-11)
       a.   **What**$_1$ (exactly/precisely) did you suppose t_1 (**exactly/precisely**) (that) they wanted t_1 (exactly/precisely)?
       b.   Tell me [**how much flour**]$_1$ (to the nearest pound) you said t_1 (**to the nearest pound**) (that) the bakery wants t_1 (to the nearest pound).

Some relevant test cases are shown in (73). These examples are complex since they require three clauses—two to provide landing sites for each overt A′-movement and one more for those movements to pass through and attempt stranding. There is nevertheless an expected contrast between an example where material is stranded in a clause edge by the inner moved phrase (73a) versus stranding by the outer one (73b). As the traces in (73) show, I have assumed that these sentences involve order-reversal in the initial CP, followed by order-preserving movement onward, as this paper predicts. Both sentences in (73) have this path structure, but importantly, only stranding by the outer moved phrase is degraded.

(73)   a.   *Intermediate stranding by inner moved phrase*
            This is the bakery which$_2$ they reported [**how much flour**]$_1$ (to the nearest pound) they thought [$_{CP}$ t_2 t_1 (**to the nearest pound**) that they sent t_1 to t_2].
       b.   *No intermediate stranding by outer moved phrase*
            Tell me [**how much flour**]$_2$ (to the nearest pound) they said who$_1$ they thought [$_{CP}$ t_2 (***to the nearest pound**) t_1 that they should ask t_1 to buy t_2].

8   The behavior of A-movement

This paper's proposals predict that the PCC should hold in a language like English regardless of whether the movements involved are A′-movement, A-movement, or a combination of the two. This is because any overlapping movements whatsoever should result in a multiple specifier structure, which will be subject to the general concerns of linearization discussed here. While I am not aware of a way to test overlapping A-movements, there is evidence that the PCC holds when A- and A′-movements overlap.

Holmberg et al. (2019) examine the properties of languages with symmetrical ditransitives, in which either the direct or indirect object is available for movement. As they discuss, some dialects of North-West British English are languages of this sort, and unlike standard English, allow passivization of the theme of a ditransitive (74):

(74)   % [A book]$_1$ was given/sent/handed him t_1 (by Mary).
         (North-West British English, Holmberg et al. 2019, ex. 8b)

2003; Bošković 2005; Abels 2003, 2012; Erlewine 2016, 2017, a.o.).

[Holmberg et al. (2019)](#) show that in many languages of this variety, North-West British English included, passivization of the theme of a ditransitive removes the possibility of A′-moving the indirect object, as in (75a). A′-movement of the theme combined with passivization of the indirect object is acceptable, however, as in (75b).

(75)   a.   * Who$_1$ was [the book]$_2$ given/sent/handed t_1 t_2 (by Mary)?
              (North-West British English, [Holmberg et al. 2019](#), ex. 9d)
        b.   [Which book]$_2$ was John$_1$ given/sent/handed t_1 t_2?
              (North-West British English, [Holmberg et al. 2019](#), ex. 9b)

[Holmberg et al. (2019)](#) argue for an approach to phase theory and multiple specifiers that predicts this restriction. However, an examination of the structure of the paths in the above examples reveals a straightforward alternative explanation for the unacceptability of (75a): this example, unlike (75b), violates the PCC. The contrast in (75) thus provides potential evidence that overlapping A- and A′-movements obey the PCC, as predicted.

Furthermore, for at least some speakers example (75b) is also acceptable in mainstream English. Since here the A′-moved phrase has the higher final landing site, then given CL and the OCG, we also expect it to have been the higher of the two phrases at the time when they passed through the vP edge. The fact that the *wh*-phrase in such a context can license a PG verifies this expectation (76):

(76)   [**Which book**]$_2$ was John$_1$ given t_1 t_2 for free [despite not being interested in **PG**$_2$]
        / [after buying **PG**$_2$ online]?

Though A-movement is independently incapable of PG-licensing ([Culicover and Postal 2001](#)), importantly, if the A′-moved phrase in (76) had not formed the outer specifier of the vP here, then we would not expect it to license a PG either. The fact that it does suggests that the OCG is not unique to A′-movement, but is applicable to the interaction between A- and A′-movement as well. This is what we expect under CL, since for this theory, the considerations of linearization that yield the OCG should apply to all forms of displacement.

9 Conclusion

In this paper, I examined several aspects of structures with overlapping phrasal movement. Using a prediction about PGs and multiple specifiers from [Nissenbaum (2000)](#), I argued that successive cyclic movement in such derivations is constrained in the way described by the OCG, which CL predicts. This result raised a question about the distribution of crossing and nesting multiple specifier formation, which I argued is resolved by understanding crossing paths as required only for multiple movements triggered by the same feature. I argued that this view facilitates an account of the PCC in English when combined with independent concepts about the distribution of covert movement, as well as the absence of the PCC in the multiple *wh*-fronting languages Bulgarian and Romanian. I also argued that these considerations make correct predictions about stranding in overlapping movement contexts,

as well as about the interaction of overlapping A- and A′-movement.

References

Abels, Klaus. 2003. *Successive-Cyclicity, Anti-locality, and Adposition Stranding*. Doctoral dissertation, University of Connecticuit.

Abels, Klaus. 2012. *Phases: An Essay on Cyclicity in Syntax*. Berlin: De Gruyter.

Arregi, Karlos, and Andrew Nevins. 2012. *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. Dordrecht: Springer.

Barbiers, Sjef. 2002. Remnant stranding and the theory of Movement. In *Dimensions of movement: From features to remnants*, ed. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers, and Hans-Martin Gärtner. John Benjamins. Https://doi.org/10.1075/la.48.04bar.

Billings, Loren, and Catherine Rudin. 1996. Optimality and Superiority: A new approach to overt multiple-*wh* ordering. In *Proceedings of Formal Approaches to Slavic Linguistics 3*, 35–60. Ann Arbor: Michigan Slavic Publications.

Bobaljik, Jonathan D. 1995. *Morphosyntax: The Syntax of Verbal Inflection*. Doctoral dissertation, Massachusetts Institute of Technology.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studica Linguistica* 59.

Branan, Kenyon. 2017. In-situ Wh-phrases in Superiority Violating Contexts don't have to be In-Situ. In *A Pesky Set: Papers for David Pesetsky*, ed. Claire Halpert, Hadas Kotek, and Coppe van Urk, 353–359. Cambridge, MA: MITWPL.

Brody, Michael. 1995. *Lexico-Logical Form*. Cambridge, MA: MIT Press.

Bruening, Benjamin. 2001. *Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy*. Doctoral dissertation, Massachusetts Institute of Technology.

Bruening, Benjamin. 2006. Differences between wh-scope marking and wh-copy constructions in Passamaquoddy. *Linguistic Inquiry* 37:25–49.

Chen, Sherry Yong. 2017. Movement constraints on the relative order of double topics in mandarin chinese. In *Toronto working papers in linguistics*, ed. Ruth Maddeaux, volume 38, 1–21. Linguistics Graduate Course Union (LGCU).

Chomsky, Noam. 1973. Conditions on Transformations. In *A Festschrift for Morris Halle*, ed. Stephen Andersopn and Paul Kiparsky, 232–288. New York: Holt.

Chomsky, Noam. 1977. On wh-movement. In *Formal syntax*, ed. Adriam Akmajian, Peter Culicover, and Thomas Wasow, 71–132. New York: Academic Press.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press. Linguistic Inquiry Monographs.

Chomsky, Noam. 1992. A minimalist program for linguist theory. In *MIT occasional papers in linguistics #1*. Cambridge, MA: MITWPL.

Chomsky, Noam. 1995. *The minimalist program*. MIT Press.

Chomsky, Noam. 2000. Minimalist Inquiries. In *Step by step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. Roger Martin, David Michales, Juan Urigareka, and Samuel Jay Keyser, 89–155. MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In *Ken hale: A life in language*, ed. Michael Kenstowicz. MIT Press.

Chomsky, Noam. 2008. On Phases. In *Foundational issues in linguistic theory: essays in honor of jean-roger vergnaud*, ed. Robert Freidin, Carlos P Otero, and Maria Luisa Zubizarreta. MIT Press.

Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8:425–504.

Cole, Peter, and Gabriella Hermon. 2000. Partial wh-movement: Evidence from Malay. In *Wh-Scope Marking*, ed. Gereon Müller and Armin von Stechow, 101–130. Amsterdam: Benjamins.

Comorovski, Ileana. 1986. Multiple WH Movement in Romanian. *Linguistic Inquiry* 17:171–177.

Culicover, Peter, and Paul Postal, ed. 2001. *Parasitic Gaps*. Cambridge, MA: MIT Press.

Davis, Colin. 2019. Possessor extraction in colloquial English: Evidence for successive cyclicity and cyclic linearization. *Linguistic Inquiry* just accepted. Https://doi.org/10.1162/ling_a_00369.

Davis, Colin. 2020. Crossing and stranding at edges: On intermediate stranding and phase theory. *Glossa* 5:1–32. Http://doi.org/10.5334/gjgl.854.

Davis, Colin. To appear. Possessor Extraction in Colloquial English: Evidence for Successive-Cyclicity and Cyclic Linearization. *Linguistic Inquiry* .

Deal, Amy Rose. 2015a. *Interaction and satisfaction in φ-agreement*. Handout of talk presented at Interaction of Grammatical Building Blocks colloquium, Universität Leipzig.

Deal, Amy Rose. 2015b. Interaction and satisfaction in φ-agreement. In *Proceedings of North East Linguistic Society 45*, ed. Thuy Bui and Deniz Ozyıldız, volume 1, 179–192. Amherst, MA: Graduate Linguistics Student Association.

Doggett, Teal B. 2004. *All Things Being Equal: Locality in Movement*. Doctoral Dissertation, MIT.

Doron, Edit, and Caroline Heycock. 1999. Filling and licensing multiple specifiers. In *Specifiers: Minimalist perspectives*, ed. D Adger, S Pintzuk, B Plunkett, and G Tsoulas, 69–89. Oxford University Press: Oxford.

Du Plessis, Hans. 1977. WH Movement in Afrikaans. *Linguistic Inquiry* 8.

Embick, David. 2010. *Localism versus Globalism in Morphology and Phonology*. MIT Press, Cambridge.

Engdahl, E. 1983. Parasitic Gaps. *Linguistics and Philosophy* 6.

Erlewine, Michael Y., and Hadas Kotek. 2018. Focus association by movement: Evidence from binding and parasitic gaps. In *Proceedings of Sinn und Bedeutung*, 1–9.

Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language & Linguistic Theory* 34. Https://doi.org/10.1007/s11049-015-9310-z.

Erlewine, Michael Yoshitaka. 2017. Why the Null Complementizer is Special in Complementizer-Trace Effects. In *A Pesky Set: Papers for David Pesetsky*, ed. Claire Halpert, Hadas Kotek, and Coppe van Urk, 232–288. Cambridge, MA: MITWPL.

Fanselow, G., and D. Lenertová. 2011. Left peripheral focus: mismatches between syntax and information structure. *Natural Language & Linguistic Theory* 29:169–209. Https://doi.org/10.1007/s11049-010-9109-x.

Fox, Danny, and John Nissenbaum. 2018. *Pied-piping and covert movement: Evidence from parasitic gap licensing*. Talk given at Workshop on Altaic Formal Linguistics 14: Workshop in honor of Shigeru Miyagawa.

Fox, Danny, and David Pesetsky. 2005a. Cyclic Linearization and its interaction with other aspects of the grammar. *Theoretical Linguistics* 31. Https://doi.org/10.1515/thli.2005.31.1-2.235.

Fox, Danny, and David Pesetsky. 2005b. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31. Https://doi.org/10.1515/thli.2005.31.1-2.1.

Groat, Erich, and John O'Neil. 1996. Spell-out at the LF interface. In *Minimal ideas*, ed. Werner Abraham, Samuel D. Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart. Amsterdam: John Benjamins.

Grohmann, Kleanthes. 2003. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Henry, Alison. 2012. Phase Edges, Quantifier Float, and the Nature of (Micro-) Variation. *Iberia: An International Journal of Theoretical Linguistics* 4:23–39.

Henry, Alsion. 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford,UK: Oxford University Press.

Holmberg, Anders, Michelle Sheehan, and Jenneke van der Wal. 2019. Movement from the Double Object Construction Is Not Fully Symmetrical. *Linguistic Inquiry* 50:677–721. Https://doi.org/10.1162/ling_a_00322.

Ishii, Toru. 1999. Cyclic spell-out and the *that*-trace effect. In *Proceedings of WCCFL 18*, ed. A. Carnie, J. D. Haugen, and P. Norquest, 220–231. Somerville, MA: Cascadilla Press.

Jenks, Peter. 2011. *The Hidden Structure of Thai Noun Phrases*. Doctoral dissertation, Harvard University.

Jenks, Peter. 2013a. *Accounting for a Generalization about Quantifier Float and Word Order in Classifier Languages*. Handout from presentation at SEALS 23, May 30.

Jenks, Peter. 2013b. Quantifier Float, Focus, and Scope in Thai. In *Proceedings of Berkeley Linguistics Society 39*, 90–107. Linguistic Society of America. Http://dx.doi.org/10.3765/bls.v39i1.3872.

Kayne, Richard. 1983. Connectedness. *Linguistic Inquiry* 14:223–249.

Kitahara, Hitatsugu. 1991. *Target-α: A unified thecky of movement and structure-building*. Doctoral Dissertaion, Harvard University.

Kitahara, Hitatsugu. 1997. *Elementary operations and optimal derivations*. Cambridge, MA: MIT Press.

Ko, Heejeong. 2007. Asymmetries in Scrambling and Cyclic Linearization. *Linguistic Inquiry* 38. Https://doi.org/10.1162/ling.2007.38.1.49.

Ko, Heejeong. 2011. Predication and edge effects. *Natural Language & Linguistic Theory* 29. Https://doi.org/10.1007/s11049-011-9143-3.

Ko, Heejeong. 2014. *Edges in Syntax: Scrambling and Cyclic Linearization*. Oxford University Press.

Koizumi, Masatoshi. 1995. *Phrase structure in minimalist syntax*. Doctoral dissertation, MIT.

Korsah, Sampson, and Andrew Murphy. 2019. Tonal reflexes of movement in Asante Twi. *Natural Language & Linguistic Theory* .

Legate, Julie Anne. 2003. Some Interface Properties of the Phase. *Linguistic Inquiry* 34. Https://doi.org/10.1162/ling.2003.34.3.506.

Longobardi, Giuseppe. 1984. Connectedness, scope and c-command. *Linguistic Inquiry* 16:163–192.

McCloskey, James. 2000. Quantifier Float and Wh-movement in an Irish English. *Linguistic Inquiry* 51. Https://doi.org/10.1162/002438900554299.

McCloskey, James. 2001. The morphosyntax of wh-extraction in Irish. *Journal of Linguistics* 37:67–100.

McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In *Derivation and explanation in the minimalist program*, ed. SD Epstein and TD Seeley, 184–226. Oxford, UK: Blackwell.

McGinnis, Martha. 1998. *Locality in A-movement*. Doctoral dissertation, MIT.

Medeiros, David. 2013. Hawaiian VP-remnant movement: A cyclic linearization approach. *Lingua* 127:72–97. Https://doi.org/10.1016/j.lingua.2013.01.008.

Nissenbaum, John. 1998. Movement and Derived Predicates: Evidence from Parasitic Gaps. In *The Interpretive Tract*, 247–295. Cambridge, MA: MIT Working Papers in Linguistics.

Nissenbaum, John. 2000. *Investigations of covert phrase movement*. Doctoral dissertation, MIT.

Nissenbaum, John, and Bernhard Schwarz. 2011. Parasitic degree phrases. *Natural Langauage Semantics* 19:1–38. Https://doi.org/10.1007/s11050-010-9061-7.

Ostrove, Jason. 2018. Stretching, spanning, and linear adjacency in Vocabulary Insertion. *Natural Language & Linguistic Theory* 36:1263–1289.

Overfelt, Jason. 2015. Extraposition of NPIs from NP. *Lingua* 164:25–44. Http://dx.doi.org/10.1016/j.lingua.2015.06.006.

Partee, Barbara, and Mats Rooth. 1983. Generalized conjunction and type ambiguity. In *Meaning, use, and interpretation of language*, ed. Rainer Baüerle, Christoph Schwarze, and Arnim von Stechow, 361–383. Berlin: de Gruyter.

Pesetsky, David. 1982. *Paths and categories*. Doctoral dissertation, MIT.

Pesetsky, David. 2000. *Phrasal Movement and Its Kin*. Cambridge: MIT press.

Podobryaev, Sasha. 2009. Postposition Stranding and related phenomena in Russian. In *Studies in formal Slavic Phonology, Morphology, Syntax, Semantics and Information Structure*, ed. G Zybatow, U Junghanns, D Lenertová, and P Biskup. Frankfurt am Main: Peter Lang.

Rackowski, Andrea. 2002. *The structure of Tagalog: Specificity, voice, and the distribution of arguments*. Doctoral dissertation, MIT.

Rackowski, Andrea, and Norvin Richards. 2005. Phase Edge and Extraction: A Tagalog Case Study. *Linguistic Inquiry* 36. Https://doi.org/10.1162/002438905774464368.

Reinhart, Tanya. 2006. *Interface Strategies*. The MIT Press, Cambridge, MA.

Richards, Norvin. 1997. *What moves where in which language?*. Doctoral dissertation, MIT.

Richards, Norvin. 1999. Featural cyclicity and ordering of multiple specifiers. In *Working minimalism*, ed. Samuel D. Epstein and Norbert Hornstein, 127–158. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar: A handbook of generative syntax*, 281–337. Dordrecht: Kluwer.

Rudin, Catherine. 1988. On Multiple Questions and Multiple WH Fronting. *Natural Language & Linguistic Theory* 6.

Sabbagh, Joseph. 2007. Ordering and linearizing rightward movement. *Natural Language & Linguistic Theory* 25. Https://doi.org/10.1007/s11049-006-9011-8.

Sauerland, Uli. 2003. Intermediate adjunction with A-movement. *Linguistic Inquiry* 34.

Stroik, Thomas. 2009. *Locality in Minimalist Syntax*. Cambridge, MA: MIT Press. Linguistic Inquiry Monographs. Vol. 51.

Torrence, Harold. 2012. The morpho-syntax of silent wh-expressions in Wolof. *Natural Language & Linguistic Theory* 30:1147–1184.

Torrence, Harold. 2018. *Left Edge Agreeing Elements inWolof Wh-Questions*. MS, UCLA.

Ura, Hiroyuki. 1996. *Multiple feature-checking: A theory of grammatical function splitting*. Doctoral dissertation, MIT.

Urban, Emily. 1999. *Exactly stranding*. Ms., University of California, Santa Cruz.

van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A case study of Dinka Bor*. Doctoral dissertation, MIT.

van Urk, Coppe, and Norvin Richards. 2015. Two Components of Long-Distance Extraction: Successive Cyclicity in Dinka. *Linguistic Inquiry* 46:113–155.

Wiland, Bartosz. 2010. Overt Evidence from Left-Branch Extraction in Polish for Punctuated Paths. *Linguistic Inquiry* 41. Https://doi.org/10.1162/ling.2010.41.2.335.

Zwart, Jan-Wouter. 1997. Transitive expletive constructions and the evidence supporting the multiple specifier hypothesis. In *German: Syntactic problems*, ed. W Abraham and van Gelderen, 105–134. Niemeyer: Tubingen.

Zyman, Erik. To appear. *Phase Constrained Obligatory Late Adjunction*. Syntax.