The nature of overlapping A-bar chains as revealed by parasitic gaps[*]

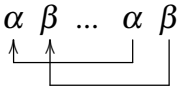
Colin Davis

Massachusetts Institute of Technology

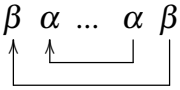
1. Introduction

In this paper, I investigate some properties of sentences in which multiple movement paths overlap. Scenarios of this variety minimally involve two moving phrases, for which there are two possible outcomes. Either the paths of the two moving phrases will cross, resulting in no change in their relative order (1), or nest, reversing their order (2).

(1) *Crossing paths (Order-preserving)*



(2) *Nesting paths (Order-reversing)*



Both of these movement patterns are attested in human language (see Richards 1997 for numerous examples). In English, overt overlapping A′-movements always nest, as described by the Path Containment Condition (PCC; Pesetsky 1982). The least complex way to form overt overlapping paths in English is to combine topic/focus movement and *wh*-movement, in which case the topicalized/focused phrase precedes the *wh*-phrase (3):

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

This paper's first goal is to examine successive cyclicity in overlapping movement configurations like that in (3), and beyond. If recent work in phase theory (Chomsky 2000, 2001, and many following) is correct that there is a clause internal phase vP, which movement from must pass through the edge of, then the moving phrases in (3) must form intermediate landing sites in the vP edge. Since there are two moving phrases here, there are

[*]Thanks to Kenyon Branan, Justin Colley, Patrick Elliott, Michael Yoshitaka Erlewine, Danny Fox, Naomi Francis, Sabine Iatridou, Laura Kalin, Martia Martinović, David Pesetsky, Norvin Richards, the audiences of NELS 50, WCCFL 38, various events at MIT, and to the people I compelled to judge the complex sentences investigated here. For more detail on the topics discussed here, see chapter 5 of Davis (2020b).

two conceivable ways in which this can occur. That is, as shown in the simplified diagrams in (4) below, these phrases' movements to the left periphery of the clause could in principle be preceded by either the topic/focus phrase forming an outer spec-vP, with the *wh*-phrase forming an inner one (4a), or the reverse (4b).

(4) *Two conceivable intermediate orders in the vP edge for (3)*

a. [$_{CP}$ *this book*$_2$ *who*$_1$ ... S [$_{vP}$ ***this book***$_2$ ***who***$_1$ v-V ... *who*$_1$ *this book*$_2$]]
b. [$_{CP}$ *this book*$_2$ *who*$_1$ ... S [$_{vP}$ ***who***$_1$ ***this book***$_2$ v-V ... *who*$_1$ *this book*$_2$]]

I argue that the behavior of parasitic gaps (PGs; Engdahl 1983, a.o.), when considered in the context of certain predictions from Nissenbaum (2000), reveals that a derivation like (4a) is correct for (3) and similar sentences. Notice that in (4a) the A′-moved phrase with the higher surface position (*this book*$_2$) was also the higher specifier formed by successive cyclic movement through the vP. This result is a manifestation of the following generalization, which I defend in this paper:

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

I argue that this result is evidence for the *Cyclic Linearization* theory of phases (CL; Fox and Pesetsky 2005a,b, Ko 2014, Davis 2019, 2020a, a.o.), which predicts that the linear orderings phase-by-phase spell-out establishes must be consistent throughout a derivation.

This paper's second goal is to clarify the nature of the PCC, a phenomenon closely related to considerations about multiple specifier formation. The sentence in (3) above obeys the PCC because its moving phrases form nesting paths. Even if the surface order topic/focus $<$ WH is maintained, it is necessary for the moved phrases' final order to be the reverse of their original order in this context, and similar ones:

(6) *Forced reversal of original order*

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]

If the generalization in (5) above is correct, then the final reversed order in PCC-obeying sentences must have been derived immediately within vP, since final orders and intermediate orders must match. While I argue that this is indeed so, it is also necessary to understand what forces this order-reversal to occur. Richards (1997) hypothesizes that PCC effects are correlated with the possibility of there being only one overt specifier of CP in a given language, and hence, an absence of the PCC in languages where multiple overt specifiers of CP are permitted. Extending this hypothesis, I argue that the PCC emerges from the CL theory 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.).

2. Background: On the ordering of multiple specifiers

Richards (1997, 1999) and much work following argues that when two phrases move to specifiers of one head, the closest phrase is attracted first (7a), while the lower one which is attracted second "tucks-in" to a specifier below the one formed by the first instance of movement (7b). This is due to an economy constraint favoring shorter movement paths.

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

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

b. $[_{XP}\ \alpha\ \beta\ \mathrm{X}\ \ldots\ [_{YP}\ \ldots\ t_{\alpha}\ t_{\beta}\ ]]$

Consider what this theory predicts for successive-cyclic movement through vP in the multiple A′-movement scenario we saw in (3) above, repeated once more below:

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

First, the initially higher phrase *who*$_1$ should move to spec-vP:[1]

(9) *Higher phrase forms first spec-vP*
$[_{vP}$ *who*$_1$ S v-V $[_{VP}$ t_1 *this book*$_2$]]

Second, the lower phrase *this book*$_2$ should tuck-in to a lower spec-vP below *who*$_1$:

(10) *Lower phrase forms second, lower spec-vP*
$[_{vP}$ *who*$_1$ *this book*$_2$ S v-V $[_{VP}$ t_1 t_2]]

After this tucking-in occurs within the vP, these moving phrases would then each continue on to their landing sites in the left periphery:

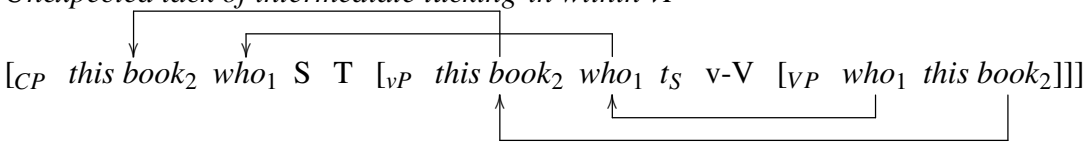
(11) *Movement to left periphery after tucking-in within vP*

$[_{CP}$ *this book*$_2$ *who*$_1$ S T $[_{vP}$ *who*$_1$ *this book*$_2$ t_S v-V $[_{VP}$ *who*$_1$ *this book*$_2$]]]

In contrast, the tucking-in theory rejects the possibility of a derivation in which, at the vP level, *this book*$_2$ moves above *who*$_1$ rather than tucking-in beneath it:

[1] Following Nissenbaum (2000), I represent the in situ subject below specifiers of (non-subject) A′-movement within vP. This order of constituents in the vP is also predicted by the CL approach that this paper will argue for, since such a configuration keeps the ordering of the A′-moved phrase(s) and the subject consistent both within vP and after these phrases' later movement from it.

(12) *Unexpected lack of intermediate tucking-in within vP*



As previewed above, I will argue that a derivation without tucking-in in vP is correct for sentences of this variety. This result reveals that multiple movements to one head do not always require crossing paths. In response to this finding, I posit that multiple specifiers can be ordered freely outside of certain specific contexts, as discussed at the end of this paper.

3. Phase theory and CL

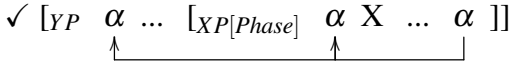
Chomsky (2000, 2001) argues that phrasal movement must pass in successive-cyclic fashion through the edge of certain constituents termed *phases* (including at least vP and CP). For Chomsky, the phase edge is an escape hatch for movement because it is not subject to phase-level spell-out, which targets only the phase head's complement. After spell-out, the complement is impenetrable to further syntactic operations. Thus any material moving out of a phase's complement must get to the edge of the phase before spell-out applies. Otherwise, it will be trapped in the complement, and unavailable for further movement.

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 requires the concept that spell-out does not make constituents impenetrable, because 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*:

(13) *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 that end up with contradictory linearization instructions. Therefore syntax must be able to form configurations that end up with linearization information that is consistent across all phases within the derivation in question. Otherwise, there will be a crash at PF. Exiting a phase by moving out via its linear edge is one way to keep the linearization for a given derivation contradiction-free:

(14) *Successive-cyclic movement via linear edge of phase*



By exiting via the linear edge of each phase passed, phase-exiting phrases are determined by PF to precede the content of each phase in question. This is ultimately consistent with a final representation where the moved material precedes all phases that it has exited.

Consider what CL predicts for a sentence where multiple overt A′-movements pass through the vP, as in (8) above. Notice that in this sentence, the phrase that has undergone topic/focus movement occupies a position in the left periphery higher than, and thus left of, the moved *wh*-phrase. The PCC also holds here, as (6) showed, though leave this fact aside for the moment. Since in (8) the topicalized/focused phrase has a final position above the *wh*-phrase, the former phrase should also have been the higher of the two when these phrases formed specifiers of vP. Otherwise, this sentence would involve a linearization contradiction. In the derivation in (11) above that Richards' theory of multiple specifier formation predicts, the vP internal order and final order for these phrases do not match. In contrast, in the derivation in (12) above that I will argue is revealed by the behavior of PGs in such sentences, the moved phrases' final and intermediate orders do match. This is as described in the generalization in (5) above, which the CL approach to phases predicts.

4. PG evidence for matching of final and intermediate orders

In this section, I provide the relevant evidence from PG licensing. PGs are gaps inside of islands which take on the interpretation of a phrase undergoing A′-movement external to and structurally across that island, among various other constraints (Culicover and Postal 2001). An initial example of a PG is provided in (15) below:

(15) *PG in adjunct island licensed by wh-movement*
Who$_1$ should we talk to t_1 [before buying a present for PG$_1$]?

The above example shows a PG licensed by *wh*-movement. PGs can be licensed by any A′-movement, including topic/focus movement:

(16) *PG in adjunct island licensed by topic/focus movement*
This book$_2$, we should really read t_2 [before citing PG$_2$ in our paper].

We've seen that topic/focus movement and *wh*-movement can co-occur in one clause. In this context, only the moved phrase with the higher surface position can license a PG, despite both of these A′-movements being capable of PG licensing when alone (17):

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

While combining topic/focus movement and *wh*-movement within one clause is not the only way to yield overlapping overt A′-chains in English, it is the simplest one. The alter-

native is to use a multi-clausal derivation, which contains multiple clause 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 (18-20) below, each of which includes a sentential adjunct with a PG. We see here that, as in (17) above, only the outermost of the two overtly A′-moved phrases is able to license the PG:

(18) a. *Successful PG licensing by higher moved phrase*
Tell me [what article]$_2$ she said [who$_1$ I should talk to t_1 about t_2 [before reviewing PG$_2$]]?

b. *Failed PG licensing by lower moved phrase*
*Tell me [what article]$_2$ she said [who$_1$ I should talk to t_1 about t_2 [after introducing myself to PG$_1$]]?

(19) a. *Successful PG licensing by higher moved phrase*
[Which 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. *Failed PG licensing by lower moved phrase*
*[Which car]$_2$ do you know [who$_1$ to persuade t_1 to buy t_2 [after having a few drinks with PG$_1$]]?

(20) a. *Successful PG licensing by higher moved phrase*
This is the book [which$_2$ I asked [who$_1$ you gave t_1 comments on t_2 [after making yourself a copy of PG$_2$]].

b. *Failed PG licensing by lower moved phrase*
*This is the book [which$_2$ I asked [who$_1$ you gave t_1 comments on t_2 [after having a discussion with PG$_1$]].

The above examples show contexts where both moving phrases originate in the same clause. The same PG asymmetry holds for moving phrases that originate in separate clauses:

(21) a. *Successful PG licensing by higher moved phrase*
These cookies$_2$, who$_1$ did you tell t_1 [that you adore t_2] [while buying a package of PG$_2$]?

b. *Failed PG licensing by lower moved phrase*
*These cookies$_2$, who$_1$ did you tell t_1 [that you adore t_2] [while having a conversation with PG$_1$]?

(22) a. *Successful PG licensing by higher moved phrase*
These are the beans [which$_2$ I know [who$_1$ you told t_1 [that you revile t_2] [after eating an expired can of PG$_2$]]].

b. *Failed PG licensing by lower moved phrase*
*These are the beans [which$_2$ I know [who$_1$ you told t_1 [that you revile t_2] [after running into PG$_1$ in the hallway]]].

(23) a. *Successful PG licensing by higher moved phrase*
[That book]$_2$, who$_1$ did you tell t_1 [that you hate t_2] [while burning a copy of PG$_2$ in the office yesterday]?

b. *Failed PG licensing by lower moved phrase*
*[That book]$_2$, who$_1$ did you tell t_1 [that you hate t_2] [while discussing politics with PG$_1$]?

The pervasive arising of this PG licensing asymmetry in overlapping movement configurations informs us about their corresponding derivation, given certain predictions from Nissenbaum (2000). Nissenbaum argues that the licensing of PGs in sentential adjuncts depends on the licensing phrase successive-cyclically moving through vP, and builds a theory of PGs which makes a more general prediction about vPs with multiple specifiers:

(24) *Multiple Specifier Single Parasitic Gap Prediction*
In multiple specifier constructions in which 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 that vP.[2]

The examples in (17-23) above all involve merge of the PG-containing adjunct to the first vP in which both moving phrases were present. The fact that the higher phrase on the surface in these sentences is always the only available licenser for the PG indicates, given the above prediction, that this phrase was also the outer specifier formed by movement from the relevant vP. This is exactly as CL predicts: whatever order these phrases have on the surface must be matched as soon as those phrases occupy the same phase and are thus linearized with respect to each other.[3]

[2]This statement of the prediction is taken from Fox and Nissenbaum (2018), who use it to analyze certain constructions involving pied-piping that are beyond the scope of this paper. While space constraints prevent fully explaining this prediction here, in brief, Nissenbaum argues that the semantic status of a vP targeted for two instances of successive cyclic movement permits a sentential adjunct containing one PG to be interpreted only when merged to the node of vP that is the sister of its outermost specifier. As a result, only the outer specifier is able to license the PG. Hence, whichever moved phrase successfully licenses a lone PG in an overlapping movement derivation must have been the outer specifier of the vP in that context.

[3]The sentences in (17-23) show PGs inside of adjunct islands. However, the important PG licensing asymmetry shown in this environment also holds for PGs in subjects. This fact was observed by Pesetsky (1982) by way of the following contrast:

(i) 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)

(25) *Higher surface position → higher position in vP edge → can license PG*
[This book]$_2$, who$_1$ should we [$_{vP}$ $\boxed{t_2}$ t_1 talk to [$_{CP}$ t_1 about t_2 before commenting on $\boxed{\text{PG}_2}$]]]

5. Extension: The Path Containment Condition

So far, we have seen evidence that the final relative order of two successive-cyclically moving phrases is matched in their intermediate landing sites. I argue that this is so because, given CL, PF rejects derivations that do not ultimately have consistent linearization information across all phases. However, CL does not by itself clarify the closely related fact that derivations with overlapping overt A′-chains in English must obey the PCC:

(26) *PCC: Order of landing sites must be reversed relative to origination sites*

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?

I argue that combining CL with an independent proposal about covert movement in English clarifies the source of the PCC.

Recall Richards' proposal that non-initial instances of multiple movement to one head tuck-in. Richards' evidence for this is based on languages like Bulgarian, where all *wh*-phrases in a multiple *wh*-question overtly move to specifiers of CP, as in (27a) below. For Richards and for Nissenbaum, what distinguishes (27a) from its English equivalent in (27b) is that for English, tucking-in to a lower specifier yields covert movement:

(27) *Overt versus covert tucking-in in multiple wh-questions*

a. *Bulgarian: Overt*
Kogo$_1$ **kavo**$_2$ e pital Ivan t_1 t_2?
whom what AUX asked Ivan

b. *English: Covert*
Who$_1$ $\varnothing_2$ did Ivan ask t_1 **what**$_2$?

Combining this view of tucking-in in English with CL reveals a reason why order-reversal is necessary for overlapping overt A′-chains in this language.

Consider (28) below, where the phrases α and β move to the vP edge with crossing paths. Here β tucks-in below α. If tucking-in movement in English is covert, then β will be linearized at the tail of its movement chain, at the right edge of this vP:

This is precisely what we expect if Nissenbaum (2000) is correct that the configuration which permits a PG in a subject is the same as that involved with a PG in a sentential adjunct. Though certain details of semantic composition are more complex for PG-bearing subjects, the relevant interaction between the PG-containing domain and successive cyclic movement through vP is exactly the same in either case.

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

Since α occupies the left linear edge of the vP in (28), CL allows α to overtly move on, as in (29). But we see in (29) that the same does not apply to β. Because β did not overtly reach the left edge of vP in (28), its further leftward overt movement as in (29) yields linearization information that contradicts what was established in (28). In particular, the ordering of (28) demands that β follow the (rest of the) material originating in vP, while (29) requires that β precede that same material. This is a contradiction:

(29) *Later overt movement of α and β from (28)*
$[\alpha \ldots \beta \ldots [_{TP}\ \text{S}\ [_{vP}\ t_{\alpha}\ t_{\beta} \ldots \text{v-V} \ldots t_{\alpha}\ t_{\beta}\ ]]]$
Linearization: $= \alpha < \beta < S < V$ (contradicts order in (28) above)

This contradiction is avoided if instead of tucking-in below α as in (28), β moves to a specifier above it within vP, as in (30). In (30) neither β nor α are treated by PF as having covertly moved, since neither tucked-in.[4] From here, further overt movement leftward of β and α is licit, since they both occupy the left linear edge of the vP.

(30) *Not tucking-in within vP yields two overt movements to vP edge*
$[_{vP}\ \ \beta\ \ \alpha\ \ \text{S}\ \ \text{v-V}\ \ [_{VP}\ \ t_{\alpha}\ \ t_{\beta}\ ]]$
Linearization: $= \beta < \alpha < \text{S} < \text{V}$

Notice that by avoiding tucking-in, these moving phrases end up reversing their order within vP. Provided that their subsequent landing sites preserve their reversed ordering established within vP, the result is a CL-satisfying derivation fitting what the PCC describes as licit, as we see in (31) below. Thus I argue that the PCC holds because avoidance of tucking-in within vP causes two overtly A′-moving phrases to end up in reversed order.[5]

[4]That is, I propose that α in (30) is not treated as having moved covertly because, despite ultimately being an inner specifier of vP, α did not come to occupy this position due to tucking-in. This is in contrast to a proposal under which inner specifiers are covert regardless of the manner in which they are derived.

[5]CL requires the order of moving phrases derived post-reversal to be maintained for the rest of the derivation. Thus it is necessary for two phrases overtly A′-moving from an embedded CP after reversing order within vP, for instance, to move with crossing paths to the edge of that CP before exiting it, as in (i). Doing otherwise would reverse the order of these phrases once more, causing an ordering contradiction:

(i) [This book]$_2$, who$_1$ did you say $[_{CP}\ t_2\ t_1$ that we should $[_{vP}\ t_2\ t_1$ talk to t_1 about $t_2]]$?

Such overt tucking-in in the CP edge would be an exception to the covertness of tucking-in discussed earlier in this section. I hypothesize that while tucking-in is indeed covert in the basic case, this PF rule is overridden by a need to maintain the covertness/overtness of a movement chain for its entirety. This can be understood as an extension of the following principle hypothesized by Fox and Nissenbaum (2018):

(31) *PCC obeyed: Final positions reverse original order*
✓ [XP_2 ... YP_1 ... [$_{vP}$ t_2 t_1 ... t_1 t_2]]

I hypothesize that crossing and nesting multiple specifier formation are both possible for overlapping movement configurations, provided that an independent factor doesn't further constrain the derivation. For the abstract scenario just examined, which is representative of English, I propose that the requirements of linearization force a nesting derivation for overlapping overt A′-movement, despite a crossing derivation presumably being possible in principle if the needs of PF were set aside. There are circumstances where crossing paths are required, however, which I will briefly address at the end of this section.

5.1 When the PCC does not apply

Richards (1997) shows that the PCC is not obeyed in Bulgarian sentences involving multiple *wh*-movements to separate CP edges, unlike analogous English sentences:

(32) *Anti-PCC effect in Bulgarian*

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

b. * **Kogo**$_2$ se opitvat da razberat [$_{CP}$ **koj**$_1$ e t_1 ubil t_2]]?
whom SELF try to find.out who AUX killed
(Richards 1997: 41 (41))

The above analysis of the PCC in English correctly predicts that Bulgarian should be capable of circumventing this constraint, since as we saw in the multiple *wh*-question in (27a) above, tucking-in in Bulgarian is overt, unlike in English (27b). However, I have hypothesized that crossing versus nesting paths are both freely available for multiple specifier formation, unless one option is ruled out due to independent factors, such as the needs of linearization. Under this view, it is unclear why in the Bulgarian (32), a PCC-style derivation is not only unnecessary (32a) but banned (32b).

I speculate that this result may stem from the fact that (32) involves two interrogative *wh*-phrases, since when the two *wh*-phrases move for separate reasons, either final order appears to be possible, as the following examples from Rudin (1988) show:[6]

(ii) *Early Determination:* If a first step in successive cyclic movement is marked as covert, a second step is going to be covert, even if the position targeted is one that would otherwise be overt.

I suggest that this principle applies not only to covert movement, but to movement in general, preserving the overtness or covertness of any movement chain until it arrives at its terminal position. This allows a phrase to move overtly within one phase and then continue to move overtly despite tucking-in within the next phase.

[6]Another possibility is that something independent mandates the ordering in (32a), following Billings and Rudin (1996), who show that the order of *wh*-phrases in Bulgarian is governed by several constraints relating to animacy, case, and other factors. It is unclear how general such constraints truly are, however.

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

a. *Order-preserving relativization with interrogative wh-movement*

Čoveka [**kojto**$_1$ ne znaeš [**kakvo**$_2$ kazvat [če e kupil t_1 t_2]]]
man who.REL NEG know.2S what say.3P that has bought
'The man who you don't know what they say that he bought'
(Adapted from Rudin 1988 ex. 8c)

b. *Order-reversing relativization with interrogative wh-movement*

Edna kniga [**kojato**$_1$ se čudja [koj$_2$ t_2 znae [**koj**$_3$ t_3 prodava t_1]]]
a book which wonder-1S who knows who sells
'A book which I wonder who knows who sells (it)'
(Adapted from Rudin 1988 ex. 19)[7]

Rudin shows that Romanian has the same *wh*-movement properties as Bulgarian, for instance, in requiring multiple overt *wh*-movement to spec-CP in multiple *wh*-questions (34):

(34) *Romanian multiple wh-fronting*

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?'
(Adapted from Rudin 1988, ex. 9a)

As expected if Romanian is like Bulgarian and unlike English in allowing overt tucking-in, multiple distinct A′-movements in Romanian can freely cross or nest, as in (35) below:

(35) *Freedom of ordering for relativization + interrogative wh-movement in Romanian*

a. 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"
(Adapted from Comorovski 1986, ex. 4b)

b. ?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"
(Donca Steriade, p.c.)

[7]Example (33b) shows relativization crossing not one but two interrogative *wh*-phrases. While the crossing of just one such phrase would be enough for the point here, Rudin chose to include more than one to emphasize the unboundedness of relativization in this particular configuration.

5.2 When multiple specifier order is rigid

In English as well as Bulgarian and Romanian, two *wh*-phrases in a multiple *wh*-question must move with crossing paths, as is clearest in the latter two languages where all such movements are overt. This is shown for Bulgarian in (36) below:

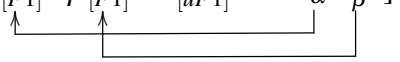
(36) *Obligatory crossing paths: Bulgarian multiple wh-question*

a. Kogo$_1$ kakvo$_2$ e pital Ivan t_1 t_2?
who.NOM what.ACC AUX asked Ivan
'What did Ivan ask who?'

b. * Kakvo$_2$ kogo$_1$ e pital Ivan t_1 t_2?
what.ACC who.NOM AUX asked Ivan
'What did Ivan ask who?'
(Richards 1997, p. 277 ex. 75)

This result is not compatible with the above hypothesis that the order of multiple specifiers is in principle free. To resolve this tension, I hypothesize that crossing paths are required in contexts where multiple movements to one head are triggered by the same attracting feature, as schematized in (37) below.

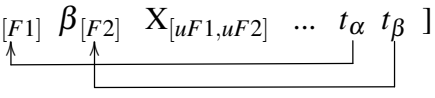
(37) *One attracting feature for multiple goals → tucking-in*



I thus propose that in multiple *wh*-questions, all *wh*-phrases are attracted by one *wh*-sensitive feature on C (and a corresponding one on v), and that these phrases thus move with obligatorily crossing paths. The same analysis (though involving a different attracting feature in each context) can be applied to the other obligatory tucking-in scenarios reported in Richards (1997). These include but are not limited to local A-scrambling in Japanese, negative fronting in Bulgarian, multiple object shift in Germanic languages, and multiple clitic clustering in Tagalog and Serbo-Croatian.

Additionally, I propose that when a head bears multiple movement-triggering features, the goals respectively attracted by those features may form crossing or nesting paths:[8]

(38) *Attraction by separate features on one head → variability*

a. 

[8]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, Doggett 2004, Rackowski and Richards 2005), many of which assume that tucking-in cannot occur in this situation. Here I maintain that tucking-in is possible in this context, but not required.

b. $[_{XP}\ \beta_{[F2]}\ \ \alpha_{[F1]}\ X_{[uF1,uF2]}\ \ ...\ \ t_\alpha\ \ t_\beta\ ]$

This proposal can accommodate the freedom of ordering in the Bulgarian and Romanian scenarios respectively shown in (33) and (35) above, as well as the obligatory nesting paths we saw in the English (17-23). All of these examples involve overlapping A′-movements of different types, or at least to different positions, such that they are plausibly triggered by distinct features. While in Bulgarian and Romanian we end up with freedom of ordering in these contexts, in English the covertness of tucked-in specifiers rules out a crossing paths derivation of (17-23) for linearization reasons, as argued in the previous section.

At least one further issue remains unresolved, however: Nothing I have said here explains why obligatory tucking-in contexts must involve only a single movement-triggering feature. For instance, we can conceive of a multiple *wh*-question which involves multiple attracting features and thus allows the formation of nesting paths, though this is evidently impossible for *wh*-questions in reality. See Davis (2020b) for an investigation of what restricts the distribution of attracting features.

6. Conclusion

In this paper, I examined the nature of derivations with overlapping A′-movement paths. Using certain predictions 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 following generalization, which CL predicts:

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

I went on to argue that the CL theory predicts the arising of the PCC in English when combined with the concept that tucked-in specifiers are covert in this language, and consequently also predicts the absence of the PCC in the multiple *wh*-fronting languages Bulgarian and Romanian. Additionally, I proposed that crossing paths are only required for multiple movements to one head triggered by the same attracting feature. More thoroughly evaluating and refining this hypothesis must be left to future work.

References

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*, ed. by Jindřich Toman, 1–24.

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

Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1–52. Cambridge: MIT Press.

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

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

Davis, Colin. 2019. Possessor Extraction in Colloquial English: Evidence for Successive-Cyclicity and Cyclic Linearization. *Linguistic Inquiry.* Early access article.

Davis, Colin. 2020a. Crossing and Stranding at Edges: Intermediate Stranding and Phase Theory. *Glossa.* 5:1–32.

Davis, Colin. 2020b. *The Linear Limitations of Syntactic Derivations*. Doctoral Dissertation, MIT.

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

Engdahl, Elisabet. 1983. Parasitic Gaps. *Linguistics and Philosophy* 6:5–34.

Fox, Danny, and John Nissenbaum. 2018. *Pied-piping and covert movement: Evidence from parasitic gap licensing*. Talk given at WAFL 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* 31:235–262.

Fox, Danny, and David Pesetsky. 2005b. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31:1–48.

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

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

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

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

Rackowski, Andrea, and Norvin Richards. 2005. Phase Edge and Extraction: A Tagalog Case Study. *Linguistic Inquiry* 36:565–599.

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. by Samuel D. Epstein and Norbert Hornstein, 127–158. Cambridge: MIT Press.

Rudin, Catherine. 1988. On Multiple Questions and Multiple WH Fronting. *Natural Language & Linguistic Theory* 6:445–501.

Colin Davis
colind@{mit.edu/usc.edu}