THE NATURE OF OVERLAPPING A′-CHAINS AS REVEALED BY PARASITIC GAPS **(Colin Davis, MIT)**

1 INTRODUCTION: The intricate properties of parasitic gaps (PGs) have long enriched research on the syntax of (A′) movement (Engdahl 1983, Nissenbaum 2000, Legate 2003, Overfelt 2015, Branan 2017, Kotek & Erlewine 2018, Fox & Nissenbaum 2018). In this work, I use PG licensing in English to reveal a novel generalization about derivations where two A′-movement chains pass through vP, stated in (1).

(1) *Generalization: If XP_1 and XP_2 move to become A′-specifiers of vP such that XP_1 c-commands XP_2, then the final surface position of XP_1 c-commands the final surface position of XP_2.*

While (1) is surprising for some theories of movement, I argue that it is a natural consequence of *Cyclic Linearization* (CL; Fox & Pesetsky 2005, a.o.), which predicts that the relative order established for the constituents of a given phase must be preserved throughout the derivation. I also show that CL interacts with the distribution of covert movement to yield Pesetsky's (1982) *Path Containment Condition* (PCC).

Data preview: The argument for (1) arises from the PG licensing asymmetry in sentences like (2), and many other contexts where A′-movement chains overlap. While both topicalization and *wh*-movement can normally license a PG, when these movements are combined in one clause as in (2), only the outermost A′-moved phrase can serve as the licenser. As we'll see, building on Nissenbaum (2000), this asymmetry indicates that the outer A′-moved phrase was also the outer A′-specifier of vP, as (1) describes.

(2) [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$_1$?

2 BACKGROUND ON SPECIFIER ORDER AND PGS: Richards (1997, 2001, a.o.) argues that when two phrases move to become specifiers of one head, the closest phrase moves first (3a), while the lower phrase moves second, "tucking-in" to a lower specifier below that formed by the prior movement (3b). Richards argues that tucking-in is motivated by *Shortest Move*, which keeps movement paths as short as possible:

(3) a. $[_{XP}\ \underset{\curvearrowright}{\alpha}\ \text{X}\ \ldots\ [_{YP}\ \ldots\ t_\alpha\ \beta\ ]]$ b. $[_{XP}\ \alpha\ \underset{\curvearrowright}{\beta}\ \text{X}\ \ldots\ [_{YP}\ \ldots\ t_\alpha\ t_\beta\ ]]$ (*Tucking-in*)

Building on Richards, Nissenbaum (2000) argues that PG licensing depends on successive-cyclic movement through the vP phase, and further, that certain constraints on PG licensing emerge because covert movement involves tucking-in. I argue that Nissenbaum's work makes a prediction (4) which reveals (1):

(4) *Prediction: A single PG in a clausal adjunct of vP can only be licensed by the outermost A′-specifier of that vP (but a lower A′-specifier of vP can license an additional PG if present).*

The puzzling properties of (2): If movement must pass through the edge of vP (Chomsky 2000, 2001, a.o.), then Richard's theory expects the following for (2): the initially higher moved phrase *who*$_1$ should have moved to spec-vP (5a), with the lower phrase *this book*$_2$ tucking-in to a lower spec-vP. However, given (4), the fact that *this book*$_2$ licenses the PG in (2) indicates that *this book*$_2$ actually moved to the outermost spec-vP (5b). (Following Nissenbaum, I represent the in situ subject below specifiers of A′-movement.)

(5) a. $[_{vP}\ \underset{\curvearrowright}{\textit{who}_1}\ \text{S}\ \text{v-V}\ [_{VP}\ t_1\ \textit{this book}_2\ ]]$ b. $[_{vP}\ \textit{this book}_2\ \textit{who}_1\ \text{S}\ \text{v-V}\ [_{VP}\ t_1\ t_2\ ]]$

This result fits (1), since both in vP and in their final position, *this book*$_2$ c-commands *who*$_1$. This is a puzzle for tucking-in theory. A related puzzle is that the order which these phrases had in their original position must be reversed in their final position, as Pesetsky's (1982) *Path Containment Condition* (PCC) describes:

(6) * [This student]$_1$, what$_2$ should we talk to t_1 about t_2? (*No order reversal, contra (2) above*)

3 A THEORY OF (1) AND THE PCC: I argue that a linearization problem posed by the covertness of tucking-in in English forces two phrases overtly A′-moving from vP to reverse order within vP (5b), rather than forming a tucking-in configuration. CL then preserves that reversed order for the rest of the derivation.

Tool #1 - Cyclic Linearization (CL): CL posits that movement is constrained by coherency of linearization. For CL, spellout linearizes whole phases. Once a phase's constituents are linearized, they may move on. However, their later positions must preserve the relative order decided at their earlier spellout, to avoid a contradiction at PF. This theory predicts that overt phrasal movement must exit a phase via its linear edge. Similarly, if a phrase covertly moves within a phase, and then overtly moves out later, a PF crash results.

Tool #2 - Covert tucking-in in English: Richards argues for tucking-in in part based on languages like Bulgarian, where all *wh*-phrases in a question overtly move to the left periphery (7a). For Richards and for

Nissenbaum, what distinguishes (7a) from its English equivalent in (7b) is that for English, tucking-in to a lower specifier yields covert movement, which is linearized at the tail of its chain. I adopt this view as well.

(7) a. **Kogo**$_1$ **kavo**$_2$ e pital Ivan t_1 t_2? b. **Who**$_1$ $\varnothing_2$ did Ivan ask t_1 **what**$_2$?
whom what AUX asked Ivan

4 WHEN TUCKING-IN FAILS: I take vP and CP to be phases (Chomsky 2000, 2001, a.o.) and hence domains where CL applies (Ko 2014, a.o.). If the originally lower moved phrase *this book*$_2$ in (2) tucked-in at vP, then when the vP is linearized, English linearization rules would regard this as covert movement. Thus as (8) shows, *this book*$_2$ would be linearized in its base position. (Note that "X $<$ Y" = "X precedes Y".)

(8) [$_{vP}$ *who*$_1$ $\varnothing_{\textit{this book}}$ S v-V [$_{VP}$ t_1 *this book*$_2$]] (= *who* $<$ S $<$ V $<$ *this book*)

Later in the derivation, I assume following Rizzi's (1997 a.o.) CP structure that *who*$_1$ moves to spec-Foc(us)P, and *this book*$_2$ to spec-Top(ic)P, where their movement ends. Since these moved phrases aren't tucked-in at their final positions in the left periphery, they are linearized as having overtly moved, as in (9):

(9) [$_{TopP}$ *this book*$_2$ [$_{FocP}$ *who*$_1$ [$_{TP}$ S vP ...]]] (= *this book* $<$ *who* $<$ S $<$ [content of vP])

Importantly, *this book*$_2$ can't be pronounced both linearly right of the material originating in vP, as (8) established, and in the clause's left edge, as (9) requires. Thus tucking-in would cause a PF crash. However, if in (8) *this book*$_2$ instead moves to the outermost spec-vP (5b), then both it and *who*$_1$ are linearized in the left edge of vP, **since neither tucks-in**. From here, their further overt movement leftward is licit under CL:

(10) [$_{TopP}$ [this book]$_2$ [$_{FocP}$ who$_1$ should [$_{TP}$ we [$_{vP}$ t_2 t_1 t_S talk to t_1 about t_2]]]]] (✓*No tucking-in*)

5 DERIVING THE PCC: We've just seen why tucking-in must be avoided when two phrases overtly A′-move from vP. By avoiding tucking-in, those phrases reverse order in vP. If their reversed order in vP matches the order of their final positions, this satisfies CL, and yields a structure fitting what the PCC describes as licit (2/10/11a). But if the phrases' original positions are such that their reversed order in vP doesn't match that of their final positions, the derivation violates CL, and fits what the PCC describes as illicit (6/11b):

(11) a. ✓ [XP$_2$ YP$_1$ ... [$_{vP}$ t_2 t_1 ... t_1 t_2]] b. * [XP$_2$ YP$_1$ ... [$_{vP}$ t_1 t_2 ... t_2 t_1]]

The same holds for (12) below, where two *wh*-phrases must reverse order and land in separate CPs. CL predicts that this reversal was derived in the lowest vP, and preserved afterwards. Notice that the PG-containing adjunct in (12) has an anaphor forcing it to adjoin in the lower clause due to Principle A. The outer moved phrase *what book*$_2$ licenses the PG, showing that it was indeed the lower vP's outer specifier:

(12) [**What book**]$_2$ did she wonder **who**$_1$ you should [$_{vP}$ t_2 t_1 ask t_1 about t_2 ...
a. [✓before making yourself a copy of PG$_2$]] / b. [*after introducing yourself to PG$_1$]]

Deriving anti-PCC effects: This account correctly predicts no PCC effect in English multiple questions like (7b), where two *wh*-phrases target the same CP. Since in (7b) the lower *wh*-phrase ultimately tucks-in covertly in CP, it also can do so in vP. This account also predicts Richards' finding that Bulgarian sentences analogous to (12) show no PCC effect (13). Since tucking-in in Bulgarian yields overt movement (7a), there is no reason to avoid tucking-in at vP, hence the two moving phrases don't need to reverse their order:

(13) **Koj**$_1$ se opitvat da razberat [$_{CP}$ **kogo**$_2$ C e t_1 ubil t_2]]?
who SELF try to find.out whom AUX killed (*Anti-PCC in Bulgarian*)

6 CONSEQUENCES: If the lower moved phrase on the surface in (2/12) was the lower A′-specifier of vP as (1) states, then following (4), it should be able to license a second PG. I show that this is correct. I also show that this account accurately predicts *surfing* but not *diving* movement patterns in English (Sauerland 1999).

Generally, these findings raise questions about locality and requirements like *Shortest Move*. Erlewine (2016) posits that anti-locality is violable, and analogously, I suggest that *Shortest Move* is as well, such that longer paths are licensed when necessary for convergence. I also detail a distribution of movement probes under Richards' (2016) *Contiguity Theory* that yields the needed movement paths in another way. This issue remains open, however, and hinges in part on independent questions about how derivation by phase relates the properties of a movement chain's intermediate steps, to properties of the final landing site.