Wholesale Late Merger in A'-movement: evidence from preposition stranding[*]

Juliet Stanton, MIT

December 2014

Abstract

To account for asymmetries between A- and A'-movement, Takahashi & Hulsey (2009) generalize the late merge option (Lebeaux 1988, Chomsky 1995) as Wholesale Late Merger (WLM). In particular, allowing NP to merge with a head D as late as (but no later than) its case position explains why A'- but not A-movement displays Principle C reconstruction effects. This paper claims that WLM is also responsible for pervasive asymmetries *within* the class of A'-extractions. The evidence comes from restrictions on English preposition stranding. I document a correlation between a preposition's complementation properties and its ability to be stranded: prepositions that disallow pronominal complements can only be stranded by a subset of A'-extractions. I argue that the extractions allowing pronoun-rejecting prepositions to be stranded disallow WLM, while those that disallow the stranding allow (and require) WLM.

1   Introduction

It is well-known that A- and A'-movement differ in the reconstruction properties that they exhibit. For example, while both types of movement allow reconstructed bound variable readings, Principle C reconstruction effects are usually obligatory for A'-movement but absent for A-movement[1]. As noted by many authors (e.g. Sauerland 1998, Fox 1999), this pattern poses a challenge for existing theories of movement: while it appears that A'-movement must leave behind a contentful copy, A-movement can act as if it leaves behind a contentless trace.

Takahashi (2006) and Takahashi & Hulsey (2009) argue that this asymmetry arises as a consequence of the possibility for limited countercyclic Merge, known as *Late Merge* (Lebeaux 1988, Chomsky 1993). Building on work by Fox (2002), they propose that late merge is possible whenever an output representation can be interpreted by the semantics (as is the case for the late-merged relative clauses posited by Lebeaux and Chomsky). It is therefore possible for an NP to merge countercyclically to its head D, as the result is semantically interpretable (see Takahashi & Hulsey 2009 for details). They name this optional operation *Wholesale Late Merger*, or WLM.

An additional proposal, that WLM is constrained by the Case Filter (applying to NP as well as DP), explains the differences in A vs. A' reconstruction properties. Since A-movement is movement to a case position, an NP can merge with its head D in the derived position, and bleed Principle C. Because A'-movement is generally movement to a non-case position, this option is not available. In short, independent

[1]Principled exceptions to this statement do, of course, exist. For example, Principle C is bled in A'-movement when the relevant R-expression is contained within an adjunct that has undergone Late Merge (Lebeaux 1988, Chomsky 1995). Principle C effects can be present in A-movement when the moving constituent is forced to reconstruct for other reasons, i.e. for scope (Fox 1999).

differences between A- and A'-movement control the applicability of WLM, which in turn leads to distinct patterns of reconstruction effects.

Since WLM offers a principled reason for the reconstruction contrast between A- and A'-movement, it can be seen as an independent argument for the somewhat controversial proposal that Late Merge is possible in the first place (cf. Chomsky 2004). The present paper strengthens this argument by offering an independent set of arguments for WLM itself (see also Bhatt & Pancheva 2004, 2007; Nikolaeva 2014). I argue that WLM not only explains the well-known contrasts between A- and A'-extractions investigated by Takahashi (2006) and Takahashi & Hulsey (2009), but is also responsible for lesser-known but pervasive differences within the class of A'-extractions: in particular, a set of asymmetries in English preposition stranding (P-stranding). In §2, I document a correlation between two seemingly independent facts about prepositions (Ps): their ability to take a pronominal complement (see also Kuroda 1964, 1968), and their ability to be stranded (see Ross 1967: 119; also Postal 1998, for closely related work focusing on similar constraints on movement out of other antipronominal contexts). A preposition that cannot take a pronominal complement, like temporal *in* (1), allows us to distinguish two types of A'-extraction: those allowing a pronoun-rejecting preposition to be stranded (e.g. wh-movement, restrictive relatives; 2a,b), and those forbidding a pronoun-rejecting preposition from being stranded (e.g. topicalization, *tough*-movement; 2c,d).

(1) Temporal *in* is pronoun-rejecting
   a. * I went swimming in December, and John went swimming in it, too.
   b. * I will be arriving in a few days, and John will be arriving in them, too.

(2) Stranding temporal *in*: restricted[2]
   a. Which month did John go swimming in?   *wh-movement*
   b. The month that John went swimming in was very cold.   *restrictive relative*
   c. * December, John went swimming in.   *topicalization*
   d. * December is tough to swim in.   *tough-movement*

By contrast, a preposition that takes a pronominal complement, like locative *in* (3), can be stranded by any type of A'-extraction (4).

(3) Locative *in* is pronoun-accepting
   a. Michelle's cat hid in the cardboard box, and my cat hid in it too.
   b. I stored my cereal in the pantry, and Chris stored his cereal in it too.

(4) Stranding locative *in*: no restrictions
   a. Which box did Michelle's cat hide in?   *wh-movement*
   b. The box that Michelle's cat hid in was made of cardboard.   *restrictive relative*
   c. That cardboard box, Michelle's cat hid in.   *topicalization*
   d. Cardboard boxes are easy for cats to hide in.   *tough-movement*

I propose that these two types of A'-extraction differ according to whether or not WLM applies. An extraction like *wh*-movement, which can strand a pronoun-rejecting P, leaves behind a fully constructed copy. An extraction like topicalization, which cannot strand a pronoun-rejecting P, leaves behind only a determiner; the rest of the DP merges after movement. In §7, I attribute these differences to interactions

---

[2]The majority of speakers consulted find these contrasts extremely sharp, but some speakers find them subtler (or they detect the contrasts for some but not all predicates or extraction types). I have not found a pattern to this variation, and leave it as a matter for further investigation.

among the Case Filter, a violable constraint favoring WLM, and restrictions on the timing of countercyclic merger (see also Sauerland 1998: 2.2).

Assuming that pronouns are bare determiners (e.g. Postal 1966, Abney 1987), I argue that the link between a preposition's ability to take a pronominal complement and its ability to be stranded is structural. I propose that a pronoun-rejecting preposition requires its complement to contain a certain kind of DP: temporal *in*, for example, requires its complement DP to contain an NP denoting an interval of time (§3). Because a pronominal DP does not contain an NP, it is an unacceptable complement for a preposition like temporal *in*. The base position of extractions like topicalization likewise lacks an NP, explaining the inability of these extractions to strand pronoun-rejecting prepositions.

I provide further support for the proposed analysis by showing that it correctly predicts contrasts not only among extractions, but also contrasts internal to individual extractions. For example, P-stranding with *wh*-pronouns, but not full *wh*-phrases, is sensitive to whether or not a given preposition is pronoun-rejecting (§8). Contrasts of this kind show that the (in)ability to strand pronoun-rejecting prepositions should not be linked to intrinsic properties of individual extractions (cf. Postal 1998): all that matters is the structural configuration of the gap site.

2 P-stranding asymmetries

This section focuses on P-stranding asymmetries in temporal (§2.1) and locative (§2.2) PPs. I show that if a given preposition cannot accept a pronoun as its complement, it can be stranded by only a subset of A'-extractions (see Postal 1998 for similar phenomena).

2.1 Temporals

It is possible to divide temporal DPs into two classes. There are *interval* DPs, referring to points or spans in time (e.g. *Monday, the last five hours*). There are also *event* DPs, referring to events that occupy certain portions of time (e.g. *John's party, Sue's talk*). A basic difference between these two types of temporal DPs is that while an interval DP is defined by their location and extent along a timeline, an event DP is defined by other properties. Some further examples of intervals and events are in (5).

(5) Interval vs. event DPs

    a.   *Intervals*: Monday, 5:00, Christmas, Chris's youth, Mary's birthday...

    b.   *Events*: John's party, Sue's talk, Mary's progress meeting, Christmas dinner...

In temporal PPs, interval and event DPs differ in a fundamental way: an event but not an interval complement to P can be pronominalized with it (6, 7).

(6) Event DP complements: ✓it

    a.   I left after John's party and Mary left after it, too.

    b.   I left before Christmas dinner and John left before it, too.

(7) Interval DP complements: *it[3]

    a.   * I left after 5:00 and Mary left after it, too.

    b.   * My family eats lamb on Easter, and John's family eats lamb on it, too.[4]

An appealing but incorrect analysis of (6, 7) could claim that an interval DP, unlike an event DP, has some property that renders it unable to be pronominalized. If this were the case, the split between the event and interval complements to P in (6, 7) would not be surprising: if an interval complement cannot be pronominalized in general, then its behavior in a PP does not require an independent explanation. Note however that *June*, an interval, can be pronominalized as a subject (8a) or as the complement of *spend* (8b)[5], but not the complement of *in* (8c).

(8) Interval DPs can be pronominalized

    a.   I spent June at the pool. It is my favorite month.

    b.   I spent June at the pool, but John spent it in his office.

    c.   * John visited his family in June, and Mary visited her family in it, too.

The contrast between (8a, b) and (8c) shows that the potential for a given DP to be pronominalized is, at least in part, dependent on the head that selects for it. An interval can be pronominalized as the complement to *spend*, for example, because *spend* accepts a pronominal complement. An interval cannot be pronominalized as the complement to temporal *in* because temporal *in* does not accept a pronominal complement[6].

The pronominalization contrasts in (6, 7), then, do not point to a difference between event and interval DPs per se, but localizes the difference to properties of the prepositions that select for them. A preposition selecting for an event DP is a pronoun-accepting preposition (a P_A): it can accept either a pronoun or a lexical noun as its complement. A preposition selecting for an interval DP, by contrast, is a pronoun-rejecting preposition (a P_R): it accepts a lexical noun, but not a pronoun, as its complement. Examples of pronoun-accepting and pronoun-rejecting temporal Ps are given in (9). Note that many prepositions in English, like before and after in (9), have P_A and P_R uses.

(9) Pronoun-accepting Ps (P_As) vs. pronoun-rejecting temporal Ps (P_Rs).

P_As	P_Rs
during (e.g. during the movie)	on (e.g. on Christmas)
after (e.g. after John's talk)	in (e.g. in three hours)
before (e.g. before Sue's party)	after (e.g. after 5:00)

P_As are the group of prepositions that can be stranded under any type of A'-extraction; representative examples with the P_As *before* (10) and *after* (11) follow.

---

[3]It is possible to replace an interval PP with *then*, i.e. *My family eats turkey on Thanksgiving, and John's family eats turkey then, too*. Since *then* has the distribution of a PP (i.e. it can appear in non-Case positions, as in *John arrived <u>then</u>*), I assume it is a pro-PP: a prepositional element that conflates a temporal P with a nominal whose reference is salient in the discourse. I hypothesize that the reason why *then* (a pro-PP) can replace an interval PP is because, although *then* can refer to a PP headed by an interval-selecting P, it is not itself an interval-selecting P. Some support for the idea that *then* is not interval-selecting comes from the fact that *then* can also refer to event PPs, e.g. *We left after John's talk, and Mary left then too.*

[4]Replacing *Thanksgiving* with a demonstrative, *this* or *that*, is similarly ungrammatical (*My family eats Turkey on Thanksgiving, and John's family eats turkey on that, too.*). Thus the inability of *it* to replace an interval complement to P also cannot be attributed to its status as a weak pronoun (see Cardinaletti & Starke 1994). For an interesting case where the weak vs. strong status of a pronoun does figure into a preposition's complementation properties, see Zribi-Hertz (1984) on French.

[5]Some other verbs that allow interval complements are *waste, kill,* and *use*.

[6]It is worth asking whether or not there are other structural positions that display the same interval vs. event pronominalization asymmetry. To the best of my knowledge, there aren't. We saw in (8) that both object and subject intervals can be pronominalized.

(10)   Stranding *before* (P_A)
  a. Which act does our band play before?   *wh-movement*
  b. The act we're playing before will be really good.   *restrictive relative*
  c. That band's set, we're playing before.   *topicalization*
  d. That band's set will be tough to leave before. *tough-movement*

(11)   Stranding *after* (P_A)
  a. Which talk are we leaving after?   *wh-movement*
  b. The talk we're leaving after should be really good. *restrictive relative*
  c. John's talk, we're leaving after.   *topicalization*
  d. John's talk will be easy to leave after.   *tough-movement*

P_Rs are the group of prepositions that can be stranded in only a subset of these extractions. I illustrate with the P_Rs *on* (12) and *in* (13).

(12)   Stranding *on* (P_R): severely constrained[7]
  a. Which holiday do people eat lamb on?   *wh-movement*
  b. The holiday that most people eat lamb on is Easter. *restrictive relative*
  c. * Easter, we ate lamb on.   *topicalization*
  d. * Easter is easy to eat lamb on.   *tough-movement*

(13)   Stranding *in* (a P_R): severely constrained
  a. Which month do people usually swim in? *wh-movement*
  b. The month that most people swim in is June. *restrictive relative*
  c. * June, we swam in.   *topicalization*
  d. * June is easy to swim in.   *tough-movement*

The contrasts in (10, 11) vs. (12, 13) document a correlation between the ability of a given preposition P to take a pronominal complement, and the ability of that P to be stranded. Some A'-extractions allow a P_R to be stranded, but others do not. Following Postal (1998)'s terminology for extraction types, based on observations of other antipronominal contexts[8], I refer to an extraction allowing a P_R to be stranded as an A-type extraction (not to be confused with A-movement). An extraction forbidding a P_R from being stranded is a B-type extraction. The full partition is given in (14), and examples are provided in (15-16). Throughout the rest of the paper, I use *wh*-movement and restrictive relatives as representative examples of A-types, and topicalization and *tough*-movement as examples of B-types.

(14)   Partition of extractions

A-type extractions	**B-type extractions**
(stranding P is always okay)	(if P is a P_R, don't strand it)
wh-movement	appositive relatives
restrictive relatives	negative inversion
infinitival relatives	topicalization
free relatives	parasitic gap-formation
	tough-movement
	gapped degree phrases

---

[7]Note that pied-piping can rescue these sentences: compare *not a single holiday will I ever eat lamb on* to the more acceptable *on not a single holiday will I ever eat lamb*.

[8]Postal (1998) shows that some A'-extractions are sensitive to other antipronominal contexts (name positions, change-of-color environments, etc.). His partition differs in several ways from mine; potential reasons for these differences are not explored here.

(15) P$_A$: *after* (We left after the afternoon session, and John left after it too.)

    a.   Which session did you leave after? *wh-movement; A*

    b.   The session that we left after was very boring. *restrictive relative; A*

    c.   I hear that the session to leave after is the afternoon one. *infinitival relative; A*

    d.   Whatever session we left after was very boring. *free relative; A*

    e.   The afternoon session, which we left after, was very boring. *appositive relative; B*

    f.   The afternoon session, we decided to leave after. *topicalization; B*

    g.   Not a single session will I ever leave after. *negative inversion; B*

    h.   Which session did you think was boring without leaving after? *parasitic gap; B*

    i.   The afternoon session will be tough to leave after. *tough-movement; B*

    j.   The afternoon session will be too good to leave after. *gapped degree phrase; B*

(16) P$_R$: *on* (*I go to class on Mondays, and Audrey goes to class on them too.)

    a.   Which days do you go to class on? *wh-movement; A*

    b.   The days that I go to class on are Mondays and Wednesdays. *restrictive relative; A*

    c.   The best days to go to class on are Tuesdays. *infinitival relative; A*

    d.   Whatever days you went to class on, you disliked them. *free relative; A*

    e.   * Mondays, which Audrey goes to class on, are always cold. *appositive relative; B*

    f.   * Mondays, Audrey always goes to class on. *topicalization; B*

    g.   * Not a single weekday do I ever go to class on. *negative inversion; B*

    h.   * Mondays are easy to enjoy without going to class on. *parasitic gap; B*

    i.   * Mondays are tough to go to class on. *tough-movement; B*

    j.   * Mondays are too cold to go to class on. *gapped degree phrase; B*

The partition in (14) makes explicit an implicational generalization. A P$_A$ can be stranded under A-type and B-type extractions, but a P$_R$ can only be stranded under A-types. Therefore any P that can be stranded under a B-type extraction must also be able to be stranded under all A-type extractions, but not vice versa:

(17) Generalizing over extractions
If a given preposition P can be stranded by a B-type extraction, that P can also be stranded by all A-type extractions.

A counterexample to the generalization in (17) would be a preposition that could be stranded by a B-type extraction (e.g. topicalization) but not an A-type extraction (e.g. *wh*-movement). To the best of my knowledge, such a preposition does not exist.

2.2 Locatives

Similar to temporal DPs, we can divide locative DPs into two classes. There are *location* DPs, which refer to points or regions of space (e.g. *the second balcony, the fourth floor*). There are also *entity* DPs, referring to physical entities that occupy certain portions of space (e.g. *the box, the car*). Like the difference between an interval and an event (temporal) DP, the difference between a location and an entity (locative) DP is semantic. A location DP like *the fourth floor* is defined by its spatial coordinates, while an entity DP like *the box* are defined by other properties. Further examples of location and entity DPs follow (18).

(18) Location vs. entity DPs

  a. *Locations*: the ground floor, the fourth floor, 10,000 feet, the sky...

  b. *Entities*: the box, the forest, the hut, the television, the car...

The distinction between an entity and a location DP is seen clearly in locative PPs. A preposition that selects for an entity complement is pronoun-accepting (19), while a preposition selecting a location complement is pronoun-rejecting (20).

(19) Entity DP complement: ✓it

  a. I ate dinner on the wooden table and John ate dinner on it, too.

  b. I climbed to the summit of the mountain and John climbed to it, too.

(20) Location DP complement: *it[9]

  a. * I ate dinner on the fourth floor and John ate dinner on it, too.

  b. * My airplane climbed to 10,000 feet and John's airplane climbed to them, too.

Note that the sentences in (20) are only ungrammatical under the location readings of *the fourth floor* and *10,000 feet*. (20a) is grammatical if John ate dinner on the surface of some floor, which happens to be fourth in some series of floors; (20b) is grammatical if 10,000 feet is a visible destination that can be climbed to, like the top of a mountain. In other words, (20a-b) are grammatical if the DP is treated as an entity. This is consistent with the generalization that an entity DP can be pronominalized as P's complement, but a location DP cannot be.

The stranding facts parallel the observations concerning temporals in §2.1. A preposition selecting for an entity DP can be stranded under any type of A'-extraction; examples are given with *in* (21) and *on* (22), in their P_A uses.

(21) Stranding *in* (a P_A)

  a. Which box was the cat hiding in?   (A)

  b. The box that the cat was hiding in is brown.   (A)

  c. Michelle's box, the cat likes to hide in.   (B)

  d. Michelle's box is easy for the cat to hide in.   (B)

(22) Stranding *on* (a P_A)

  a. Which table did Colleen dance on?   (A)

  b. The table that Colleen danced on is gone now.   (A)

  c. The coffee table, Colleen danced on last night.   (B)

  d. The coffee table was tough to dance on.   (B)

A preposition selecting for a location DP, however, can only be stranded under A-type extractions. I illustrate with *to* (23) and *on* (24), in their P_R uses.

(23) Stranding *to* (a P_R)

  a. What height did the plane fly to?   (A)

  b. The height that the plane flew to was 10,000 feet.   (A)

---

[9]Like temporal PPs and *then*, it is possible to replace locative PPs (or at least, part of them; see Svenonius 2008 for some discussion) with *there*. Thus *I ate dinner on the fourth floor and John ate dinner there, too* is grammatical. Similar to *then*, I assume *there* to be a pro-PP that takes as its antecedent any kind of locative PP.

    c.   * 10,000 feet, the plane flew to. (B)

    d.   * 10,000 feet is tough to fly to. (B)

(24)   Stranding *on* (a P_R)

    a.   Which floor did we find cake on? (A)

    b.   The floor that we found cake on was deserted. (A)

    c.   * The fourth floor, we found cake on. (B)

    d.   * The fourth floor is easy to find cake on. (B)

Place names, a type of locative DP, are ambiguous between locations and entities. A place name like *Boston* can be interpreted in two ways: as a set of coordinates (either geographical, or a point on a cognitive map) or as a physical entity, as Chomsky (1999: 42) notes: "if London is reduced to dust, it ... can be re-built elsewhere and still be the same city." We can say that London was reduced to dust because place names, in one sense, refer to a collection of entities that occupy specific points in space. We can also say that London can be rebuilt elsewhere with different buildings because, in another sense, place names refer to a set of abstract coordinates, not necessarily linked to their physical contents.

To the extent that a place name can be pronominalized as the complement of a preposition, the entity reading is preferred (25).

(25)   Pronominalizing place names

    a.   I traveled to France, and John traveled to it too. (*location, ✓entity)

    b.   I ate lunch on the Charles River. John ate lunch on it too. (*location, ✓entity)

The first conjunct in (25b) has a location reading under which I sat on the banks of the Charles River, near the water, and ate my lunch. This reading is absent in the second conjunct: the only licit interpretation here is one in which John sat on the surface of the Charles River and ate his lunch. In other words, the sentence is only licit if the Charles River is treated as an entity.

In (26), we see that the location reading of on the Charles River is preserved when *on* is stranded by A-type extractions (26a-b), but not by B-type extractions (26c-d).

(26)   Place names and stranding (* = lack of a location reading)

    a.   Which river in Boston did you eat lunch on? (A)

    b.   The river that we ate lunch on was the Charles. (A)

    c.   * The Charles River, we ate lunch on. (B)

    d.   * The Charles River is easy to eat lunch on. (B)

The lack of an available location reading in (26c,d) is consistent with the generalization that a preposition selecting a location DP cannot be stranded by B-types[10].

2.3 Summary

The previous two subsections demonstrate a link between two things: (1) whether or not a preposition P can accept a pronoun as its complement, and (2) whether or not that P can be stranded by B-type extractions. The relevant generalizations are as follows. If P is pronoun-accepting, then P can be stranded by both A-type and B-type extractions.

---

[10]The V+P combination *live in* is a counterexample to the generalizations established here. It is impossible to pronominalize *Boston*, for example, as the complement of *in* (**I live in Boston, and John lives in it, too*), but stranding *in* with B-type extractions, for most speakers, is possible (*Boston is easy to live in.*). This is an isolated counterexample, and its status may have more to do with properties of *live in*, yet to be defined.

(27) If ✓[P pronoun], then ✓A-type and ✓B-type.

If P is pronoun-rejecting, P can be stranded under A-type extractions only.

(28) If *[P pronoun], then ✓A-type but *B-type.

Although the empirical focus of this paper concerns temporal and locative PPs, the generalizations in (27-28) hold more widely. In (29), for example, instrumental *with* is a P_A, and can be stranded under both A-type and B-type extractions. In (30), by contrast, circumstantial *under* is a P_R, and can be stranded under A-type extractions only.

(29) Anthony eats his oatmeal with (that spoon / it).
  a. Which spoon does Anthony eat his oatmeal with? (A)
  b. The spoon that Anthony eats his oatmeal with is large. (A)
  c. That spoon in the sink, Anthony eats his oatmeal with. (B)
  d. Large spoons are easy to eat oatmeal with. (B)

(30) David will only celebrate under (a restricted set of circumstances / *them).
  a. What set of circumstances will David celebrate under? (A)
  b. The set of circumstances that David celebrates under are unknown. (A)
  c. * That set of circumstances, David celebrates under. (B)
  d. * That set of circumstances is easy to celebrate under. (B)

The remainder of the paper seeks to explain why some prepositions are P_Rs, and why this should have anything at all to do with their potential to be stranded.

3 On antipronominal contexts

In §2, we established that temporal and locative DP complements to P can be divided into two classes, according to their ability to be pronominalized with *it* (or *them*). As the complement to a temporal preposition, an event but not an interval DP can be pronominalized. As the complement to a locative preposition, an entity but not a location DP can be pronominalized.

Interval and location DPs – the types of DPs that cannot be pronominalized as P's complement – are *coordinate-denoting* DPs. A coordinate-denoting DP is a predicate ranging over either spatial or temporal dimensions; such a DP is defined by the amount of time or space that it occupies. Events and entities – the types of DPs that can be pronominalized as P's complement – are *concrete* DPs, in the sense that they are predicates ranging over entities that occupy certain portions of temporal or spatial dimensions. The particular type of space that a concrete DP occupies determines whether it is treated as a locative or a temporal, but it is defined by properties other than its spatial coordinates. The event DP *John's party*, in (31), illustrates: John's party takes up a portion of Monday, but its length does not define it.

(31) *John's party*: an event DP

Prepositions selecting for coordinate-denoting DPs are P_Rs: they can take DP containing lexical nouns (i.e. *the box, John's party*), but not pronouns, as their complements. Prepositions that select for concrete DPs are P_As: they can take either pronouns or DPs containing lexical nouns as their complements.

9

What does it mean, though, for a preposition to be pronoun-rejecting? What is the relevant difference between a pronoun and a coordinate-denoting DP, such that a preposition can discriminate between the two? One potentially relevant difference is that pronouns and lexical nouns are not members of the same syntactic category: pronouns are determiners that do not have an NP complement (Postal 1966, Abney 1987, Uriagereka 1995, though cf. Elbourne 2005), while lexical nouns are Ns. The first argument for this difference in category membership comes from the fact that pronouns and determiners are in complementary distribution (32a), but lexical nouns and determiners are not (32b, c).

(32)   Pronouns and determiners are in complementary distribution (Abney 1987: 281)

    a.   * The she that I talked to was nice.

    b.   The Mary that I talked to was nice.

    c.   The woman that I talked to was nice.

The fact that determiners are in complementary distribution with pronouns (but not lexical nouns) suggests that pronouns (but not lexical nouns) have the same structural position as determiners. Further evidence for this comes from the fact that pronouns can act as overt determiners under certain circumstances, e.g. *we religious ones* (see Postal 1966[11]).

If the analysis of pronouns as determiners is on the right track, then a pronoun and a phrasal DPs are structurally distinct: a phrasal DP like *the box* contains an NP, but a pronoun like *it* does not. Returning to the properties of P_Rs, the basic proposal here is that a P_R is a preposition requiring its complement DP to contain something more than just a bare D. In essence, it appears that the complement of a P_R must contain an NP. But when we look more closely at the selectional restrictions of P_Rs like temporal *in* and *on*, it becomes clear that it isn't just any sort of NP that a P_R requires: it's a certain semantic type of NP. In the examples below, we see that temporal *in* (33) and *on* (34) require interval DP complements, and forbid event DP complements.

(33)   Interval-selecting *in*

    a.   I'll be there in five hours.   (✓interval DP)

    b.   * I'll be there in John's party.   (*event DP)

(34)   Interval-selecting *on*

    a.   I'll be there on Monday.   (✓interval DP)

    b.   * I'll be there on Sue's talk.   (*event DP)

The only situation in which (33b) and (34b) are marginally acceptable is one in which the speaker is using *John's party* and *Sue's talk* to denote units of time – in other words, if the DPs are treated as intervals. Thus we cannot simply say that a P_R is a preposition requiring its complement to contain an NP, because sentences like (33b) and (34b), where a temporal P_R takes an event complement, are ungrammatical.

Understanding why the complements of temporal P_Rs are restricted to intervals requires us to find a property of intervals that both events and pronouns lack. Following Kayne (2005a), I propose that an interval DP contains a nominal category belonging to a particular semantic category, abbreviated here as TIME. TIME can be either null (as in *December the 25th* (DAY)) or pronounced (as in *the* MONTH *of June*, *Avery's wedding* DAY, etc.), and this TIME element is what lends the DP its intervalic denotation. For evidence that TIME resides in N, consider the alternation in (35)[12].

---

[11]Others argue that examples like *we religious ones* should be analyzed as appositive modifiers to a pronominal head. See Delorme & Dougherty (1972), and Pesetsky (1978) for a criticism.

[12]For temporals, I assume that *day, month*, and other interval-denoting nouns are sub-species of TIME. Similarly, for locatives, I assume that *city, state*, and other place-denoting nouns are sub-species of PLACE.

(35) TIME is a noun

    a.   the twenty-fifth (DAY) of December

    b.   December twenty-fifth (DAY)

In (35a), we see that the null DAY acts as a noun, as it has a determiner, an ordinal modifier, and a PP modifier. (35b) is derivationally related to (35a): *December* has moved into a position structurally higher than D and lost its preposition (perhaps because *December* is now in Spec, DP, and no longer needs the case assigned by *of*). Note that even though pronouncing the TIME element in (35b) is marginally acceptable at best, the presence of the ordinal modifier *twenty-fifth* strongly suggests that something occupies N.

We can say something similar for location DPs. Following Kayne (2005b), I propose that a location DP contains the noun PLACE. Like TIME, PLACE can be either null (as in *Boston* (CITY)) or overt (as in *the Charles* RIVER, *New York* CITY). PLACE lends a locative DP a location denotation by signifying that the DP occupies a span of coordinates along a spatial plane. Some evidence that PLACE occupies N comes from the alternation in (36), structurally identical to (35).

(36) PLACE as a noun

    a.   the CITY of New York

    b.   New York CITY[13]

Returning now to P-stranding, we can say that an interval-selecting temporal P_R is a preposition that requires its complement to contain TIME, and a location-selecting locative P_R is a preposition that requires its complement to contain PLACE. Presumably this is a semantic requirement, as there is no reason to believe that an interval DP that contains TIME is syntactically distinct from an event DP[14].

In sum, we can now argue that a preposition requiring its complement to contain either TIME or PLACE is a coordinate-selecting P. When a coordinate-selecting P takes a pronominal complement, its need for a TIME or PLACE element is not met. This is because a pronominal DP does not contain an NP, and therefore cannot possibly contain TIME or PLACE[15].

4   Linking complementation and stranding

We can now re-characterize the link between complementation and stranding in light of the previous section. A temporal or locative P_R, whose semantics require its complement to contain TIME or PLACE, can be stranded only under A-type extractions. A temporal or locative P_A, which does not require its complement to contain TIME or PLACE, can be stranded under both A-type and B-type extractions. The question, then, is this: how can we distinguish the A-type extractions from the B-type extractions in a way that allows us to capitalize on the specific needs of P_Rs?

I propose that the A-type and B-type extractions differ in their sensitivity to the selectional restrictions of stranded prepositions because they leave behind copies of different sizes. A-type extractions (e.g. *wh-*movement, restrictive relatives) leave behind fully constructed copies (see Chomsky 1993, 1995). A-type extractions can strand P_Rs because the copy in the base position can contain an NP. If the copy in the base

---

[13] An interesting difference between the temporal example in (35b) and the locative example in (36b) is that the presence of *the* is grammatical in the former (*December 25th* vs. *December the 25th*), but not in the latter (*New York City* vs. **New York the city*). It is also worth noting using *the* in temporal DPs is impossible for holidays named by their dates (*July 4th* vs. **July the 4th*).

[14] An anonymous reviewer notes that the restriction could also be due to a syntactic constraint that penalizes a failure to meet a head's complementation property. This is a possibility; what would need to be worked out is how it is possible for a P to see 'through' the structure of its complement DP to discern the identity of the embedded NP.

[15] Although the only types of PPs I discuss in this section and in what follows are locative and temporal PPs, I anticipate that the analysis proposed here could extend to circumstantial and other types of PPs (see §2.3).

position can contain an NP, it can contain TIME or PLACE, and the selectional restrictions of P_Rs can be satisfied (37).

(37) A-type extractions leave full copies

    a. Which day of the month did you eat lamb on$_R$ [which day of the month]?

    b. Which city are we eating lunch in$_R$ [which city]?

B-type extractions, by contrast, leave behind something smaller than a full DP – crucially, something that does not contain an NP. Because TIME and PLACE reside within the NP, the selectional restrictions of P_Rs cannot be satisfied (38).

(38) B-type extractions do not leave NPs

    a. * No holiday will I ever eat lamb on$_R$ t.[16]

    b. * The fourth floor is fun to eat dinner on$_R$ t.

The proposed link between complementation and stranding, then, is structural. The reasons why a P_R cannot take a pronoun as its complement, or be stranded under B-type extractions, are one and the same. Neither of these configurations allows P's complement to contain an N, where TIME and PLACE reside[17].

To make sense of this idea, we need to allow for the possibility that some A'-extractions leave behind full copies, while others do not. This is where the theory of Wholesale Late Merger (WLM; Takahashi 2006, Takahashi & Hulsey 2009) comes in. In short, WLM proposes that certain extractions (i.e. A-movement constructions) permit countercyclic merger of an NP to its head D, while other extractions (i.e. *wh*-movement) do not. I provide some necessary background on WLM (§5), and extend the theory to account for the differences among A'-extractions discussed here (§6-7).

5 Wholesale Late Merger

As noted in the introduction, WLM addresses a well-known problem for the copy theory of movement (see Chomsky 1993, 1995): extractions appear to differ in whether or not they leave a contentful copy behind. Takahashi (2006) and Takahashi & Hulsey (2009) (hereafter T±H) focus on the difference between A- and A'-movement: this basic distinction will be the focus of this section's discussion, though later there will be reason to look more closely at differences within the class of A'-extractions.

The presence of reconstructed bound variable readings in both A- and A'-movement suggests that both types of movement can leave behind a full copy. In the examples below (39), the QP *every professor* binds the variable *his*, suggesting that a copy of the variable is c-commanded by the QP at LF.

(39) Reconstructed bound variable readings (a: T&H 2009: 391, 390)

    a. Someone from his$_i$ class seems to [every professor]$_i$ to be a genius.   (A)

    b. Which of his$_i$ students did John think that [every professor]$_i$ talked to?   (A')

The data in (39) show that leaving a full copy behind is at least optional for A- and A'-movement. But is it obligatory? For A'-movement, Principle C reconstruction effects suggest that the answer is yes. As shown in (40a), sentences where an R-expression in the moving NP is c-commanded by a coindexed pronoun at LF are ungrammatical. For A-movement, however, the answer appears to be no: Principle C reconstruction effects are absent (40b).

---

[16]t denotes that what is left behind is less than a full copy; exactly what is left behind will be discussed later.

[17]The idea that some extraction sites behave like pronouns because they are, in some sense, pronouns, builds on previous work drawing the same conclusion, such as Perlmutter 1972, Cinque 1990, and Postal 1998. While the idea has antecedents, the proposed formalization is novel. See §8.2 for a brief comparison with Postal 1998.

(40) Principle C: A- vs. A'-movement

    a. Which corner of John$_i$'s room was he$_i$ sitting in?   (T&H 2009: 391; A')

    b. John$_i$'s mother seems to him$_i$ to be wonderful.   (Lebeaux 1998: 23; A)

The A-movement facts pose a problem for the copy theory of movement. While it appears that A'-extractions must always leave a copy behind, A-extractions do so only optionally. The presence of reconstructed bound variable readings for A-extractions suggests that leaving a full copy in the base position is a possibility, but the absence of Principle C effects suggests that leaving a full copy behind is not obligatory.

To solve this problem, T±H propose that it is possible for an NP to countercyclically merge with its head D (see 41). Their proposal builds on work by Fox (2002), who proposes that late merge is possible whenever the output representation is semantically interpretable. Fox's (2002) proposal differs from Lebeaux's (1988, 1998) theory of late merge because it does not limit the operation to adjunct phrases. It allows an NP to merge late with its head D, because the result is interpretable (see T±H for details).

(41) Application of WLM
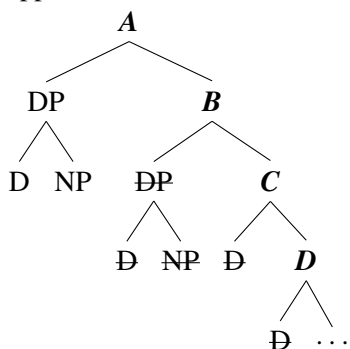
In the diagram above, the bare D is externally merged in Position D. It then moves to Position C, and again to Position B. Position B is the site of WLM: here, the NP merges with its head D. The entire DP then moves to Position A, where it is pronounced.

(41) shows how WLM allows for A-extractions to bleed Principle C. If an NP containing an R-expression does not merge with D until D is outside the c-command domain of a co-indexed nominal, Principle C effects will be absent.

The simple mechanism of WLM, however, does not yet allow us to explain why the reconstruction properties of A- and A'-movement should differ. Why, for example, is it possible for A-movement to bleed Principle C through WLM, while this option is not available to A'-movement? In other words: if we grant that WLM is an option available to the grammar, what constrains it?

T±H's answer to this question is that WLM is constrained by the Case Filter. To move towards an explanation of the differences between A- and A'-movement, T±H advocate an approach in which "DPs as a whole demand Case because both determiners and nouns, which constitute DPs, must receive case" (T&H 2009: 401)[18]. Thus WLM is blocked when an NP would merge outside the domain of its case-assigner.

The appeal to Case allows T±H to capture the difference between A- and A'-movement. A-movement, which is movement from a non-case to a case position, permits the application of WLM (though cf. Nikolaeva 2014: 5 for a discussion of reconstruction effects in Russian scrambling, a form of A-movement to a non-case position). This means that a restrictor NP can merge outside of the c-command domain of a coindexed nominal. In (42), Principle C is bled because there is no R-expression in the base position.

---

[18]Support for the idea that both Ds and Ns receive case comes from many Indo-European languages, i.e. Greek (Ancient (Smyth 1956) and Modern (Sabine Iatridou, p.c.)), where Ds and Ns bear case morphology.

(42) WLM permitted in A-movement   ((38b) from T&H 2009: 402)

    a. The corner of John$_i$'s room seems to him$_i$ [~~the~~] to be very dusty.

    b. Every argument that John$_i$ is a genius seems to him$_i$ [~~every~~] to be flawless.

A'-movement, by contrast, is movement from a case to a non-case position. T±H propose that WLM is blocked in this context because the NP must merge in the base position to get case. In the examples in (43), Principle C effects are present because *he* c-commands *John* at LF.

(43) WLM blocked in A'-movement

    a. Which corner of John$_i$'s room was he$_i$ sitting in [~~which corner...~~]?

    b. Which proof that John$_i$ is a genius did he$_i$ believe [~~which proof...~~]?

In sum, the reconstruction properties of A- and A'-movement differ because the two types of movement place differing restrictions on the application of WLM. A-movement permits WLM because an NP can merge outside of its base position and still receive case. A'-movement blocks WLM because the NP must merge in its base position to receive case. What is important is that, depending on the extraction in question, the size of copy that a DP leaves behind can differ.

6   Relating WLM to P-stranding

Although the only A'-extraction that T±H consider is *wh*-movement, they implicitly assume that all varieties of A'-movement exhibit Principle C effects. Looking more broadly across the variety of A'-extractions discussed in relation to P-stranding, we see that this assumption is mostly well-founded. A representative sample of A'-extractions displaying Principle C effects is in (44).

(44) a. * Which criticism of Mary$_i$'s proposal did she$_i$ reject?   *wh-movement*

    b. * The student whose criticism of Mary$_i$'s proposal she$_i$ rejected...   *restrictive relative*

    c. * That student, whose criticism of Mary$_i$'s proposal she$_i$ rejected...   *appositive relative*

    d. * Not a single criticism of Mary$_i$'s proposal has she$_i$ ever rejected.   *negative inversion*

    e. * The harsh criticism of Mary$_i$'s proposal, she$_i$ rejected.   *topicalization*

Parasitic gaps (45a), *tough*-movement (45b), and gapped degree phrases (45c) behave differently, however: like A-extractions, Principle C effects are absent.

(45) Principle C effects absent

    a. Which criticism of Mary$_i$'s proposal did John endorse after she$_i$ rejected?

    b. The criticism of Mary$_i$'s proposal is hard for her$_i$ to accept.

    c. The criticism of Mary$_i$'s proposal is too harsh for her$_i$ to believe.

Crossing the presence/absence of Principle C effects with a given extraction's ability to strand a P_R yields the table in (46). We can now distinguish among three different types of extractions. The A-type extractions (e.g. *wh*-movement) are extractions that exhibit Principle C effects and permit stranding of P_Rs. The B_1-type extractions (e.g. topicalization) exhibit Principle C effects, but ban stranding of P_Rs. The B_2-type extractions (e.g. *tough*-movement) bleed Principle C and ban stranding of P_Rs.

(46) Principle C effects vs. stranding

	Principle C effects: Present	Principle C effects: Absent
✓ [P_R t]	**A-type extractions** *wh*-movement restrictive relatives infinitival relatives free relatives	
*[P_R t]	**B$_1$-type extractions** appositive relatives topicalization negative inversion	**B$_2$-type extractions** *tough*-movement gapped degree phrases parasitic gap-formation

If we allow for the possibility that WLM can occur in A-movement and A'-movement alike, WLM can help us understand why A-type, B$_1$-type, and B$_2$-type extractions differ according to the properties in (46). I sketch this analysis below.

For A-type extractions (e.g. *wh*-movement), the NP must merge maximally early, in the base position. Thus in (47a) the pronoun *she* c-commands *Mary* at LF, and Principle C reconstruction effects result. Because merging the NP in the base position means that a full copy is left behind, a P_R can be stranded (47b): because the preposition's complement can contain an NP, it can contain TIME or PLACE.

(47) Properties of A-types (here: *wh*-movement)

    a.   * Which criticism of Mary$_i$'s proposal did she$_i$ reject [which criticism...]?

    b.   Which holiday does your family eat lamb on$_R$ [which holiday]?

In the B$_1$-type extractions (e.g. topicalization), WLM must occur early, but not too early. We know from the presence of Principle C effects (48a) that the site of WLM cannot be maximally late, as in A-movement, because the R-expression *Mary* must be merged within the c-command domain of the coindexed pronoun *she*. We also know from the ban on stranding P_Rs (48b) that WLM is not blocked entirely: the fact that B$_1$-type extractions are sensitive to the distinction between P_As and P_Rs suggests that their base positions do not contain NPs. Thus the site of WLM must be above the base position, but within the c-command domain of a coindexed pronoun.

(48) Properties of B$_1$-types (here: topicalization)

    a.   * The latest criticism of Mary$_i$'s proposal, she$_i$ rejected *t*.

    b.   * Easter, I would never eat lamb on$_R$ *t*.

The analysis of B$_2$-type extractions adopted here (see §8.1 for some justification) assumes that all B$_2$-type extractions involve movement of a null operator (see e.g. Chomsky 1977, 1981 for *tough*-movement; Brillman (2014) for gapped degree phrases; Nissenbaum 2000 and references there for parasitic gaps). Because null operators do not contain lexical material, this analysis correctly predicts that Principle C effects are absent (49a) and that P_Rs cannot be stranded (49b). The selectional restrictions of P_Rs can't be satisfied because their complements do not contain TIME or PLACE[19].

(49) Properties of B$_2$-types (here: *tough*-movement)

    a.   The criticism of Mary$_i$'s proposal is tough for her$_i$ to accept [∅].

    b.   * Easter is easy to eat lamb on$_R$ [∅].

---

[19] An alternative is to assume an improper movement analysis of *tough*-movement and GDPs (e.g. Brody 1993, Hornstein 2000, Hicks 2003) and say that WLM occurs maximally late (this idea due to Norvin Richards, p.c.). See §8.1 for brief discussion.

Note that there is no extraction that can both bleed Principle C and strand a P_R. This gap is significant, and predicted from a WLM-style analysis of the patterns in (46). An extraction that can bleed Principle C is by definition an extraction with no NP in the base position. An extraction that can strand a P_R is by definition an extraction where the NP must merge in the base position, in order to satisfy the selectional restrictions the P_R. A WLM-style analysis of this type of extraction is impossible, because it is a contradiction: the NP needs to merge both very late (to bleed Principle C) and very early (to strand a P_R).

7 Deriving the patterns

What we need, at this point, is an explanation for why WLM is tightly constrained in B_1-type extractions and blocked in A-types. What motivates WLM, and why does its application vary across A'-extractions?

The logic of the facts we have observed suggests that WLM is not just an option, but rather a violable preference, in the Optimality Theoretic sense (Prince & Smolensky 1993 [2002]). This principle lies at the heart of the proposal developed here. I claim, furthermore, that WLM is constrained in A-type and B_1-type extractions because there are conflicting constraints that take priority over the preference for WLM. Following T±H, one such conflicting constraint is the Case Filter. To model this interaction, we need two constraints: one requiring NPs to receive case (GETCASE, defined in (50)), and one enforcing the general preference for WLM (LATEMERGE, defined in (51)).

(50)   GETCASE: one * if NP is caseless.

(51)   LATEMERGE: If NP is merged at position x, assign one * for each possible merge site $x\prime$ whose mother node c-commands x.

The presence of Principle C reconstruction effects in A'-movement shows that GETCASE dominates LATEMERGE: it is more important for an NP to receive case than it is for that NP to be merged as late as possible. The interaction between GETCASE and LATEMERGE is modeled schematically in (52); the tableau should be read from top down. A boxed letter (here, C) indicates the last site of WLM where NP can receive case; the node where WLM occurs is bolded, italicized, and underlined. I assume that successive cyclic movement occurs from positions D to A. Note that the constraints and tableaux in this section are presented only as a way of formalizing and visualizing the pressures governing the application and location of WLM. Presumably, there are deeper reasons why these constraints exist, as well as reasons why they interact in the way that they do.

(52)   Tableau illustrating GETCASE » LATEMERGE

	a.	b.	>c.	d.
GETCASE	*!	*!		
LATEMERGE		* (A)	** (A,B)	***! (A,B,C)

16

Both (52a) and (52b) incur fatal violations of GETCASE because the NP is merged above C, the last position where case can be assigned. The decision between (52c) and (52d) is made by LATEMERGE. (52d) incurs three violations of LATEMERGE, as there are three potential landing sites (A, B, and C) that c-command D, the location where the NP is merged. (52c) incurs only two violations of LATEMERGE, as there are two potential merge sites (A and B) that c-command the site of WLM. Between (52c) and (52d), (52c) best satisfies LATEMERGE, and is selected as optimal.

This analysis predicts that, all else being equal, *WLM of NP with D will occur at the latest step in the derivation at which the NP can receive case*. Looking more globally at the position of LATEMERGE in the grammar, however, it is clear that there are other constraints that outrank it. For example, the need to establish a variable binding dependency can force an NP to merge earlier than its Case position (T±H, §5). In addition to the necessity to bind a variable, it is likely that there are quite a few other semantic constraints that overrule the preference for WLM[20]. An anonymous reviewer notes three such constraints, listed here. First, scope reconstruction is possible in A-movement, suggesting that the need to derive non-surface scope can overrule LATEMERGE (53a). Second, Fox (2000) and others have noted that objects of creation verbs undergo reconstruction for what are arguably semantic reasons; see (53b). In addition, Sportiche (2006) argues that when A-movement applies to pieces of idiom chunks in French, those idiom chunks must undergo reconstruction.

(53) Reconstruction in A-movement
    a. Reconstruction for non-surface scope (Fox 1999: 160)
       Someone from New York is very likely to win the lottery (some >likely; likely >some)
    b. Reconstruction for objects of creation verbs (Fox 2000: 170)
       #For those issues to be clarified, [many more/new papers about Quine$_i$'s philosophy] seem to him$_i$ to be needed.

Although the global interaction of LATEMERGE with other constraints is an area worthy of further study, the remainder of this paper focuses more narrowly on the interaction of only those constraints necessary to explain the P-stranding facts introduced in §2: that is, for now, the interaction of LATEMERGE and GETCASE. §7.1 discusses case assignment in PPs, and shows that the schematic analysis sketched in (52) is sufficient to derive the properties of B$_1$-type extractions. In §7.2, I suggest that the moving DP in A-type extractions is embedded within an additional layer of functional structure, Cable's (2007, 2010) QP (see also Hagstrom 1998), and propose that WLM is blocked in these extractions because it would be in a sense, *too* countercyclic. While WLM is by definition a countercyclic operation, there are constraints on just how countercyclic it can be (see also Tada 1993: 63-70, Sauerland 1998: 2.2). In §7.3 I show that the analysis presented here makes correct predictions regarding different types of extraction out of other domains.
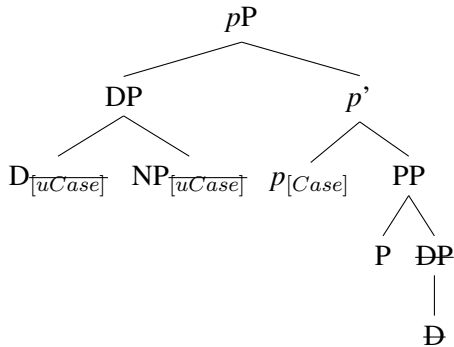
7.1 Deriving B$_1$-type extractions

Recall that, in order to derive both the P-stranding and reconstruction properties of B-types, it must be the case that the site of WLM in these extractions is very low – but not maximally low, as a B-type extraction

---

[20]An anonymous reviewer notes that the effects of LATEMERGE can only be seen under very restricted circumstances, as there are many other constraints that dominate it. While it is true that the effects of LATEMERGE are often invisible, especially in the case of A'-movement, other phenomena display the same split between A– and B-type extractions documented here. Take, for example, Lasnik & Stowell's (1991) discussion of *weakest crossover*, where they demonstrate that certain A'-extractions – but not all – behave like A-movement in that weak crossover violations appear to be absent entirely. The extractions where weak crossover is present include *wh*-movement and restrictive relatives (my A-types); the extractions where weak crossover is absent include topicalization, parasitic gap-formation, and *tough*-movement (my B-types). Various authors (e.g. Koopman & Sportiche 1983, Lasnik & Stowell 1991, Sauerland 1998) have related the presence or absence of weak crossover to the contents of the gap site (i.e. pronominal vs. non-pronominal); perhaps weak(est) crossover is an example of another phenomenon in which the effects of LATEMERGE are discernible in A'-movement.

cannot strand a P_R. One possible way to achieve this balance is to assume that PPs have a pP shell, and that p is what assigns case to P's complement. Arguments for a little p projection have been made by a number of different authors, on a number of different grounds (see van Riemsdijk 1990; Manninen 2003: 117 and references there; Ramchand & Svenonius 2004; Svenonius 2003, 2008, 2010; Levinson 2011; Oxford 2011; and many others). In A'-movement, assuming that pP is a phase (see e.g. Abels 2003), and that Case can be assigned in a Spec-head relationship[21], the last possible landing site where an NP will be able to receive case is Spec, pP. The analysis sketched in §7 then predicts that, for A'-movement out of pP, WLM should take place in Spec, p[22]. This is because Spec, p is the last possible landing site at which an NP can receive case (54).

(54) WLM in Spec, pP

This simple proposal is sufficient to derive the Principle C and P-stranding facts for B_1-type extractions. No conceivable pronoun can intervene between the complement of P and Spec, pP. Thus Principle C effects are present, just as if the NP were merged in the base position. But because what is left in the base position is just a determiner, P's complement does not contain an NP, and a P_R cannot be stranded.

An obvious question arises here: if case considerations can force NP to merge with D earlier than otherwise required, why can't the selectional restrictions of a P_R have a similar effect, requiring NP to merge with D in its base position? The answer, I believe, lies in the differing nature of the Case Filter and the selectional requirements of particular lexical items. The requirement for an NP to receive case is syntactic, but the requirement for a P_R to have a TIME- or PLACE-containing complement is presumably semantic (though cf. fn. 15). A P_R demands a DP complement endowed with a certain kind of meaning, not a certain structural configuration: there is no obvious reason to believe that an interval and an event DP, for example, are structurally distinct. It seems that while the grammar can regulate interactions among purely syntactic constraints, it cannot regulate interactions among syntactic constraints and those that interface with the semantics.

This statement, though, will ultimately, need to be qualified. While it is impossible to overrule WLM in order to satisfy the selectional restrictions of P_Rs, it is possible to overrule WLM to bind a variable. Note here that our definition of LATEMERGE predicts that, if a NP must merge early to bind a variable, it should merge in the *highest position where the variable containing the NP can be bound*. This is consistent with the P-stranding facts: even when NP is merged early to establish a variable binding dependency, it must still be merged above the base position (i.e. in Spec, p), because a P_R still cannot be stranded in a B-type extraction.

---

[21]In this derivation, it is crucial that case assignment can follow movement of the NP to Spec, p. Although we know from much recent work that case does not *have* to be assigned in a Spec-head relationship, I assume here that it can be. More precisely, when a probe with an EPP property checks case on its goal, case assignment can either precede or follow movement of the goal. In derivations where WLM is involved (as in (54)), case assignment will always follow movement, as this allows for maximal satisfaction of LATEMERGE.

[22]Note that positing a pP shell makes it possible for a DP to move successive cyclically through Spec, pP without violating Abel's (2003) stranding generalization.

(55) Variable binding in B-type extractions

    a.   * The river of his$_i$ ancestors, [every boy]$_i$ ate lunch on$_R$.

    b.   * His$_i$ mother's favorite month is easy for [every boy]$_i$ to swim in$_R$.

    c.   * Not a single one of his$_i$ favorite holidays will I eat [any butcher's]$_i$ lamb on$_R$.

As we saw earlier, WLM can also be overruled to obtain non-surface scope, or satisfy the semantic properties of certain verbs (§5). These facts suggest that some semantically relevant constraints, but not others, can overrule syntactic ones. I do not have an explanation for why this partition exists, but it is worth noting that a similar asymmetry – that some semantic constraints are more important than others – has been discovered in work by Fox (1998)[23].

7.2 Deriving A-type extractions

A-type extractions pose an interesting problem for the current analysis. The fact that A-type extractions permit a P$_R$ to be stranded, as seen in (56), suggests that some factor blocks WLM entirely. For a P$_R$'s selectional restrictions to be satisfied, the NP must merge as part of P's complement.

(56) A-type extractions can strand P$_R$s

    a.   Which floor are we eating dinner on$_R$ [~~which floor~~]?

    b.   Which month did you vacation in$_R$ [~~which month~~]?

To explain why the examples in (56) are good, we must identify another constraint, specific to A-type extractions, that overrides the preference for WLM. To identify this constraint, we need to find some property that the A-types share. More than this, it must also be some property that all B-types lack.

One possibility is that the moving DP in A-type extractions is embedded within an additional layer of structure: a QP, in Cable's (2007, 2010) sense. Though the bulk of Cable's (2007, 2010) discussion focuses on *wh*-movement, he notes that restrictive relatives and free relatives in English abide by similar constraints, and posits that the moving constituents in these extractions are embedded within a larger QP shell[24]. He speculates (2007: 369-375) that a Q-based analysis could be extended to all other types of A'-extraction. My proposal is more restrictive: the moving DP in the A-type A'-extractions, but not the B-types, are embedded within a QP shell.

Some evidence that this is a promising line of analysis comes from a difference in acceptable pied-piping size between the A-type extractions and appositive relatives (a B-type extraction). English is an example of what Cable terms a *limited pied-piping language*: in *wh*-movement, very little material can be pied-piped along with the *wh*-phrase. This limitation is illustrated in (57). The sentence in (57a), where the *wh*-phrase is embedded within a larger DP, is much less acceptable than (57b), where it is not.

---

[23]The relevant examples are ones in which inverse scope is generally impossible for reasons of Scope Economy and parallelism, but permitted when necessary to bind a variable. The following contrast, from Fox 1998: 39, demonstrates:

  (i)   Someone in the audience knows the capital of every country. The person who was invited to talk about it does too. (*someone >every, every >someone*)

  (ii)   Someone in the audience knows the capital of every country. The person who was invited to talk about these countries does too. (*someone >every, *every >someone*).

In (i), the inverse scope reading is licensed by the necessity for the pronoun *it* (in the second sentence) to be bound by *every country* (in the first). In (ii), the inverse scope reading is impossible when there is no need to bind a variable. See Fox 1998, especially pp. 37-45, for more discussion.

[24]The Qs involved in *wh*-questions and restrictive relatives are presumably different kinds of Q. Cable dubs the *wh*-variety Q$_Q$ and the relative variety Q$_{REL}$. What matters here is only the presence of the extra layer of structure.

(57) Pied piping in *wh*-movement[25]

    a. ?? Pictures of who do you display in your home?

    b. Who do you display pictures of in your home?

Cable (2007: 279) argues that restrictions on pied-piping size, as in (57), are due to the fact that Q-*wh* agreement is obligatory in limited pied-piping languages (like English). In these languages, lexical categories, like N, cannot intervene between Q and a *wh*-word. This restriction on pied piping size also holds for restrictive relatives (58): compare ungrammatical (58a) to grammatical (58b).

(58) Limited pied piping in restrictive relatives

    a. * The scientist pictures of whom I display in my home won a Nobel Prize.

    b. The scientist who I display pictures of in my home won a Nobel prize.

Appositive relatives and other B-type extractions, however, allow what Heck (2004) terms 'massive pied piping' (see also Bresnan 1976, Emonds 1976, Jackendoff 1977, and other more recent work cited by Cable 2007: 341). While pied piping in *wh*-movement and other A-types is limited, pied piping in B-types is not. Examples of massive pied piping in appositive relatives (59a) and negative inversion (59b) follow; in (59b), I assume that the negative phrase *no famous scientist* is targeted for movement, and *pictures of* is pied-piped along with it.

(59) Massive pied piping in B-types

    a. Marie Curie, pictures of whom I display in my home, won a Nobel Prize. *appositive relative*

    b. Pictures of no famous scientist would I ever display in my home. *negative inversion*

The grammatical sentences in (59) blatantly violate Q-*wh* agreement. Rather than trying to explain why the A-type extractions appear to require Q-*wh* agreement and the B-types do not, it is perhaps simpler to say that the moving DP in the A-types, but not the B-types, is embedded within a QP shell (though see Heck 2004, 2008 and Cable 2007: 341-348 for alternatives). Under this account, massive pied piping and apparent violations of Q-*wh* agreement can occur in B-types because there is no Q. To be clear, this proposal entails that movement in A- and B-type extractions is triggered by different things. In A-type extractions, the probe carries uninterpretable Q features, and searches for interpretable Q features on its goal (Cable 2007, 2010), meaning that a DP *must* be embedded in a QP shell for movement to occur[26]. In B-type extractions, however, the probe looks for other kinds of interpretable features that are perhaps extraction-specific: a Topic feature for topicalization, for example, or a Neg feature for negative inversion.
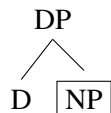
Why, though, would the presence of a QP shell in A-type extractions block WLM? In line with Tada's (1993: 63-70) modification of Lebeaux's (1988) theory of late merge, I propose that WLM cannot apply in A-type extractions because it would be *too* countercyclic. WLM is permitted when the NP merges to the daughter of the current root node, but prohibited when its application would modify the internal structure of the daughter's complement. In (60a), WLM is permitted because NP merges with D, the daughter of the current root node. In (60b), WLM is forbidden, because NP merges inside DP embedded within a QP.

---

[25]Some speakers report that the contrast in (57) is not strong. What's important is that there is a contrast in (57), and that it disappears for the corresponding sentences with appositive relatives. Note also that in the most natural pronunciation of (57a), the *wh*-word is focused ('pictures of *who*...?') as if it were a *wh*-in-situ question ('you displayed pictures of *who*...?'). Without focusing the *wh*-word, as is usual for *wh*-questions with movement, (57a) is much less natural.

[26]Concretely, I assume that there is a constraint requiring agreement between the probe and the goal (AGREE), and that AGREE dominates LATEMERGE. We would expect that other constraints can dominate AGREE, causing agreement to fail (see Preminger 2011). What remains to be worked out under this approach is a theory of the constraints that can interact with, and dominate, AGREE (though cf. Preminger: 130ff on violable constraint models).

(60)   WLM in A-types is too countercyclic

    a.   ✓WLM inside DP (B-types)   b.   *WLM inside QP (A-types)

        DP   QP

      D NP   Q   DP

                                           D NP
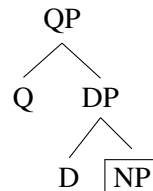
The proposed restriction on WLM bears a striking resemblance to restrictions on late adjunction described by Sauerland (1998: 2.2). Sauerland shows, in brief, that there is an ordering effect among relative clauses, predicted by Tada's (1993: 63-70) modification of Lebeaux's (1988) proposal. While it is possible to force reconstruction of an inner modifier without forcing reconstruction of an outer modifier, the reverse is not possible: forcing reconstruction of an outer modifier forces reconstruction of an inner modifier as well. Compare, for example, (61a) and (61b) (from Sauerland 1998: 52):

(61)   Sauerland (1998) and the timing of late merge

    a. [Which computer <u>compatible with his$_j$</u> that Mary$_k$ knew how to use]$_i$ did she$_k$ tell every boy$_j$ to buy t_i?

    b. [Which computer compatible with Mary$_k$'s <u>that he$_j$ knew how to use</u>]$_i$ did she$_k$ tell every boy$_j$ to buy t_i?
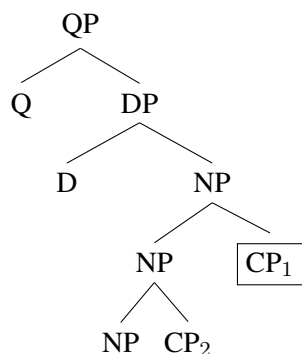
In (61a), the inner relative clause (underlined) must merge at t_i for *every boy* to bind *his*. It is not necessary for the outer relative clause (*that Mary knew how to use*) to merge at t_i, so a Principle C violation is avoided through late adjunction. In (61b), however, reconstruction of the outer modifier is forced: the underlined clause must merge at t_i for *every boy* to bind *his*. The fact that a Principle C violation is present in (61b) suggests that if an outer modifier is forced to reconstruct to bind a variable, the inner modifier must reconstruct as well. This suggests that it is impossible for an inner modifier to merge countercyclically between a noun and another modifier; that is, the inner modifier must merge derivationally prior to the outer modifier. See Sauerland 1998: 2.2 for more examples and discussion.

Though the sources of these two effects might ultimately be unrelated[27], we could achieve a unified account by stating the constraint on excessive countercyclicity in relational, rather than absolute, terms. Consider again the difference between the structure in (60a), where WLM is permissible, and (60b), where it is not. In the impossible (60b), notice that the NP that has undergone WLM is asymmetrically c-commanded by an element, Q, which already stands in a complementation relation with DP. Now consider the difference between the structures described by Sauerland: the acceptable configuration where an an outer but not an inner modifier undergoes late merge (62a), and the ungrammatical configuration where an inner but not an outer modifier undergoes late merge (62b). Below, the CP that has undergone late merge is boxed.

---

[27]If the speculation that follows is ultimately incorrect, we are left with a contradiction: it is possible for Lebeaux-style late adjunction to target an embedded constituent, but not for WLM to do so. I leave open the possibility that the potential for countercyclic merge of complements is just more limited than that of adjuncts. We know already that there is an asymmetry between late merger of complements and adjuncts: while late adjunction is always interpretable, late merger of a complement is not always interpretable (cf. Fox 2002: 69). Perhaps they are also differentiated according to the degree of countercyclicity allowed.

(62) Embedded late adjunction is too countercyclic

    a. ✓Outer modifier merges late   b. *Inner modifier merges late
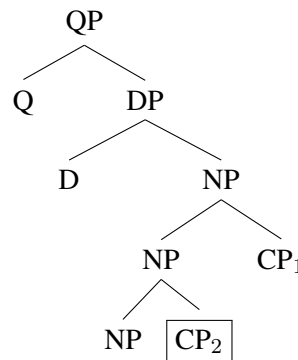
In the acceptable (62a), the modification relation established by late merge is the structurally highest relationship of its type: CP_1 is not asymmetrically c-commanded by another element that is either a modifier or sister to a modifier. In the unacceptable (62b), however, the modifier that has undergone late merge is c-commanded by an element, NP, already standing in a modification relationship with CP_1. To unify the restrictions on WLM observed here with those on late adjunction observed by Sauerland, we could say that countercyclic merger is permitted *only* when the relationship established is the structurally highest of its type (i.e. complementation vs. modification) within the current root node. In other words, if a complement undergoes WLM, it cannot be asymmetrically c-commanded by another element in a complementation relationship; if a modifier undergoes Late Merge, it cannot be asymmetrically c-commanded by another element in a modification relationship. (60b) is unacceptable, then, because the late-merged complement is asymmetrically c-commanded by Q, which takes a DP complement; (62b) is unacceptable because the late-merged adjunct, CP_2 is asymmetrically c-commanded by CP_1, which is a modifier of NP.

Let's now integrate the above discussion into the analysis. The A-type extractions show us that the constraint requiring WLM to create the structurally highest relationship of its type (*TooLate) dominates LateMerge, as WLM of NP to D within a QP is impossible. A possible definition for *TooLate, based on the above discussion, is given in (63). The desired result is in (64): WLM is blocked when it would result in a violation of *TooLate.

(63) *TooLate: a * if the relationship established by late merge isn't the structurally highest of its type.

(64) Tableau demonstrating *TooLate » LateMerge

	a.	b.
*TooLate	*!	
LateMerge	n	$n+1$

Candidate (64a), where WLM applies, incurs a violation of *TooLate: NP merges with D, creating a complementation relation that is not the structurally highest of its type. Even though candidate (64b), where WLM is blocked, incurs more violations of LateMerge, it is selected as optimal. NP must merge with the D before the DP merges with Q; otherwise, the derivation fatally violates *TooLate[28]. This modification to the analysis allows us to derive the Principle C and stranding properties of A-type extractions. WLM is blocked due to the presence of a QP shell, so the NP merges in the base position. Because A-type extractions leave behind a full copy, Principle C effects are present, and P_Rs can be stranded.

Note that this modification to the analysis makes a broader prediction: *all* movement of *wh*-phrases (and other elements in the same category), not only A-extractions to A'-positions, should exhibit reconstruction effects. This is not the case: when A-movement is followed by *wh*-movement, Principle C is bled (e.g. *Which of John$_i$'s siblings seems to him$_i$ to be smart?*). We can solve this problem by allowing for the possibility that Q, like NP, prefers to merge late, and by making two additional claims. First, the moving DP in an A-type A'-extraction must be embedded within a QP shell before movement begins, as Q is targeted for movement in these extractions (see argumentation in Cable 2007, 2010). On the other hand, I propose that the moving DP in A-movement need not be embedded within a QP shell. The result is that while Q prefers to be merged late, it must be merged immediately prior to the first step of A'-movement. This allows for the possibility that WLM is impossible in an A-type A'-extraction, but possible when A-movement feeds A'-movement[29].

7.3 Further predictions

The interaction of LateMerge and GetCase makes a very basic prediction: all else equal, WLM should occur at the last position in which the NP can receive Case. Here I expand the body of evidence supporting this prediction by investigating properties of other extraction types out of different domains. We will begin with the English pseudopassive, a type of A-movement that can strand prepositions. Several examples are provided in (65).

(65) Examples of the pseudopassive

    a.   The bed was slept in.

    b.   I was talked about.

    c.   The desk was written on.

The only crucial assumption I will make regarding the mechanics of the pseudopassive is that, as is the case for the "true" passive, the moved DP receives Case not from P or *p*, but from the matrix T (cf. Abels 2003: 234). This is evident in (65b) from the fact that the subject *I* bears nominative case, not accusative,

---

[28]There is an additional candidate, in which only Q merges as P's complement, and DP merges to Q after Q moves to Spec, *p*P. As noted by an anonymous reviewer, this derivation does not incur a violation of *TooLate, and incorrectly predicts that WLM should occur in A-type extractions. It's not immediately obvious, though, that the result of the operation just described would be semantically interpretable. WLM of NP is possible because what is left behind – a quantificational determiner – can be converted to a definite description through Trace Conversion. Cable's Q serves as a variable over choice functions (see also Hagstrom 1998, Yatsushiro 2001); it is unclear that the operations of variable insertion and determiner replacement are even applicable here, or, if they are, that they would result in a semantically interpretable structure.

[29]An anonymous reviewer notes a potential loophole in this proposal: suppose it is possible for DP to move to Spec, *p* for Case reasons, and for (i) late merge of Q and (ii) A'-movement to follow. Given the preference for WLM, this would predict that late merge of Q should be the desired state of affairs, and furthermore that stranding a P_R should be impossible for all extractions, even A-types. We can avoid this unwanted prediction by asserting that, at least in English, Case alone cannot trigger movement to Spec, *p*; *p* lacks an EPP feature. It just so happens, however, that if a DP lands in Spec, *p* on its way up, Case can be assigned there. Under this analysis, there is no reason for a DP to land in Spec, *p* unless Spec, *p* is an intermediate A'-step – in which case, in A-type extractions, the Q element must be present. See Kayne 1984:2ff for a similar argument from case assignment into embedded Comp: a DP cannot land in embedded Comp unless it has other (non-case-related) reasons for being there.

as would be assigned by the preposition *about* (compare **Me was talked about* to *They talked about me*). In addition, I assume that movement to Spec, T in the psuedopassive is driven by the same factors that drive movement in the true passive: the standard analysis is that the lower DP moves to check case (e.g. Baker, Johnson & Roberts 1989: 222).

If T assigns case to the moving DP, then the preference for late merge dictates that NP must merge to its head D in Spec, T. Because the NP undergoes WLM, i.e. is not present in the base position, this analysis predicts that the pseudopassive should behave like a B-type extraction: stranding a P_R should be impossible. This is in fact the case. In the locative examples[30] below, we see that pseudopassivization is possible when the preposition is entity-selecting (66), but not when it is location-selecting (67).

(66) Pseudopassives licit for entity-selecting Ps

    a.   The house was lived in.
        cf. We lived in {the house, it}.

    b.   Boxes were peered into.
        cf. We peered into {boxes, them}.

    c.   Mice were stepped on.
        cf. We stepped on {mice, them}.

    d.   The wall was written on.
        cf. We wrote on {the wall, it}.

(67) Pseudopassives banned for location-selecting Ps (* = lack of a location reading)

    a.   * The second floor was worked on.
        cf. We worked on {the second floor, *it}.

    b.   * The sky was flown into.
        cf. We flew into {the sky, *it}.

    c.   * The Charles River was dined on.
        cf. We dined on {the Charles River, *it}.

    d.   * The fourth floor was celebrated on.
        cf. We celebrated on {the fourth floor, *it}.

This is the pattern observed for B-type extractions. Because the moving NP is not present in the base position, the complement of a location-selecting P does not contain an element of the semantic category PLACE. In B-type extractions and pseudopassivization alike, a P_R cannot be stranded[31].

I turn now to A'-extraction of an object out of *v*P. Recall the claim in this section that case assignment can be procrastinated until Spec, *x*, where *x* is the relevant case-assigner. Assuming that accusative case is

---

[30]I focus on locatives because stranding a temporal preposition under A-movement is generally degraded, regardless of whether it is a P_R (**Christmas was eaten lamb on*) or a P_A (**John's party was left after*).

[31]An anonymous reviewer notes that we would expect for pseudopassives of the sort in (67) to become grammatical when the subject contains a bound variable, as the constraint penalizing a failure to bind a variable will force NP to merge in the complement of a pronoun-rejecting P. However, this is not the case: even when the bound variable reading is possible, stranding a P_R is impossible (i-ii).

    (i)   *[The floor that his$_i$ mother liked] was jumped on by [every child]$_i$.

    (ii)   *[The river that his$_i$ mother swam in] was dined on by [every politician]$_i$.

If however the *by*-phrase is merged in Spec, *v*P (cf. Collins 2005: 84-85), and the NP restrictor can undergo WLM at the VP-adjoined position (cf. Takahashi & Hulsey 2009: 401), the data in (i-ii) are exactly what we would expect: reconstruction is forced to the matrix VP-adjoined position, where the variable *his* can be bound, but not all the way to the base position. Because the NP is not present in the base position, a P_R cannot be stranded.

assigned by v, the analysis predicts that WLM of an object NP can be delayed until its head D reaches Spec, v. This prediction has potential implications for our analysis of movement out of double object constructions: if the NP *report on John's student* undergoes WLM at Spec, v, we would expect for (68a) to be grammatical but (68b) to be ungrammatical (judgments suppressed, for now).

(68) a. That report on John$_i$'s student, Mary showed him$_i$.

b. That report on John$_i$'s student, he$_i$ showed Mary.

We might expect to find a contrast here because the position of the coindexed pronominal differs. In (68a), the coindexed pronominal is VP-internal. If the NP *report on John's student* undergoes WLM at spec, vP, then *him* should never c-command *John*, and a Condition C violation should be absent (69a). If however the coindexed pronominal resides in Spec, vP, *he* will c-command *John*, and a Condition C violation should be present (69b). (I assume that movement of the object to an inner Spec, v must precede Merge of the external argument to an outer Spec, v; see Müller 2010.)

(69) Positioning of the coindexed pronominal matters

a. Coindexed object: WLM applies above c-command domain
$[_{vP}$ Mary $[_{vP}$ that <u>report on John$_i$'s student</u> $[_{VP}$ showed him$_i$ ~~that~~ $]$ $]$ $]$ $]$

b. Coindexed subject: WLM applies within c-command domain
$[_{vP}$ He$_i$ $[_{vP}$ that <u>report on John$_i$'s student</u> $[_{VP}$ showed Mary ~~that~~ $]$ $]$ $]$ $]$

What we find, however, is that both of the sentences in (68) are ungrammatical: a Condition C violation results regardless of whether the coindexed pronominal is a subject or an object. Rather than posing a problem for the current analysis, however, the failure of (69a) to bleed Condition C has the potential to teach us something about the structure of double object constructions. Pesetsky (1996: 123ff) argues that the second argument in a double object construction does not receive Case from any verbal projection; rather, it is introduced and Case-licensed by a null preposition, called here P (see also Harley 2002 for a similar proposal). Assuming this analysis, WLM of the NP *report on John's student* can occur no later than Spec, PP – a position which is crucially *within* the c-command domain of either a coindexed object (70a) or subject (70b).

(70) Positioning of the coindexed pronominal doesn't matter

a. Coindexed object: WLM applies within c-command domain
$[_{vP}$ Mary $[_{vP}$ that report … $[_{VP}$ showed him$_i$ $[_{PP}$ ~~that report on John$_i$'s student~~ $]$ $]$ $]$ $]$

b. Coindexed subject: WLM applies within c-command domain
$[_{vP}$ He$_i$ $[_{vP}$ that report … $[_{VP}$ showed Mary $[_{PP}$ ~~that report on John$_i$'s student~~ $]$ $]$ $]$ $]$

We can therefore interpret the failure of (68a) to bleed Condition C not as a problem for the current analysis, but rather as an indication that the analysis proposed above, following Pesetsky and Harley, is on the right track. If we assume that the second object in a double object construction is introduced and assigned Case by a null prepositional element, then the failure of both examples in (68) to bleed Condition C is not surprising.

8   Further support: B$_2$-type extractions, and more pronouns

The analysis proposed in §7 claims that the link between a P's selectional restrictions and its potential to be stranded is structural. Interval and location P$_R$s reject all DP complements, pronounced or not, that do not contain TIME or PLACE. This approach makes a prediction: if P's complement does not contain an NP for other reasons, stranding a P$_R$ should be impossible, regardless of the extraction in question.

Here I verify this prediction with evidence from two domains. §8.1 returns to the properties of B_2-type extractions, and argues that their behavior is best explained as the result of null operator movement. Because P's complement never contains an NP, a P_R cannot be stranded. §8.2 shows that the analysis correctly predicts contrasts *within* individual extraction types. When *wh*-movement involves movement of *wh*-pronouns, *wh*-movement acts like a B-type extraction: it is sensitive to whether or not a given P is pronoun-rejecting.

8.1 B_2-types and null operators

If we assume that the B_2-type extractions (*tough*-movement, gapped degree phrases, and parasitic gaps) are structurally similar – that is, they all involve movement of a null operator – then the fact that they bleed Principle C and cannot strand P_Rs is predicted by the current analysis. Null operators do not contain lexical material (i.e. NPs), so there is nothing in their structure that could satisfy the selectional restrictions of P_Rs, or drive a Principle C effect. But the assumption that these three extractions all involve null operator movement is not universally accepted. Below I briefly sketch the basis of the controversy, and provide some evidence in favor of the null operator style analyses.

While it is common to propose that parasitic gaps involve movement of a null operator or some other null pronominal element (see Nissenbaum 2000: 24-29 for an overview), the analysis of *tough*-movement, and gapped degree phrases, is more controversial. The existence of defective intervention effects (see Hartman 2009, also Brillman 2014) and the presence of reconstructed bound variable readings support an improper movement analysis, in which the moving constituent originates in the embedded clause. The absence of scope reconstruction from the TC subject into the embedded clause (see Postal 1974, Epstein 1989, Fleisher 2013) and the absence of Principle C effects, however, supports an account where the DP originates in the matrix clause, like Chomsky's (1977, 1981) null operator analysis.

It is beyond the scope of this paper to provide a conclusive argument for either improper movement or null operators, nor do I attempt to further differentiate within these two classes of analyses. When we look at interactions between the B_2-type extractions and *wh*-movement, however, some evidence for a null operator analysis of the B_2-type extractions emerges. Stranding P_Rs using a combination of *wh*-movement and a B_2-type extraction is impossible (71a, 72a) and these extractions bleed Principle C (71b, 72b). In short, hybrids of A-type and B_2-type extractions behave like B_2-types.

(71)   Hybrid type 1: *tough-* and *wh*-movement

    a.   * Which holiday is easy to eat lamb on?
        (cf. Which holiday did we eat lamb on?)

    b.   Which criticism of Mary$_i$'s proposal is hard for her$_i$ to accept?
        (cf. *Which criticism of Mary$_i$'s proposal did she$_i$ reject?)

(72)   Hybrid type 2: GDPs and *wh*-movement

    a.   * Which holiday is too joyous to eat lamb on?
        (cf. Which holiday did we eat lamb on?)

    b.   Which criticism of Mary$_i$'s proposal is too harsh for her$_i$ to accept?
        (cf. *Which criticism of Mary$_i$'s proposal did she$_i$ reject?)

These differences between the hybrid extractions and *wh*-movement are predicted by the null operator analysis. If the matrix subject never occupies any position in the embedded clause, then it follows that a P_R cannot be stranded, because its complement cannot contain TIME or PLACE. Similarly, in (71b, 72b) Principle C is bled because *Mary* is never within the c-command domain of co-indexed *her*.

An improper movement style analysis, where the QP originates in the embedded clause, would face some problems in accounting for (71-72). Because the hybrid extractions are *wh*-interrogatives, I assume

that the fronted DP is embedded within a QP shell. An analysis assuming that the QP originates in the embedded clause would need to explain why it is apparently acceptable for excessively late WLM to occur in the hybrid extractions, but not in pure *wh*-movement[32].

8.2 *Wh*-pronouns

We saw in §8.1 that if a given extraction never allows a P_R to have a complement containing TIME or PLACE, that extraction will never allow a P_R to be stranded. The analysis outlined in §7, in turn, suggests that the ability to strand a P_R is linked to the structural configuration of the gap, not to the extraction that creates it. Given this, we would expect, under certain conditions, to find contrasts within extractions. Extractions that are in principle capable of leaving behind full copies should be sensitive to P_Rs if the full copy is pronoun-sized (i.e., a D).

One way to assess this prediction is to compare the behavior of *wh*-movement with *wh*-pronouns and full *wh*-phrases. It is a plausible assumption that *wh*-pronouns like *who* and *what* do not contain NPs, like their non-*wh* counterparts *her* and *it*. *Wh*-movement with *wh*-pronouns should then be sensitive to whether or not a preposition is a P_A or a P_R, while *wh*-movement with full *wh*-phrases should not discriminate. This is because A-type extractions leave behind a full copy, but the full copy of a *wh*-pronoun does not contain an NP, and therefore cannot contain TIME or PLACE.

This prediction holds. The examples in (73) show that movement of both *wh*-pronouns and *wh*-phrases is grammatical when stranding a P_A; the examples in (74) show that movement of *wh*-phrases is grammatical when stranding a P_R, but movement of *wh*-pronouns is not.

(73) Stranding a P_A

   a. My family eats dinner on the green table. John's family eats dinner on it too.

   b. What table does your family eat dinner on?

   c. What does your family eat dinner on?

(74) Stranding a P_R

   a. * My family eats turkey on Thanksgiving. John's family eats turkey on it too.

   b. What holiday does your family eat turkey on?

   c. * What does your family eat turkey on?

Note that the status of (74c) is not due to a restriction on movement: in a multiple question, the phrasal DP *which holiday* but not the pronoun *what* is an acceptable complement to the temporal P_R *on*, as predicted[33]:

(75) *What* vs. *which holiday* in a multiple question

   a. Which person thinks that John's family eats turkey on which holiday?

   b. * Which person thinks that John's family eats turkey on what?

---

[32]An anonymous reviewer suggests an analysis along these lines. As the initial step of A'-movement in *tough*-movement may well not be Q-driven, late merge of Q could be possible – meaning that NP is free to merge after D has left the infinitival clause. For the purposes of this section, the two accounts make identical predictions (Condition C is bled, stranding a P_R is impossible), so the choice between them is not crucial here.

[33]In addition, the *what* vs. *what x* contrast appears under sluicing: compare the ungrammatical **John's family eats turkey on some holiday, but I don't know what* to the grammatical *John's family eats turkey on some holiday, but I don't know which one*. The emergence of this contrast under sluicing is consistent with the analysis of antipronominal contexts as a consequence of selectional restrictions, rather than island-like constraints on movement, which can in some cases be circumvented by sluicing (see Ross 1969 for the original observation, and Merchant 2001 for recent work). Thanks to an anonymous reviewer for pointing this out.

In addition, the ungrammaticality of (74c) cannot be attributed to a more general restriction on the types of DPs that *wh*-pronouns can replace. *Thanksgiving* can be pronominalized with *it* when it is the complement of *celebrate* (76a). In this context, movement of both *wh*-pronouns and *wh*-phrases is permitted (76b-c).

(76)   Interval DPs and *wh*-pronouns

      a.   My family celebrates Thanksgiving, and John's family celebrates it too.

      b.   Which holiday does your family celebrate?

      c.   What does your family celebrate?

The analysis proposed in §7 bears some resemblance to Postal's (1998) account of antipronominal contexts, but the contrasts in (73-74) reveal that the WLM-style analysis makes accurate predictions that Postal's (1998) analysis does not. Postal proposes that certain extractions are sensitive to antipronominal contexts because they leave behind null resumptive pronouns. Because his account proposes that extractions themselves are linked to certain types of null categories, it is unable to predict contrasts within extractions, like the contrast between *wh*-pronouns and *wh*-phrases documented in (73-74). Under an analysis where the ability to strand a P_R is not linked to intrinsic properties of individual extractions, however, but to the structural configuration of the gap, the pattern in (73-74) falls out. The ability to strand a temporal or locative P_R depends entirely on whether or not its complement contains TIME or PLACE. Specific extractions only appear to play a role because their independent properties (e.g. presence of a QP shell) regulate the application and location of WLM.

9   Conclusion

The P-stranding asymmetries and Principle C facts discussed in this paper demonstrate a need to divide A'-extractions into two major groups. I have argued that the differences between these groups arise because they place differing constraints on the application of WLM. The A-type extractions leave behind fully constructed copies, as the application of WLM is blocked. The B-type extractions do not leave behind fully constructed copies, either because WLM is required (B_1-type extractions) or because the extractions in question involve null operator movement (B_2-type extractions). The extractions that leave behind fully constructed copies (A-types) can strand pronoun-rejecting prepositions; the extractions that do not leave behind fully constructed copies (B-types) cannot.

More broadly, all English A'-extractions have the capability to strand prepositions, but differing properties of individual extractions lead to distinct constraints on which Ps can be stranded. Thus while it is true that all A'-extractions share the same set of basic properties (e.g. Chomsky 1977, 1981, 1982; Pesetsky 1982; Engdahl 1983), asymmetries within a single property can point to subtle, but fundamental, differences among them.

References

Abels, Klaus (2003). *Successive Cyclicity, Anti-locality, and Adposition stranding*. PhD thesis, UConn.

Abney, Steven Paul (1987). *The English Noun Phrase in its Sentential Aspect*. PhD thesis, MIT.

Baker, Mark, Kyle Johnson & Ian Roberts. Passive Arguments Raised. *LI* **20**. 219-251.

Bhatt, Rajesh & Roumyana Pancheva (2004). Late merger of degree clauses. *LI* **35**. 1-45.

Bhatt, Rajesh & Roumyana Pancheva (2007). Degree quantifiers, position of merger effects with their restrictors, and conservativity. In Chris Barker & Pauline Jacobson (eds.), *Direct compositionality*, 306-335. Oxford: Oxford University Press.

Bresnan, Joan (1976). On the Form and Function of Transformations. *LI* **7**. 31-40.

Brillman, Ruth (2014). Gapped Degree Phrases are Improper Movement Constructions. Ms., MIT.

Brody, Michael (1993). θ-Theory and arguments. *LI* **24**. 1-23.

Cable, Seth (2007). *The grammar of Q: Q-particles and the nature of Wh-fronting, as revealed by the Wh-questions of Tlingit*. PhD thesis, MIT.

Cable, Seth (2010). *The Grammar of Q: Q-Particles, Wh-Movement, and Pied-Piping*. Oxford: Oxford University Press.

Cardinaletti, Anna & Michal Starke (2004). The typology of Structural Deficiency: On the Three Grammatical Classes. Ms., MIT.

Chomsky, Noam (1977). On wh-movement. In Peter Culicover, Tom Wasow, & Adrian Akmajian (eds.), *Formal Syntax*, 71-132. New York: Academic Press.

Chomsky, Noam (1981). *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.

Chomsky, Noam (1982). *Some concepts and consequences of the theory of Government and Binding*. Cambridge: MIT Press.

Chomsky, Noam (1993). A minimalist program for linguistic theory. In Kenneth Hale and Samuel Jay Keyser (eds.), *The view from Buliding 20: Essays in linguistics in honor of Sylvain Bromberger*, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam (1999). *Powers and prospects: reflections on human nature and the social order*. London: The Electric Book Company Ltd.

Chomsky, Noam (2004). Beyond Explanatory Adequacy. Adriana Beletti (ed.), *The cartography of syntactic structure*, 104-131. Oxford: Oxford University Press.

Cinque, Guglielmo (1990). *Types of A'-Dependencies*. Cambridge, MA: MIT Press.

Collins, Chris (2000). A smuggling approach to the passive in English. *Syntax* **8**. 81-120.

Delorme, Evelyne & Ray C. Dougherty (1972). Appositive NP constructions: we, the men; we men; I, a man; ETC. *Foundations of Language* **8**. 2-29.

Elbourne, Paul D. (2005). *Situations and Individuals*. Cambridge, MA: MIT Press.

Emonds, Joseph (1976). *A Transformation Approach to English Syntax: Root, structure-preserving, and local transformations*. New York: Academic Press.

Engdahl, Elizabet (1983). Parasitic Gaps. *Linguistics and Philosophy* **6**. 5-34.

Epstein, Samuel David (1989). Quantification in null operator constructions. *LI* **20**. 647-658.

Fleisher, Nicholas (2013). On the Absence of Scope Reconstruction in Tough-Subject A-chains. *LI* **44**. 321-332.

Fox, Daniel J. (1998). *Economy and Semantic Interpretation – A study of scope and variable binding*. PhD thesis, MIT.

Fox, Danny (1999). Reconstruction, binding theory, and the interpretation of chains. *LI* **30**. 157-196.

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

Fox, Danny (2002). Antecedent-contained deletion and the copy theory of movement. *LI* **33**. 63-96.

Hagstrom, Paul (1998). *Decomposing Questions*. PhD thesis, MIT.

Harley, Heidi (2002). Possession and the double object construction. *Linguistic variation yearbook* **2**. 31-70.

Hartman, Jeremy (2009). Intervention in Tough Constructions. In *Proceedings of NELS 39*.

Heck, Fabian (2004). *A Theory of Pied Piping*. Doctoral dissertation, Universität Tubingen.

Heck, Fabian (2008). *On pied-piping: Wh-movement and beyond*. Berlin: Mouton de Gruyter.

Hicks, Glyn (2003). *So Easy to Look At, So Hard to Define: Tough movement in the Minimalist framework*. MA thesis, University of York.

Hornstein, Norbert (2000). *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.

Jackendoff, Ray (1977). *X-bar syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Kayne, Richard S. (1984). *Connectedness and binary branching*. Dordrecht: Foris.

Kayne, Richard S. (2005a). Silent Years, Silent Hours. *Movement and Silence*. Oxford: Oxford University Press.

Kayne, Richard S. (2005b). Here and There. *Movement and Silence*. Oxford: Oxford University Press.

Koopman, Hilda & Dominique Sportiche (1983). Variables and the Bijection Principle. *The Linguistic Review* **2**. 139-160.

Kuroda, Sige-Yuki (1964). A note on English relativization. Ms., MIT.

Kuroda, S.-Y. (1968). English Relativization and Certain Related Problems. *Lg* **44**. 244-266.

Lasnik, Howard & Tim Stowell (1991). Weakest Crossover. *LI* **22**. 687-720.

Lebeaux, David (1988). *Language acquisition and the form of grammar*. PhD thesis, UMass Amherst.

Lebeaux, David (1998). *Where does binding theory apply?* Technical report. Princeton, NJ: NEC Research Institute.

Levinson, Lisa (2011). Possessive WITH in Germanic: HAVE and the role of P. *Syntax* **14**. 355-393.

Manninen, Satu Helena (2003).*Small Phrase Layers: A study of Finish Manner Adverbials*. Amsterdam: John Benjamins Publishing Company.

Merchant, Jason (2001).*The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Müller, Gereon (2010). On Deriving CED Effects from the PIC. *LI* **41**. 35-82.

Nikolaeva, Liudmila (2014). *The secret life of pronouns*. Ph.D. thesis, MIT.

Nissenbaum, Jonathan W. (2000). *Investigations of covert movement*. PhD thesis, MIT.

Oxford, Will (2011). The syntax of Innu-aimun locatives. In *UBCWPL 31: Proceedings of WSCLA 16*.

Perlmutter, David M. (1972). Evidence for shadow pronouns in French relativization. In Paul M. Peranteau, Judith N. Levi & Gloria C. Phares (eds.), *The Chicago which hunt*. CLS: University of Chicago.

Pesetsky, David (1978). Category switching and so-called so-called pronouns. *CLS* **14**. 350-361.

Pesetsky, David Michael (1982). *Paths and categories*. PhD thesis, MIT.

Pesetsky, David (1996). *Zero Syntax*. Cambridge, MA: MIT Press.

Postal, Paul M. (1966). On so-called 'pronouns' in English. *Monograph series in language in linguistics*, **19**. 177-206.

Postal, Paul M. (1974). *On raising*. Cambridge, MA: MIT Press.

Postal, Paul M. (1998). *Three investigations of extraction*. Cambridge, MA: MIT Press.

Preminger, Omer (2011). *Agreement as a fallible operation.* PhD thesis, MIT.

Prince, Alan & Paul Smolensky (1993 [2002]). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University Center for Cognitive Science Technical Report **2**.

Ramchand, Gillian & Peter Svenonius (2004). Prepositions and External Argument Demotion. In Torgrim Solstad, Benjamin Lyngfelt, & Maria Filiouchkina Krave (eds.), *Demoting the Agent: Passive and other Voice-related Phenomena*, 93-99. University of Oslo.

Ross, John Robert (1967). *Constraints on variables in syntax*. PhD thesis, MIT.

Sauerland, Uli (1998). *The meaning of chains*. PhD thesis, MIT.

Smyth, Herbert Weir (1956). *Greek Grammar*. Cambridge: Harvard University Press.

Sportiche, Dominique (2006). Reconstruction, binding, and scope. In Martin Everaert & Henk van Riems-dijk, eds., *The Blackwell companion to syntax: Volume IV*, 35-93. Oxford: Blackwell.

Svenonius, Peter (2003). Limits on p: filling in holes vs. falling in holes. *Nordlyd* **31**: 431-445.

Svenonius, Peter (2008). Projections of P. In Anna Asbury (ed.), *Syntax and Semantics of Spatial P*, 63-84. Philadelphia: John Benjamins Publishing.

Svenonius, Peter (2010). Spatial P in English. In Guglielmo Cinque & Luigi Rizzi (eds.), *Cartography of Syntactic Structure*. Oxford: Oxford University Press.

Tada, Hiroaki (1993). *A / A-bar partition in derivation*. PhD thesis, MIT.

Takahashi, Shoichi (2006). *Decompositionality and Identity*. PhD thesis, MIT.

Takahashi, Shoichi & Sarah Hulsey (2009). Wholesale Late Merger: Beyond the A/A' Distinction. *LI* **40**. 487-526.

Uriagereka, Juan (1995). Aspects of the syntax of clitic placement in Western Romance. *LI* **26**. 79-123.

Van Riemsdijk, Henk (1990). Functional Prepositions. In Harm Pinkster & Inge Genee (eds.), *Unity in Diversity: Papers Presented to Simon C. Dik on his 50th Birthday*. Berlin: Walter de Gruyter.

Yatsushiro, Kazuko (2001). "The Distribution of *Mo* and *Ka* and its Implications." In Cuervo, Maria Christina, Daniel Harbour, Ken Hiraiwa, and Shinichiro Ishihara (eds.), *Formal Approaches to Japanese Linguistics* **3**. MIT Working Papers in Linguistics: Cambridge, MA.

Zribi-Hertz, Anne (1984). Orphan Prepositions in French and the Concept of 'Null Pronoun'. Bloomington, IN: Indiana Linguistics Club.