How to get off an island*

Chris O'Brien

September 27, 2015

Introduction

MERGE-based theories of syntax allow for a dizzying array of structures to be built in which a syntactic object is simultaneously dominated by two or more mothers. Because such a syntax will almost inevitably over-generate, it becomes crucial to determine how the requirements imposed on the grammar by the PF and LF interfaces might constrain the overall system.

This paper is about how a multi-dominant syntax, when combined with the appropriate theory of PF and LF interpretation, can be used to explain a variety of asymmetries in extraction from islands, dubbed the *selective island* phenomena (Cinque, 1990; Postal, 1998). Adapting certain insights from Bachrach and Katzir (2009), I argue that the cyclic interpretation of multi-dominant structures allows for the Spellout of shared syntactic material to be delayed, at least at certain domains for cyclic Spellout. In other words, at certain points in the derivation shared material is effectively invisible to the PF and LF interfaces. This invisibility not only allows for movement to avoid island effects, but also predicts various limitations on island obviation imposed at the LF interface. These semantic limitations will, in turn, predict the selective island effects.

Selective islands

Sentences which violate a syntactic island condition differ in the degree to which they are judged unacceptable. There are a variety of such *selective island* phenomena (Cinque, 1990; Postal, 1998). For instance, Postal (1998) observes that syntactic environments which forbid pronouns (1a) also render extraction from islands unusually difficult (1b):

*Thanks to Irene Heim, Kyle Johnson, Michael Kenstowicz, Richard Larson, Jim McCloskey, Juliet Stanton, Tim Stowell, the audience at MIT's Ling Lunch, GLOW 38, my classmates in MIT's Spring 2015 workshop in linguistics, and especially Danny Fox and David Pesetsky for helpful comments and discussion while writing this paper. Also thanks to Anthony Brohan for suggesting *Gilligan's Island* themed example sentences. All errors are my own.

(1) a. Gilligan speaks Hungarian, and Mary Ann speaks (*in) it, too.
b. That's the language that Gilligan knows someone who speaks ?(*in) _.

Additionally, Cinque (1990) and Postal (1998) observe that DPs (2a) are more amenable to extraction from islands than non-DPs (2b):

(2) a. ?Which country do people who who visit _ often wind up in the hospital?
b. *How sick do people who look _ often wind up in the hospital?

Finally, it has also been observed that islands block scope reconstruction (Longobardi, 1987; Cinque, 1990; Ruys, 2011):

(3) a. How many people do you think I should talk to?
b. How many people do you know a man who should talk to?

(3a) has a narrow scope reading with respect to the modal *should* for *many people* which is unavailable in (3b).

Both Cinque and Postal provided a *resumptive pronoun* analysis for this data. In one version of this analysis, a derivation is available in which *wh*-phrases are base-generated in a position above an island, and anaphorically bind a potentially null pronominal inside of the island. This is illustrated below for *which book does John know the man who wrote _?*, a sentence which violates the complex NP island constraint (Ross, 1967):

(4) which book$_i$ does John know ~~which book~~$_1$ [$_{island}$ the man who wrote pro$_i$]?

The island effect is ameliorated on this analysis because there has actually been no extraction out of the island. This derivation is assumed to be virtually identical to one with an overt resumptive pronoun, which also ameliorates island effects:

(5) ?Which book$_i$ does John know a man who wrote it$_i$?

One question raised by this analysis is why English would allow null pronouns in a sentence like (4), even though it does not normally allow them in such a position:

(6) Every boy$_i$ knows a man who talked about *pro$_i$/him$_i$.

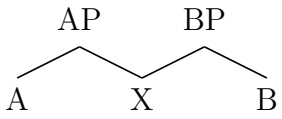
In Cinque's version of the analysis, this contrast is explained by a constraint on empty categories based upon the framework of Chomsky (1982), which is not easily translatable into other theories. Furthermore, this analysis assumes that there are two ways of generating movement structures: one which involves an actual movement operation, another which involves binding a potentially null pronoun. We might ask why this should be so, and whether the data in (1)-(3) could be accounted for in another way.

The proposal

This paper will argue that the selective island phenomena are ultimately a consequence of an operation like Citko's (2005) *parallel merge* (PM) and the manner in which structures formed by PM are interpreted at the PF and LF interfaces.

PM allows a phrase to be shared between two non-c-commanding positions. This is illustrated below:

(7)



Bachrach and Katzir (2009) have argued for a cyclic Spellout algorithm which allows Spellout of shared material to be delayed. A node X is only visible for Spellout, according to Bachrach and Katzir, if the root of the Spellout domain (SOD) dominates X in *every* position that X is merged. If a phrase is shared both inside of and outside of a given SOD, that phrase will avoid Spellout within that SOD. When combined with the hypothesis that islandhood is related to cyclic Spellout (Fox and Pesetsky, 2005, 2007, a.o.), this predicts that syntactic sharing should permit island obviation.

Johnson (2012, 2014) argues that movement structures can be built up in a manner which exploits PM. For instance, according to Johnson, successive cyclic movement of a DP involves sharing of a single NP between two null determiner heads. This simplifies the grammar, since it allows for the LF structures created by Fox's (1999; 2003) *trace conversion* rule to be generated within the derivational syntax.

Consider now what happens when Bachrach and Katzir's delayed Spellout mechanism is integrated into Johnson's theory of syntax. At first glance, the consequences appear to be pathological. Because PM is a basic part of movement for Johnson, the delayed Spellout mechanism will allow delayed Spellout, and hence island obviation, to become a possibility in virtually all movement structures. At least for certain types of movement, there should be no such thing as an island.

This paper will argue that, once the semantic consequences of delayed Spellout are taken into account, this prediction is actually desirable. If one assumes that cyclic Spellout returns both a PF and LF interpretation (Chomsky, 2000), then a phrase which avoids Spellout in a given syntactic position will not only fail to be pronounced in that position, but will also avoid semantic interpretation there. Hence island obviation, if achieved via a delayed Spellout mechanism, should also impose restrictions on reconstruction of a moved phrase into the island. The selective islands, as we shall see, are simply those configurations in which the restrictions on reconstruction imposed by delayed Spellout would lead to problems with semantic interpretation.

Selective islands and ATB-movement

Bachrach and Katzir (2009) also used delayed Spellout in a multi-dominant syntax to explain island obviation effects. However, they assumed a syntax in which PM is only used for certain constructions, such as right-node raising and ATB-movement. This predicts not only that RNR is insensitive to islands (Postal, 1993, 1998), but also an asymmetry between ATB- and non-ATB-movement: ATB-movement should ameliorate island effects. This turns out to be correct:

(8) a. ??Who did Gilligan talk to someone who admires _?
b. Who did Gilligan talk to someone who admires _ and Mary Ann go swimming with someone who can't stand _?

We will refer to sentences like (8b) as *symbiotic gap* (SG) constructions.

However, Bachrach and Katzir's analysis misses an important generalization. While island effects are indeed ameliorated in ATB-structures, they remain subject to all of Cinque and Postal's selective island effects. This is illustrated in (9) below with a sentence that violates the same anti-pronominality restriction as (1):

(9) That's one language that Gilligan knows a man who speaks *(in) _ and Mary Ann also knows a woman who speaks (*in) _.

As we shall see, the analysis developed in this paper accounts for this. Because both ATB movement and ordinary extraction make use of PM, the explanation for selective island effects in ordinary extraction will naturally carry over to ATB-movement.

This is not, then, a paper about why islands exist. Rather, it is about why, given the existence of islands, the set of asymmetries that we see in (1)-(3) arises in both ATB and non-ATB contexts. I will suggest that the unacceptability of sentences like (8a) follows from an additional constraint on derivations which prefers for MERGE to apply as quickly as possible. Hence as soon as a *wh*-DP enters a workspace where it can Merge into a larger structure, there will be a requirement that it do so. This constraint will necessarily be violated in a configuration like (8a).

The paper will be organized as follows. **Section 1** will lay out the data. In **section 2**, we will will describe Johnson's theory of movement in some detail. **Section 3** will develop our conception of cyclic Spellout. **Section 4** will show how these two proposals conspire to predict island obviation. Finally, **sections 5** and **6** will discuss some potential problems for our proposal, as well as additional predictions.

1 Selective islands

1.1 The basic phenomenon

The selective island phenomenon (Cinque, 1990; Postal, 1998) reveals that all cases of extraction from island environments are not created equal. For instance, the awkwardness of extraction from an island is magnified when the extracted element is

a non-DP. We illustrate this below with two cases of extraction which violate the subject and complex NP island condition:

(10) a. ?Which country do people who who visit _ often wind up in the hospital?
b. *How sick do people who look _ often wind up in the hospital?

The case of non-DP extraction (10b) is judged less acceptable than DP extraction (10a) (Cinque, 1990; Postal, 1998).

Cinque and Postal also observe a correlation between the ability of a given syntactic context to accept pronouns and extractability from that context across an island boundary. This is illustrated below:

(11) a. The Skipper can speak Swedish and The Professor can speak it too.
b. *The Skipper can speak Swedish and The Professor can speak in it too.

(12) a. ?That's the language that people who can speak _ usually own a boat.
b. *That's the language that people who can speak in _ usually own a boat.

(13) That's the language that Ginger said Gilligan can speak (in) _.

Example (11) demonstrates that the complement of *speak in* is inhospitable to pronouns, while the complement of *speak* is not. This contrast is reproduced in extraction across an island boundary, (12), but not in non-island extraction (13).

Stanton (2015) discusses several similar anti-pronominal contexts involving preposition-stranding. For instance, Stanton shows that certain locative prepositions forbid pronouns:

(14) a. Gilligan swam in the lagoon and Ginger swam in it too.
b. *Gilligan had lunch on the lagoon and Ginger had lunch on it too.

As we see in (14a), the preposition *in* is an instance of a locative pronoun which accepts pronouns. As we see in (14b), *on* is not. Note that the preposition *on the lagoon* is actually ambiguous. Under one reading, it locates a position on the surface of the lagoon. On another, it locates a position along the lagoon's edge. It is this latter reading which is unavailable in (14b). The former reading is pragmatically odd, though it is available if we imagine, for instance, that Gilligan and Ginger have some sort of raft which allows them to eat lunch on the lagoon's surface.

The examples in (15) illustrates how each environment behaves in cases of extraction from a subject island:

(15) a. ?That's the lagoon that swimming in _ causes strange rashes.
b. *That's the lagoon that having lunch on _ causes strange rashes.

Extraction across an island from the complement position of *on* (15b) is much less acceptable on the relevant reading than corresponding movement out of the complement position of *in* (15a).

Example (16) provides another example of this pattern, again using the preposition *on*. (16b) involves extraction from an adjunct island:

(16) a. *Mary works on the 8th floor of the Stata Center and Horatio works on it too.

b. *That's the floor of the Stata Center that Bill is worried because Mary works on _.

The PP *on the 8th floor* is ambiguous. One reading has an event taking place on the physical floor. For instance, Mary might be fixing a hole in it. On another reading, Mary's office is on the 8th floor. It is this second reading which is unavailable in (16a) and (16b).

A similar pattern shows up when we construct examples parallel to (16) using some of Stanton's other anti-pronominal contexts. Stanton discusses some pronominal restrictions involving temporal prepositions like *on Christmas eve*. First, we see that these prepositions generally forbid pronominal complements:

(17) *Mary ordered Chinese takeout on Christmas eve and Fred ordered takeout on it too.

As with its locative counterparts, a temporal preposition cannot be stranded in an island context. (18a) illustrates this with extraction from an adjunct island:

(18) a. ?*That's the holiday that Gilligan stopped visiting his family after learning that he could just order Chinese takeout on _.

b. ?That's the holiday that Gilligan stopped visiting his family after learning that he could celebrate _ at home.

Finally, extraction from an island blocks scope reconstruction:

(19) a. How many people do you think I should talk to?

b. How many people do you know a man who should talk to?

Example (19a) allows two readings. Crucially, with of one of these readings it is not necessary that the person answering the question have specific people in mind. For instance, imagine that we are conducting a survey. In this case (19a) can have as an answer a certain number of people, without any associated inference that the answerer has specific people in mind. This reading disappears in (19b).

1.2 Symbiotic gaps and selective islands

In this section we show that the full array of selective island effects discussed above reappear in Bachrach and Katzir's symbiotic gap (SG) constructions, but not in ordinary ATB-movement.

As we saw above, extraction from islands, to the extent that it is possible at all, is limited to DPs. This restriction does not hold on ordinary ATB-movement:

(20) How sick did Gilligan look _ and Ginger feel _?

However, it reappears with SG sentences:

(21) a. *How sick did Gilligan talk to people who looked _ and Mary Ann speak with people who felt _.

b. *How sick do people who look _ take the professor's medicine and people who feel _ get sick on the SS Minnow?

The DP restrictions show up in the symbiotic gap construction, just like in other cases of extraction from islands.

We also saw above that extraction from islands is limited to pronoun-accepting contexts. Below this is illustrated with another of Postal's (1998) examples of anti-pronominality:

(22) a. *Coppe painted his bike orange and Ted painted his car it.

b. *That's the shade of orange that people who paint their bikes _ usually come from the Netherlands.

c. Which shade of orange did Coppe paint his bike _?

d. Which shade of orange did Coppe paint his bike _ and Ted paint his car _?

An example of such a pronoun-rejecting context, the construction *paint his bike orange*, is provided in (22a). In (22b) we see that extraction from a subject island results in ungrammaticality, but not extraction from a non-island (22c). Finally, (22d) simply demonstrates that the gaps left behind by ordinary ATB-movement are hospitable to this particular anti-pronominal context.

However, unlike ordinary ATB-movement (22d), SG sentences forbid extraction here. This is illustrated by (23):

(23) *That's the shade of orange that Coppe knows a guy who painted his bike _ and Ted knows a guy who painted his car _.

The same effect is observed with Stanton's (2015) pronoun-rejecting contexts, such as locative prepositions:

(24) a. That's the river that John talked to some of the people who drank from _ and Mary talked to some of the people who swam in _.

b. *That's the river that John talked to some of the people who drank from _ and Mary talked to some of the people who had lunch on_.

c. That's the river that Bill drank from _ and Mary had lunch on _.

In (24a), we see a grammatical example of SG extraction from the complement position of *in*. Compare this to SG extraction from the complement position of *on* (24b). The sentence is ungrammatical (except on the pragmatically strange reading discussed in **section 1.1**). (24c) simply provides a case of ordinary ATB-extraction from complement *on*, for comparison.

This effect is also observed with Stanton's examples involving temporal prepositions. Again, extraction from the compliment of a temporal preposition is acceptable with ordinary ATB-movement:

(25) That's the holiday that my family celebrates _ and your's orders takeout on _.

But not in SG contexts:

(26) *That's the holiday that I have friends who celebrate _ and you have friends who order takeout on _.

Finally, we again observe scope reconstruction effects in symbiotic gap constructions:

(27) How many people do you know a man who thinks I should talk to _ and Mary know a woman who thinks I should avoid _?

Example (27) only allows a wide scope reading for the *many* phrase with respect to the modal embedded within the island.

2 A theory of movement

2.1 Movement and parallel merge

Johnson (2012, 2014) explores some of the consequences of a grammar with parallel MERGE. PM, Johnson shows, enables movement structures to be generated without some of the problems associated with Chomsky's (1995) internal MERGE. In particular, Johnson's system does not need to stipulate an operation like Fox's (1999) *trace conversion* as an independent LF rule. In **section 4** we shall see another argument for Johnson's theory: When combined with a theory of delayed Spellout (**section 3**), it provides a natural explanation for the selective island effects seen in **section 1**.

Internal MERGE (IM) combines one syntactic object with another object that it dominates. While this approach to movement is certainly elegant- a single syntactic operation is responsible for both movement and basic phrase structure building- IM also raises problems for semantic interpretation. If, as is typically assumed, the head and tail of a chain perform different kinds of semantic work, then how can a single syntactic object, with presumably a single denotation, perform two jobs at once?

Fox (1999, 2003) solves this problem with *trace conversion*, a post-derivational rule which converts the non-heads of a DP chain into a bound definite description at LF. While Fox assumes that trace conversion is a rule which applies at LF, Johnson shows how PM allows us to build the *effects* of trace conversion directly into the

derivational syntax.[1] Johnson's proposal simplifies the grammar, since trace conversion is no longer needed as an independent rule.[2]

We will set Johnson's system up in the following way. A derivation begins with a numeration, a set of the lexical items which will enter into the derivation, and a set of initially empty workspaces. The grammar possesses a function SELECT which moves lexical items from the numeration into one or more of the workspaces, and a function COLLECT which combines the contents of two workspaces.[3] Finally, MERGE may apply between syntactic objects within a workspace. Parallel merge is simply the application of external merge in two workspaces which happen to share the same constituent. Suppose we have a set of workspaces like the following:

(28) a. $W_1 = \{A, B\}$
b. $W_2 = \{A, C\}$

Here, the same syntactic object A is shared between the two workspaces W_1 and W_2. If MERGE(A,B) and MERGE(A,C) both apply, then A will be part of two syntactic objects $[_B$ A B] and $[_C$ A C]. If COLLECT then moves one of these syntactic objects from one workspace into the other, then further MERGE operations will be possible, resulting in a workspace in which A is shared.

[1]Johnson (2012, 2014) also raises the following conceptual objection to trace conversion. Fox's rule is designed to solve a problem which only exists for movement chains. However, there is no principled reason why one could not define a similar rule which effects the exact same structural change on any pair of c-commanding quantifiers. For instance, Johnson (2014: pg. 28-29) notes that one could imagine a grammar which applied a variant of the trace conversion rule to a sentence like (1):

(1) Every boy criticized no boy's mother.

In this hypothetical grammar, Fox's rule is allowed to apply to the lower quantificational phrase *no boy*, converting it into a definite description bound by the higher quantifier *every boy*. The sentence would then have the interpretation *Every boy x criticized the boy x's mother*. This sort of thing is clearly unattested. Instead, the mechanism must be specific to movement. While for Fox this restriction to chains derived by movement is written into the rule's structural description, Johnson argues that it would be better to eliminate TC as a distinct transformation entirely, and to capture its effects in the theory of movement itself.

[2]Another important conceptual advantage of Johnson's system is that it allows us to maintain an extremely strong version of Chomsky's (1995) *extension condition* (EC):

(1) EXTENSION CONDITION:
Both arguments of MERGE must be roots in some workspace

This version of the EC results in a highly restrictive grammar: while external MERGE remains possible, internal MERGE- which involves merger of a root syntactic object and a non-root which it dominates, is forbidden. However, parallel MERGE is still possible, if it is understood as simultaneous external MERGE.

[3]COLLECT, both the idea and name, is borrowed from Fox and Pesetsky (2007).

2.2 *Wh*-movement

With this alternative view of syntax in place, we can build the effects of Fox's trace conversion directly in the derivational syntax. We can illustrate this with Johnson's approach to *wh*-movement. For Johnson, the lower instance of the *wh*-chain is a bound defined description.

Following Fox (1999, 2003), we will assume that an function ID adjoins to the NP and introduces a variable n along with a phonetically null definite determiner *the*:[4]

(29) a. $[\![the]\!] = \lambda$f: $\exists!$x[f(x)=1]. ιx[f(x)=1]

b. $[\![\text{ID}]\!]^g = \lambda$x. x = g(7)

c. $[\![[_{DP}$ the $[_{NP}$ [ID 7] skipper]]$]\!]^g$ is defined iff $\exists!$x: skipper(x) = 1 & x = g(7),
Where defined, = ιx[skipper(x) = 1 & x = g(7)] = g(7)

Note that this analysis of the DP differs somewhat from Johnson's implementation, who follows Elbourne (2005) in assuming that an index combines directly with the D, resulting in a structure like [[D *i*] NP]. This difference in the structure of the DP will have consequences for our own analysis in **section 4** (see footnote 13).

Under this sort of analysis, the bound definite DP winds up denoting the entity introduced by g(7), and presupposes that g(7) is a skipper.The head of a chain is built by a Q morpheme (Beck, 2006; Kotek, 2014, a.o.), a function which converts the focus semantic value of its sister into the ordinary semantic value of its mother.

Below we go through a simple derivation of the sentence *Which coconut does Gilligan like?*. We begin with a numeration and a set of empty workspaces:

(30) N = {Q, *the*, ID-7, coconut, Gilligan, T, like}

a. $W_1 = \{\}$

b. $W_2 = \{\}$

The SELECT function then applies, placing the NP *coconut*, ID-7, and *the* in each workspace:

(31) N = {Q, Gilligan, T, like}

a. W_1 = *the*, ID-7, coconut

b. W_2 = *the*, ID-7, coconut

Successive applications of MERGE then form a bound DP in each workspace:

(32) N = {Q, Gilligan, T, like}

a. $W_1 = [_{DP}$ *the* $[_{NP}$ ID-7 coconut]]

[4]The notation below is borrowed from Heim (2013).

b. W_2 = [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]

Now, SELECT applies again, placing *like*, T, and *Gilligan* in W_1:

(33) N = {Q}

a. W_1 = Gilligan, T, like, [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]

b. W_2 = [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]

MERGE again applies, creating a TP in W_1:

(34) N = {Q}

a. W_1 = [$_{TP}$ Gilligan [T [$_{VP}$ like, [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]]]]

b. W_2 = [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]

Next, SELECT places Q in W_2:

(35) N = {}

a. W_1 = Gilligan, T, like, [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]

b. W_2 = Q, [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]

This is followed by MERGE(Q, DP):

(36) N = {}

a. W_1 = [$_{TP}$ Gilligan [T [$_{VP}$ like, [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]]]]

b. W_2 = [$_{QP}$ Q [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]]

Finally, COLLECT applies, combining the QP from W_2 with the material in W_1. This is followed by merger of QP and the structure in W_1:

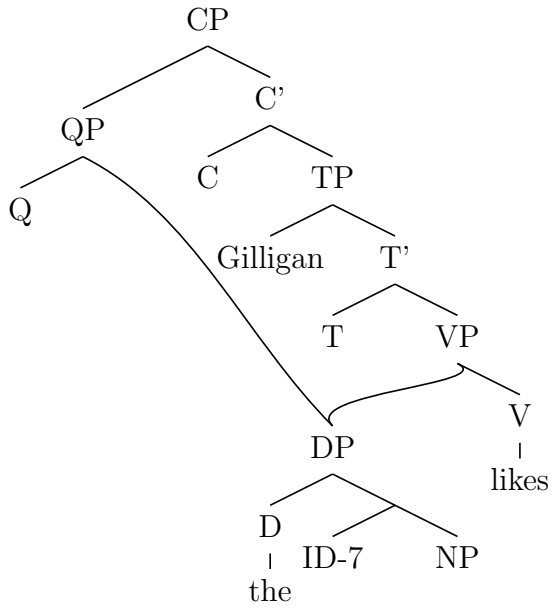
(37) N = {}

a. W_1 = [$_{CP}$ [$_{QP}$ Q [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]][$_{TP}$ Gilligan [T [$_{VP}$ like, [$_{DP}$ *the* [$_{NP}$ ID-7 coconut]]]]]]

b. W_2 = {}

The resulting structure can be graphically represented as follows:

(38)



The linearization of structures like (38) is determined by a set of interacting constraints. For instance, Johnson assumes a constraint CONTIGUITY, which has the effect of generally mapping phrases into contiguous strings. Competing is a constraint which requires the material dominated by QPs to be linearized to the left.[5]

The semantics of the QP will not be relevant to this paper. Johnson (2014) shows that his system is compatible both with an analysis based upon Beck (2006)- in which the wh-DP introduces focus alternatives, which the Q morpheme converts into the ordinary semantic value-, while Johnson (2012) an analysis in which Q denotes an unrestricted existential quantifier (Heim, 2013).

2.3 Quantifier raising

Quantifier raising will work in a similar fashion, with the bound definite DP merged in the QR structure's base-generated position. A quantificational DP forms via PM,

[5] From Johnson (2014, pg. 16):

(1) Contiguity

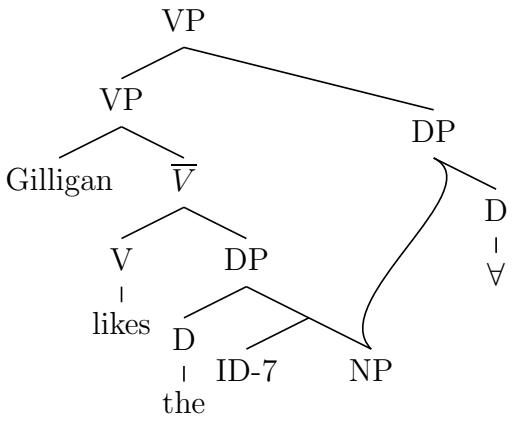
Let A be the set of vocabulary items dominated by some node A, and b a vocabulary item not in A. For every A in a phrase marker, if a linearization includes $b < a$ then it cannot include $A' < b$, for $a, b \in A$, and if it includes $a < b$, then it cannot also include $b < A'$, for $a, A' \in A$.

(2) The Wh-criterion

If a wh-QP is merged to a CP, then a grammatical linearization must position that QP so that it precedes everything in the QP.

and then merges into an interpretable position which c-commands the definite DP:[6]

(39)



The structure in (39) will receive the following interpretation:[7]

[6]Kyle Johnson (p.c.) points out the following problem with my implementation of the system. Johnson's syntax, whatever the analysis of the DP, permits a sort of anti-late merger in which a modifier is adjoined to the the lower position of the NP, but not the higher position. Thus for a sentence like *John likes every book about syntax* the following problematic LF is possible:

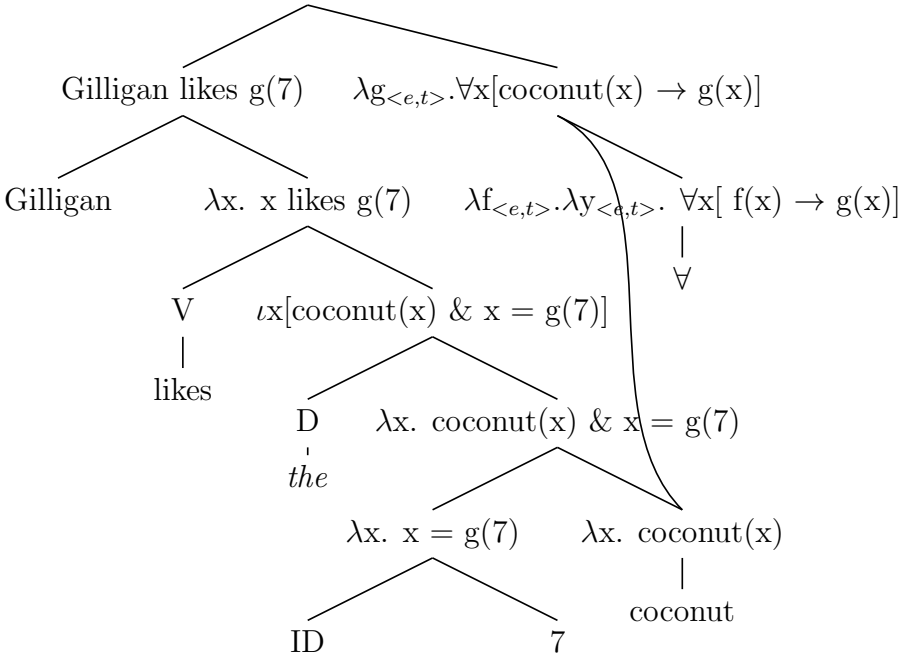
(1) [[∀ book] λx [John likes [[THE [ID-x book about syntax]]]

The issue is that whatever constraint forces the constituent *book about syntax* to be in both the restrictor and scope positions might also rule out an LF like (39) by forcing ID-n to be shared along with the NP.

[7]Johnson's theory requires that semantic binding is established via a rule like the following:

(1) **Variable binding rule**
Let α be a branching node with daughters β and γ, where β = [D δ] and γ dominates some phrase [$_{DP}$ the ID i δ]
Then, $[\![\alpha]\!]$ = FA($[\![\beta]\!]$, [λx. $[\![\gamma]\!]^{g[i \rightarrow x]}$])

(40) $\forall$x[coconut(x) $\rightarrow$ Gilligan likes ιy[coconut(y) & y = x]]

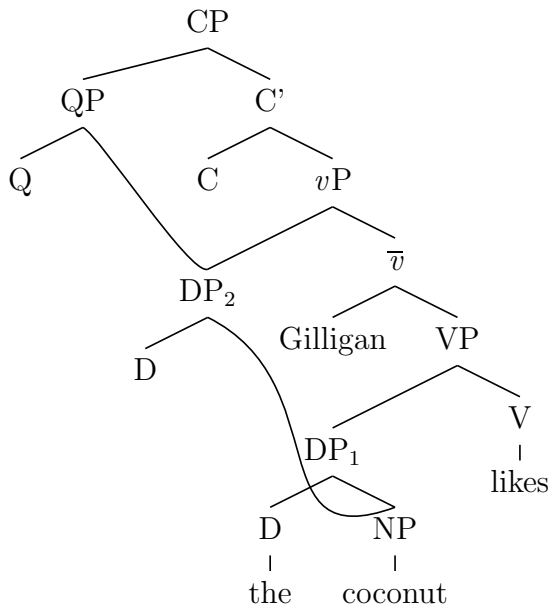


2.4 Successive cyclicity

Successive cyclic movement, finally, is a case of QR feeding *wh*-movement.[8] Consider *Which coconut does Gilligan like* again. First, a D merges with NP, and then with the complement of the VP. Next, a second D merges with the NP, before the resulting DP remerges in a c-commanding position, at the left edge of the VP. This repeats the derivation for a QR structure, except with the quantificational morpheme replaced with a second definite determiner. A Q morpheme will be merged with the highest DP:

[8]Johnson (2014) argues that this can explain why "Beck effects" (Beck, 2006) do not appear in overt movement. Beck observes that in some languages in-situ *wh*-phrases are sensitive to the presence of a focus-sensitive operator located between the position where the *wh*-phrase takes scope and the position where it is pronounced. This follows from Beck's analysis of *wh*-questions, since the Q morpheme is itself a focus sensitive operator. Such intervention effects fail to appear, however, in overt *wh*-movement. This follows if successive cyclic movement can travel over an intervening focus sensitive operator, and the binding relation established by this movement step does not depend on focus.

(41)



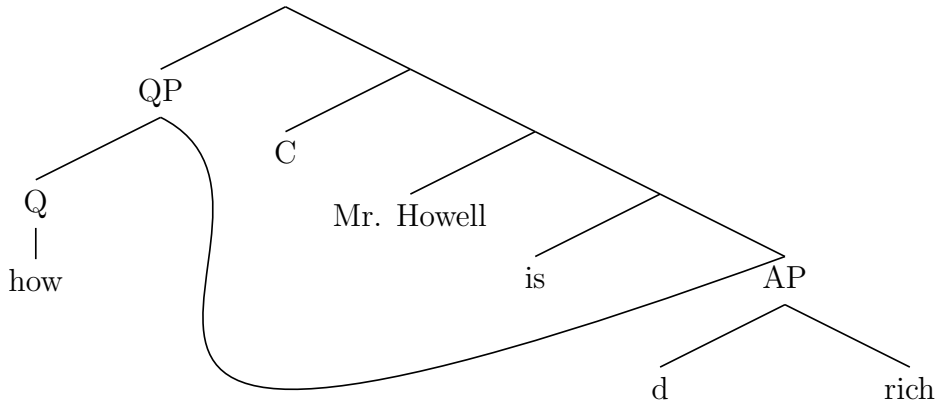
2.5 non-DP movement

Finally, consider the movement of a non-DP, as in the sentence in (42):

(42) How rich is Mr. Howell?

Johnson suggests a syntactic structure like the following for (42):

(43)



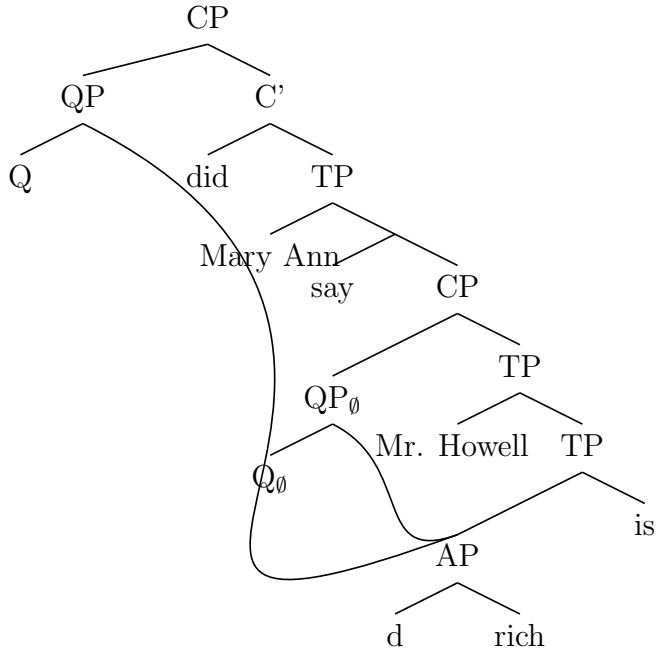
In this structure, Q merges with the AP, forcing the entire AP to pied-pipe the edge of the sentence at PF. What about successive cyclic movement? For DP movement, as we saw above, a moved NP was semantically interpretable in its derived position, thanks to the introduction of a second determiner that parallel merged with the NP. This second DP then merged into a position which syntactically c-commanded, and semantically bound, the other DP that the NP was a complement of. This sort of derivation is unavailable for movement of predicative phrases, since the syntactic material which made the lower DP interpretable as a variable, a phonetically null D and

an variable function ID-n, would not be able to produce a semantically interpretable structure.[9]

However, Johnson's system does provide a certain flexibility in which nodes are semantically interpreted. While it is required that a phrase be interpreted at some position, it does not need to be interpreted at *every* position it occupies. This allows for a derivation in which the moved non-DP merges into a higher position, but is interpreted only in its lower position. We could execute this by assuming that each step of SC movement for a non-DP involves merger of a semantically vacuous Q morpheme to the DP. This vacuous QP then merges at the edge of a Spell-out domain.

This is illustrated below for the sentence *How rich did Mary Ann say Mr. Howell is _?*. For simplicity, we assume that Spellout only applies at CP:

(44)



Besides the semantically contentful QP which merges at the edge of the matrix CP, this derivation also involves a vacuous $Q_{\emptyset}$, which also merges with the AP, and then merges at the edge of the embedded CP. When Spellout applies at the embedded CP, the AP is shared between the QP and its base-generated position. While the rules of linearization allow for everything dominated by the QP to be pronounced at the left edge of the SOD, the AP is interpreted only in its base-generated position.

Johnson points out that this predicts a basic asymmetry between DP and non-DP movement: while DP which moves successive-cyclically may be interpreted in its derived position, a non-DP may not be. In other words, reconstruction for non-DPs is mandatory. This predicts, for instance, that unlike DPs (Lebeaux, 1988), late

[9]If the definite determiner merged with an AP, it would produce a definite description, not a phrase with a predicative interpretation.

adjunction should be unable to obviate condition C effects with non-DPs. It has been claimed by various authors (Barss, 1986; Huang, 1993) that this is in fact so:

(45) a. *How satisfied by Sarah$_i$'s debate was she$_i$ _?
b. *How satisfied because of Sarah$_i$'s performance did she$_i$ say you were _?

However, the judgments of some informants that I have consulted varies for these sentences.

3 A theory of cyclic Spellout

3.1 Cyclic Spellout

Crucial to this proposal will be the assumption that syntactic islands are determined by the mapping between syntax and the interfaces (Fox and Pesetsky 2005, a.o.). This can be captured in a theory like Fox and Pesetsky's theory of *cyclic linearization*. According to Fox and Pesetsky, Spellout applies cyclically at certain designated Spellout domains (SOD). When Spellout applies to a given SOD, a set of linear precedence (LP) relations between nodes in that SOD is determined. Each subsequent round of Spellout will simply add to this set of LP relations. Fox and Pesetsky furthermore assume that ordering contradictions are forbidden. That is, for any two distinct nodes α and β, the set of LP statements may not contain both $\alpha < \beta$ and $\beta < \alpha$.

It follows that, if a phrase α is to be moved to the left edge of a sentence (as *wh*-phrases typically are in languages like English), then α must be linearized at the left edge of every SOD that it is Spelled out in. An island, according to Fox and Pesetsky, is simply an SOD which prevents this:

(46) **Island for leftward movement of α**
An island for leftward movement is a Spellout domain that does not allow α to be pronounced at its left edge
(Fox and Pesetsky 2007, pg. 2)

We will assume Fox and Pesetsky's theory, with one important addition: the syntax-to-semantics mapping is determined cyclically, along with the syntax-to-phonology mapping. While this is assumed in conceptions of cyclic Spellout like phase theory (Chomsky, 2000), it is not a direct concern of Fox and Pesetsky, or any works that build upon their theory.

We will assume that cyclic Spellout must satisfy the following principles:

(47) For some phasal node X and the Spellout domain SOD(X):
Spellout (PF):
a. For every $x, y \in$ SOD(X), the linear relation between x and y is determined at Spellout(X)

b. The total set of linear precedence statements for a given derivation must not produce an ordering contradiction at Spellout(X)

c. Once an LP relation is established, it cannot be altered by later iterations of Spellout.

(48) **Spellout (LF):**

a. For every $x \in$ SOD(X), $[\![x]\!]$ is determined at Spellout(X)

b. $[\![$X$]\!]$ must be semantically well-formed at Spellout(X)

c. Once $[\![$X$]\!]$ has been determined, it cannot be altered by later iterations of Spellout.

While (47) simply encodes the principles of cyclic linearization which we discussed above, (48) adds the assumption that the semantic interpretation of any node is fixed cyclically, and that once the interpretation is fixed, it cannot be altered by subsequent iterations of Spellout.

But what exactly gets Spelled out when Spellout applies? Following Chomsky (2000), phase theorists generally assume that Spellout is calculated with respect to the c-command domain of designated *phase heads*, such as v or C. Bachrach and Katzir (2009) suggest (i) that Spellout is calculated with respect to dominance rather than c-command and (ii) that Spellout is sensitive to what they call *complete dominance*. A version of this notion, which is slightly different from Bachrach and Katzir's, is made more precise below:[10]

(49) a. **Complete Dominance**: a node X completely dominates a node Y within some workspace W iff X reflexively dominates every mother of Y in W. The set of nodes completely dominated by X in W is called the **Complete Dominance Domain** of X in W, abbreviated CDD(X, W).

b. **Spellout Domain (SOD)**: SOD(X, W) = CDD(X, W), if X is a **Spellout node**.

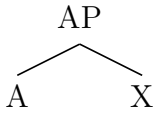
c. **Spellout Node**: A designated syntactic object that triggers spellout of its spellout domain (= VP, CP, roots of specifiers and adjuncts).

Importantly, complete dominance is defined relative to a given workspace. Consequently, if a phrase is shared between two positions in different workspaces, then that phrase will count as completely dominated within each workspace.

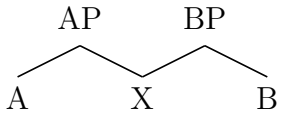
To illustrate the notion of complete dominance, consider the two syntactic structures below:

[10]This notion of complete dominance dates back to Wilder (1999), who calls it *full dominance*. Fox and Pesetsky (2009), whose analysis will become important latter on, use the term *total dominance* to distinguish their definition, virtually identical to my own, from Bachrach and Katzir's definition, which is slightly different from the definition that I'm presenting. I'm sticking with this label, though, since it is used with the definition in (49a) in other works (e.g. Gracanin-Yuksek 2007).

(50) a.



b.



In (50a), A **completely dominates** X, since AP (reflexively) dominates every mother of X (which happens to be just X itself). (50b) illustrates a sharing structure, generated by application of parallel merge. In this structure, X has two mothers- AP and BP. Supposing that AP and BP co-exist in some workspace W, X $\notin$ CDD(AP, W) or CDD(BP, W).

Now, **Spellout**(K, W), where K is the root of an SOD within a workspace W, will be defined as follows:

(51) For some **Spellout node** K and workspace W, where K $\in$ W,
At **Spellout**(K,W)

a. Form *Lin*(K), the set of linear precedence statements for all nodes X $\in$ SOD(K,W)

b. Semantically interpret every $x \in$ SOD(K,W)

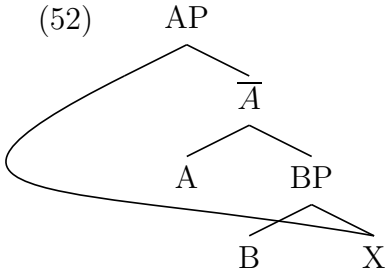
c. For all **Spellout nodes** in SOD(K,W), reuse previously computed PF and LF interpretations.

A more complete version of the Spellout algorithm is provided in the **Appendix**.

3.2 How it works

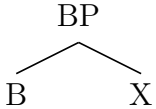
Now consider the Spellout of a schematic structure like (52):

(52)



Assume that B is an island for movement of X. The initial round of Spellout will apply to the following structure:

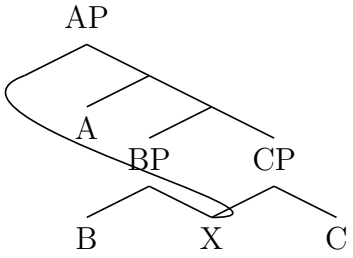
(53)



If we suppose that X is forbidden from merging at the left edge of BP, then X will have to be linearized to the right of B. If the language is head initial, then this will result in a B<X LP statement. The next round of Spellout, which will apply to the entire structure in (52), will force X to be merged at the left edge of the SOD. Crucially a statement X < B will be generated, contradicting the previous round of Spellout.

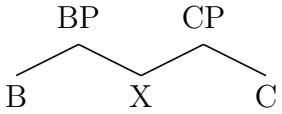
Now consider an alternative structure, which involves syntactic sharing:

(54)



In this structure, X is merged as the complement to both B and C. It has also remerged as the specifier of AP. Assume that AP, BP, and CP are SODs. Also assume BP and CP are islands. Imagine also that the **Spellout**(BP) and **Spellout**(CP) applies when the following structure has been formed within a single workspace:

(55)

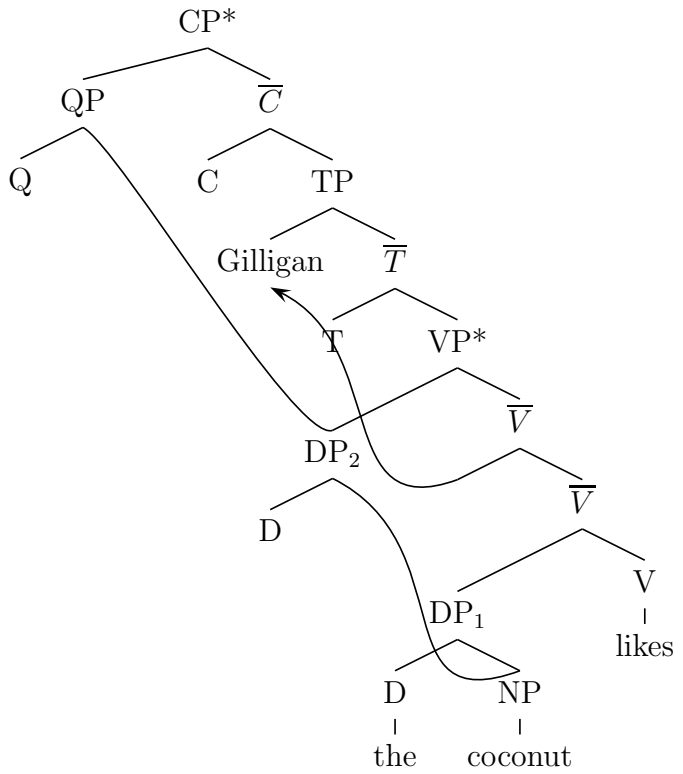


Now X will avoid Spellout inside of the BP and CP SODs, since it is incompletely dominated by both. This has two consequences. First, a linearization contradiction fails to arise, since X will not be linearized until A is merged and the AP Spellout domain is formed. Second, X will not be semantically interpreted within its lower positions as complement,AP and complement,BP, since it avoided Spellout in these two positions.

3.3 A sample derivation: DP movement

Again consider the derivation for *Which coconut does Gilligan like*, now taking into account the theory of cyclic Spellout. The structure for the entire sentence will be as follows, with * marking the Spellout nodes:

(56)



Spellout will initially apply to the maximal VP projection. In order for the linearization algorithm to work properly, two assumptions about the linearization algorithm will be necessary:

1. a linearization constraint, similar to Johnson's *wh*-criterion (see footnote 5), which forces the material dominated by DP_2 to be pronounced at the left edge of the SOD

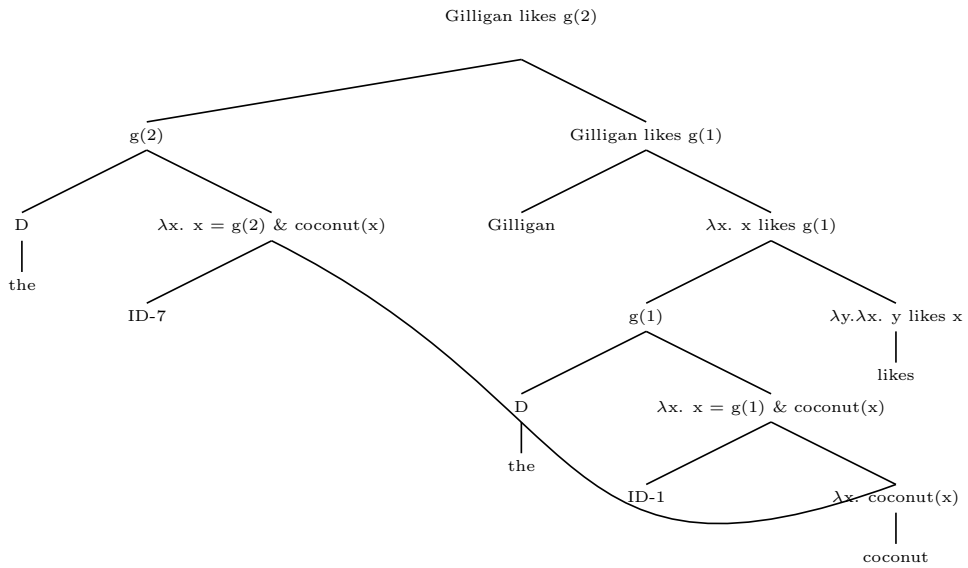
2. phonetically null determiners are not considered by linearization.

With these assumptions in place, we will arrive the following set of linearization statements:

(57) $Lin(\text{VP*}) = \{\text{coconut} < \text{Gilligan}, \text{coconut} < \text{likes}, \text{Gilligan} < \text{likes}\}$

This SOD will have the following semantic interpretation:

(58)



The next Spellout domain will be the matrix CP. It should produce the following linearization statements:

(59) $Lin(\text{CP}^*) = \{\text{which} < \text{coconut}, \text{QP} < \text{Gilligan}, \text{QP} < \text{VP}, \text{coconut} < \text{Gilligan}, \text{coconut} < \text{likes}, \text{Gilligan} < \text{likes}\}$

This set of LP statements is non-contradictory, an so causes no problems at PF. The semantic interpretation will proceed as follows. First the semantic value computed for VP* in the earlier round of Spellout will be retrieved, then this value will be used to compute the semantic value for the entire CP.

4 How to get off an island

4.1 Islands

Now consider extraction from an island structure. For instance, take a sentence like (60):

(60) ??Which skipper does Mary Ann know a professor who admires _?

In (60), the *wh*-phrase *which man* has been extracted from a relative clause, in violation of the complex NP constraint (Ross, 1967). According to the theory of islandhood adopted in this paper, the unacceptability is ultimately due to cyclic Spellout. Extraction out of the complex NP yields a linearization contradiction. However, the goal of this section will be to show that, in our Johnsonian framework, derivational possibilities exist which should allow delayed Spellout of the NP, and therefore avoid a linearization contradiction.

To see how this will work, let us zero in on the crucial step in the derivation, at which the island SOD is formed. We will assume that the SOD which determines islandhood in (60) is the relative clause itself.[11] Following the ceoncpetion of islandhood borrwed from Fox and Pesetsky, we will assume that linearization of the *wh*-phrase at the left edge of the relative clause is blocked.[12]

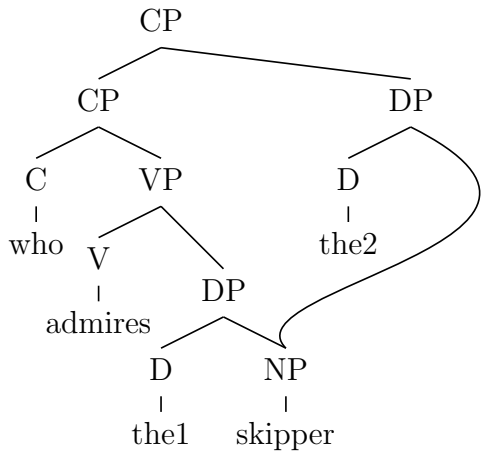
Each DP involved in the intermediate movement step initially occupies a separate workspace. Thus we might have a derivational stage like the following:

(61) $W_1 = [_{CP}$ who [admires $[_{DP}$ *the*1 skipper]]]
$W_2 = [_{DP}$ *the*2 skipper]
etc.

At this point, a number of possibilities are available. First, the island CP in W_1 could be Spelled out without first COLLECTing W_1 and W_2. However, because the NP *skipper* would be completely dominated within W_1, it would be linearized to the right of the V *admires* , which would ultimately lead to a linearization contradiction.

We could also COLLECT W_2 with W_1 and then MERGE the second DP to the CP. W_1 would then contain the following structure.

(62)



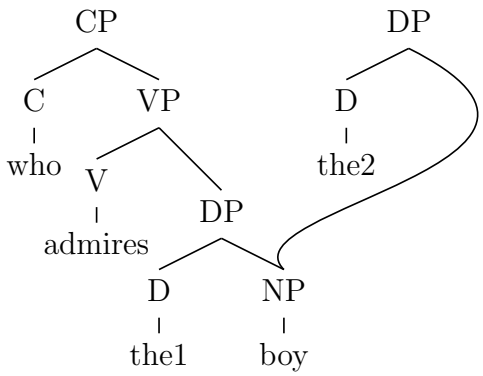
However, this derivational sequence is unavailable since, as an island, the DP may not be merged at the CP's left edge.

Crucially however, yet another possibility exists. Suppose that COLLECT applies between W_2 and W_1, but that the second DP does not Merge with the island CP. At this point the W_1 would look like the following:

[11] The question of exactly what SOD is responsible for RC islands depends upon what is responsible for the RC's islandhood. Since RCs are arguably adjuncts, and since their creation involves RC-internal $\overline{A}$-movement, it seems plausible to assume that the RC itself constitutes either an adjunct or a subjacency island.

[12] This might be because movement of the *wh*-phrase is simply forbidden to the left edge of the island SOD, or because the *wh*-phrase is required to 'tuck in' at an inner specifier,CP position underneath the relativization operator. In a system like Fox and Pesetsky's, inner specifiers in languages like English are linearized in their base-generated position. See Fox and Pesetsky (2009).

(63)

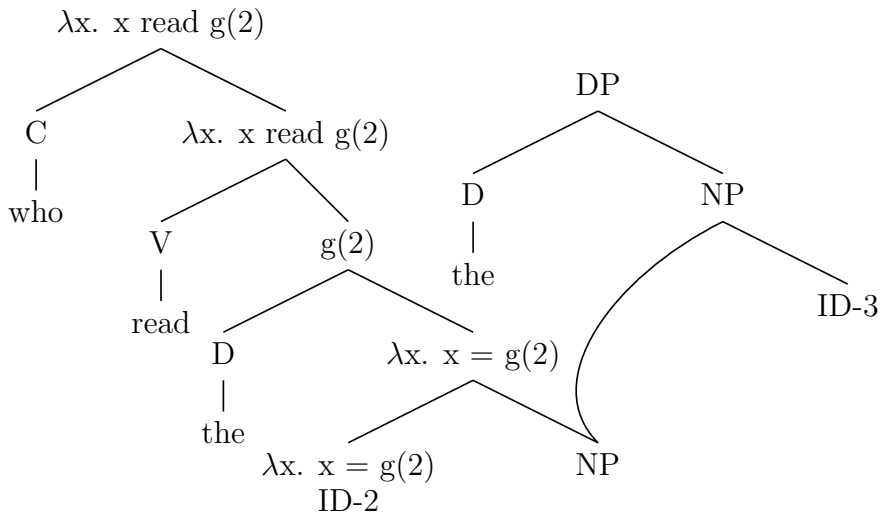


In this structure, the CP root of the relative clause completely dominates the determiner *the1*. The NP, which is also merged into a second DP, is not completely dominated by the CP. Now suppose that Spellout applies to the CP at this point.

According to the definition of SODs borrowed from Bachrach and Katzir, in which only a node which is completely dominated by the Spellout node is Spelled out, the NP should avoid Spellout. At PF Spellout, the NP will avoid linearization with respect to the other non-null terms in the SOD. Consequently, a linearization contradiction will not occur later on in the derivation. At LF Spellout, the structure which is completely dominated within the island SOD then receives the following semantic interpretation:[13]

[13]Note that it this analysis is not easily made compatible with the specific analysis of the bound definite DP assumed by Johnson (2014), which has an Elbourne (2005) inspired structure [[THE n] NP]. If the NP is not spelled out in this sort of structure, then the definite determiner will simply lack a type $< e, t >$ complement to combine with, resulting in semantic uninterpretability. However, if we assume that referential pronouns combine with a phonetically null ONE predicate, as Elbourne (2005: pg. 95) suggests, then the grammar could generate a structure comparable to (64) in which the lower DP is replaced with [[THE n] ONE NP]. If the NP, but not ONE, is shared, then the Spelled out structure [[THE n] ONE] will be semantically interpretable. However, this variant of the analysis seems less attractive than the current proposal.

(64)



With this we arrive at the major theoretical conclusion of this paper. Delayed Spellout of the NP in (64) has two consequences. First, an island violation is avoided, since Spellout of the NP within the island is delayed. Second, delayed Spellout will result in an interpretable structure. Although the NP is not interpreted semantically, the island CP still receives a well-formed semantic interpretation, with the base-generated position of the movement structure interpreted as an unrestricted variable.[14] Since this semantic interpretation constitutes the LF output for cyclic Spellout of the island, this will be the only interpretation available for the final structure. Future rounds of Spellout will be incapable of changing it. In other words, the structure in (64) illustrates the following:

(65) A phrase which is moved out of an island may not reconstruct into the island

As we are about to see, this result will allow us to derive the full array of selective island effects discussed in **section 1**.

4.2 Anti-pronominality

Among other things, the conclusion in (65) derives anti-pronominality. The anti-pronominality effects observed in **section 1** show that the material left behind in extraction from island contexts differs from that left behind in non-island extraction:

[14]Recall that the determiner denotes a function which applies to a type $< e, t >$ predicate and returns the unique entity denoted by that predicate. ID-n simply denotes the set of individuals identical to g(n). In other words, it is a set of consisting of a single individual, the one corresponding to g(n). [*the* ID-n] will simply return that individual. We assume that this is identical to the semantics of an ordinary pronoun, except that pronouns typically also carry ϕ-features.

the syntactic object left behind in island contexts appears to have something in common with a pronominal, whereas the object left behind non-island movement behaves like a fully reconstructed DP. This is illustrated by the example in (1), repeated below as (66):

(66) a. Gilligan speaks Hungarian, and Mary Ann speaks (*in) it, too.
b. That's the language that Gilligan knows someone who speaks ?(*in) _.

A similar contrast shows up elsewhere. Postal (1998) and Stanton (2015) notice that different types of $\overline{A}$-movement differ in their sensitivity to pronoun rejecting contexts. Some types of extraction, like topicalization, disallow movement out of pronoun rejecting environments:

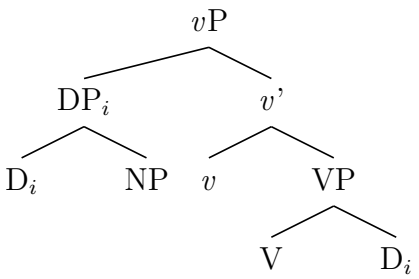
(67) Hungarian, Mary refuses to speak (*in) _.

Others, like *wh*-movement and relativization, allow this sort of extraction (provided the movement is not out of an island, of course):

(68) That's one language that Mary refuses to speak (in) _.

Stanton (2015) provides an explanation for this contrast. Assuming copy theory, she argues that topicalization (and other types of movement with similar properties) forces *wholesale late merger* (Takahashi and Hulsey, 2009) of the NP complement of the moved DP at a site structurally higher than its base-generated position. As a consequence, the copy left behind in the base-generated position in this sort of configuration will be a bare D, with the full DP copy only appearing at a higher site:

(69)



For reasons not relevant to the present discussion, Stanton argues that in *wh*-movement and relativization merger of the NP necessarily occurs in the base-generated position of the chain. The copy in the base-generated position will be a full DP copy, NP complement included.

Following Postal (1966), Stanton argues that a bare determiner is structurally identical to a pronoun. Hence the lower copy in a case of topicalization might be expected to display sensitivity to anti-pronominal contexts.[15] Conversely, the lower

[15]While Stanton does not analyze the full range of Postal's anti-pronominality effects, she argues that, for pronoun-rejecting temporal and locative PPs, specific material in the NP must be present in order for the PP to receive the appropriate interpretation. Note that Stanton's examples of

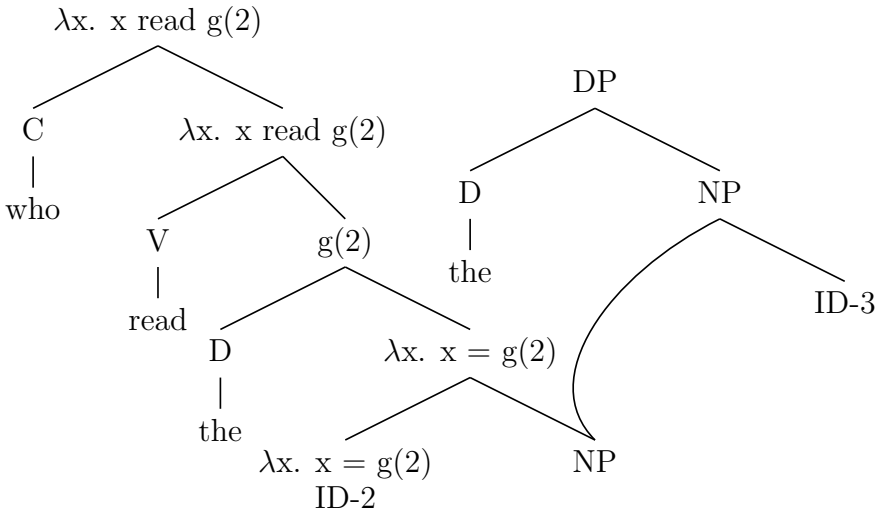
copy in *wh*-movement will be a fully reconstructed DP copy. Anti-pronominal sensitivity is not expected.

To put it in a slightly different way, Stanton argues that the grammar imposes certain restrictions on the reconstruction of the NP complement in a DP movement structure. This restriction in turn predicts anti-pronominality. We will adopt this basic insight of Stanton's paper:

(70) Anti-pronominal effects are related to the ability of the NP to reconstruct into the pronoun-rejecting site.

With that in mind, let us return to the structure in (64), repeated below:

(71)



As we have seen, a cyclic Spellout algorithm which incorporates complete dominance into its definition of SODs will predict that the NP in this structure can avoid Spellout, and therefore avoid becoming trapped in an island. Moreover, the NP's avoiding Spellout will also prevent its reconstruction into the island. However, the phrase [$_{DP}$ THE ID-2]will be Spelled out, and interpreted as a variable, in other words a D with no NP complement. This will have an effect identical to Stanton's use of WLM, except that the restriction on NP reconstruction now arises out of the interaction between a Johnsonian syntax and the Spellout algorithm, rather than the dynamics of MERGE itself.

pronoun-rejecting prepositions, like *on the lagoon*, are generally ambiguous between a literal reading (where one is actually located on the lagoon), and an idiomatic one (where one is along the banks of the lagoon. It is this idiomatic reading which is rendered unavailable in cases like extraction from an island.

4.3 DP/non-DP asymmetry

We saw in **section 2.5** that Johnson's theory posits syntactic and semantic differences between DP and non-DP movement. While successive cyclic movement of a DP may affect semantic interpretation, non-DP movement is semantically vacuous.

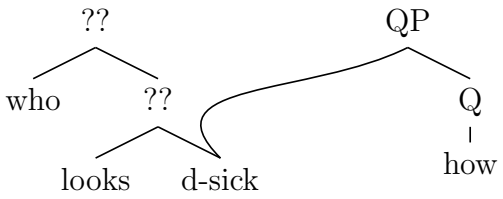
We have also seen that DPs are allowed to escape from islands, once Johnson's syntax is combined with delayed Spellout theory. But what about non-DPs?. Consider a sentence like the following:

(72) *How sick does Mary know a man who looks _?

Of course, we could imagine a derivation in which the *wh*-phrase is Spelled Out within the island SOD; however, as an island, the *wh*-phrase will not be Spelled out at the left edge of the SOD. As with DP movement, this should lead to a linearization contradiction in later Spellout domains.

But what if we attempt to delay Spellout of the *wh*-phrase? Again, this would be accomplished by creating a workspace which the wh-AP is incompletely dominated within the island Spellout domain? This sort of structure would look like (73):

(73)



While this structure will avoid a linearization contradiction it will also lack a well-formed semantic interpretation. In order for the delayed Spellout mechanism to produce a structure with well-formed semantic interpretation, there must be something Spelled out in the moved phrase's base-generated position which is semantically interpretable in that position. This requirement is satisfied in the case of DP movement, in which a DP [THE ID-n], which has the interpretation of a variable, will occupy that position. This is not possible given Johnson's syntax for non-DP movement, since the AP itself is the only thing merged in the base-generated position which can receive a semantic interpretation. In short, reconstruction into the base-generated position is obligatory for non-DPs, rendering the delayed Spellout mechanism inapplicable to avoid island violations.

4.4 Scope reconstruction

It's been observed that islands block scope reconstruction (Longobardi 1987, Cinque 1990, Ruys 2011). While this observation has typically been made regarding *wh*-islands (75a), it can also be extended to syntactic islands, such as the complex NP and adjunct island sentences in (75):

(74) How many people do you think I should talk to?

(75) a. How many people do you wonder whether I should talk to?
b. How many people do you know a man who should talk to?
c. How many people did you get mad because Bill should talk to?

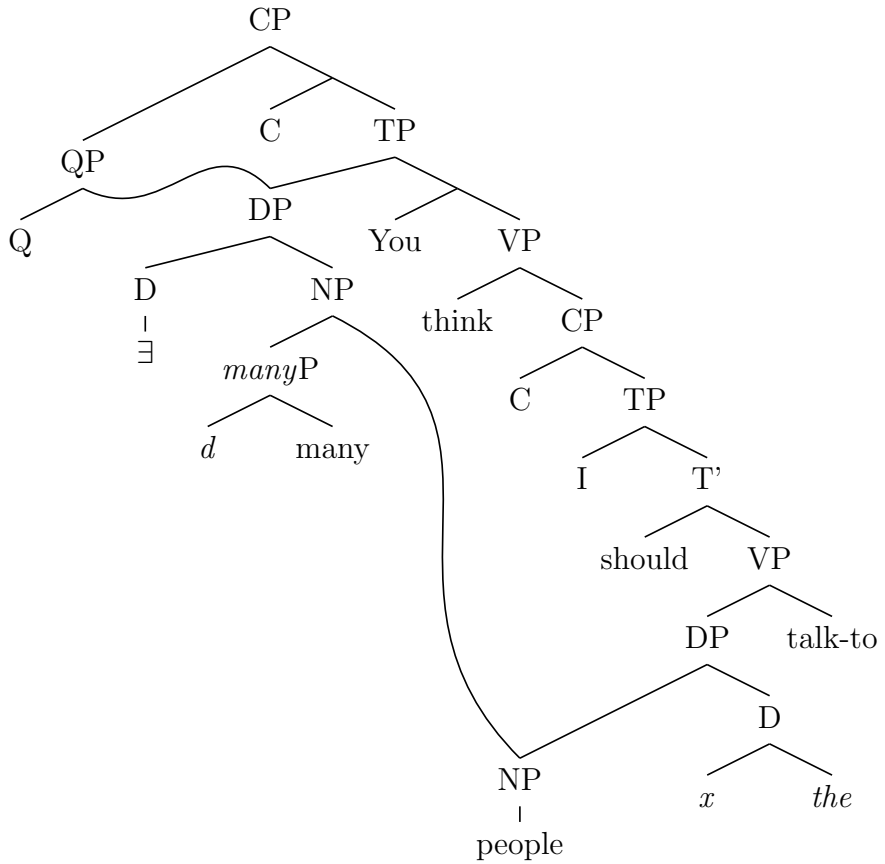
(74) is ambiguous. Under one reading, it means something like: *for what number n are there n people x such that you think I should talk to x?*. Under this reading, the *many* phrase takes scope over the modal *should*. To see what sort of LF generates this reading, we will assume that this sort of *wh*-phrase involves two sorts of quantification. First, a *wh*-operator which quantifies over degrees; second, existential quantification. Assume a semantics for the *many*-phrase similar to the analysis of Hackl (2000):

(76) $[\![\textbf{many}]\!] = \lambda d.\lambda x.\ \mathit{measure}(\mathrm{x}) = \mathrm{d}$

After *many* combines with a degree *d*, it will denote a property of individuals which can combine with an NP via predicate modification. This NP will then combine with an existential quantifier.

There are two positions in which the DP containing *many* can take scope. First, it may take scope above the modal:

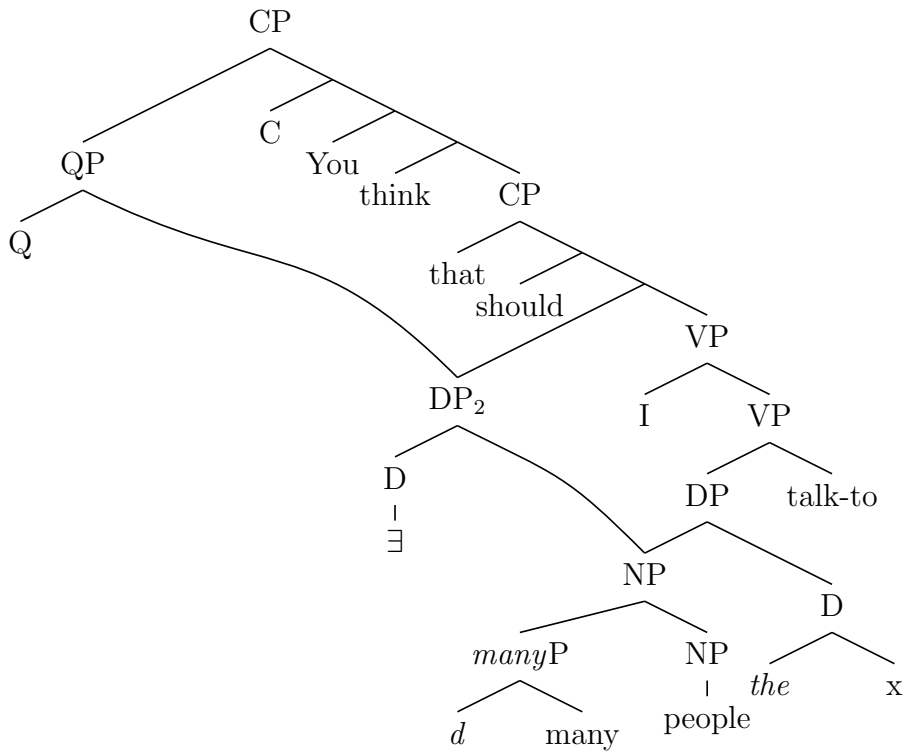
(77)



The interpretation of this structure can be paraphrased as *what is the maximal degree n such that there are n-many people that you think I should talk to?*. For instance, suppose that you ask me for advice about who to talk to about a particular problem in phonology, and I response that you should speak with Bill and Tom. In this case, the answer to the question corresponding to the LF in (77) would be: two people (Bill and Tom).

A second LF is available, in which the *many* phrase scopes below the modal:

(78)

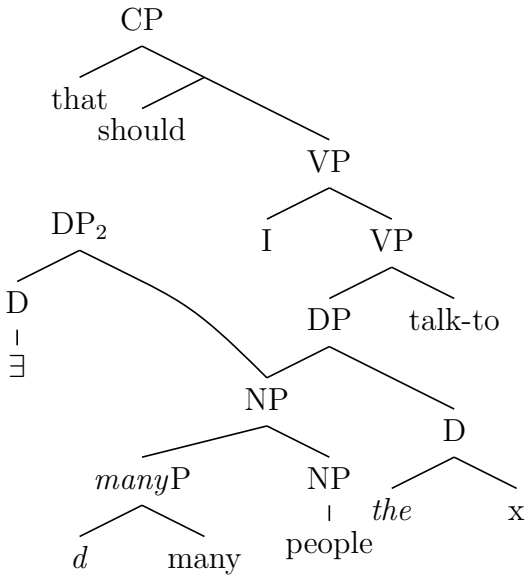


The second LF corresponds to an interpretation with the following paraphrase: *what is the number n such that you think that there should be n-many people that I talk to?*. This could correspond to a situation where I want you to talk to some number of people, but do not have any particular individual in mind.

Now let us consider these sorts of structures in island sentence those in (75). Curiously the ambiguity has disappeared in these sentences. Instead, only a wide scope reading paraphrased is available. For instance, for (75a) this can be paraphrased as *what is the number n such that there are n-many people that you wonder whether I should talk to?* is available.

This follows from our assumptions about how islandhood works. The point is that, even if the *many* phrase is merged in a position within the island, the mechanics of delayed Spellout will prevent its being interpreted inside of the island. At a relative clause island SOD, for instance, our syntax would posit a structure like (79):

(79)



In this structure, the DP will be interpreted within the island as a bare D, and hence as a variable. Interpretation of the existential quantifier and *many*-phrase will be delayed until some point above the island, and hence above the modal. We correctly predict, then, that only the wide scope reading will be available.

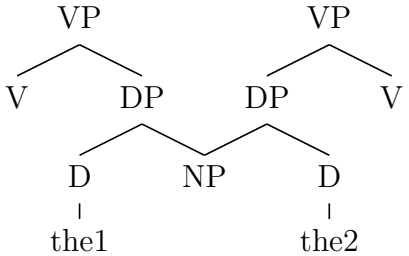
4.5 Selective islands and across-the-board movement

Next let us see how selective island effects arise in ATB-movement. Imagine a derivation like the following, in which a single NP is shared between two workspaces. Different determines then merge with the NP in each workspace:

(80) a. $W_1 = [_{DP}$ *the*1 NP]
$W_2 = [_{DP}$ *the*2 NP]

Imagine that COLLECT applies at this point, so that the two DPs now occupy the same workspace. Suppose furthermore that two VP SODs are built up, so that the structure in the workspace looks like (81):

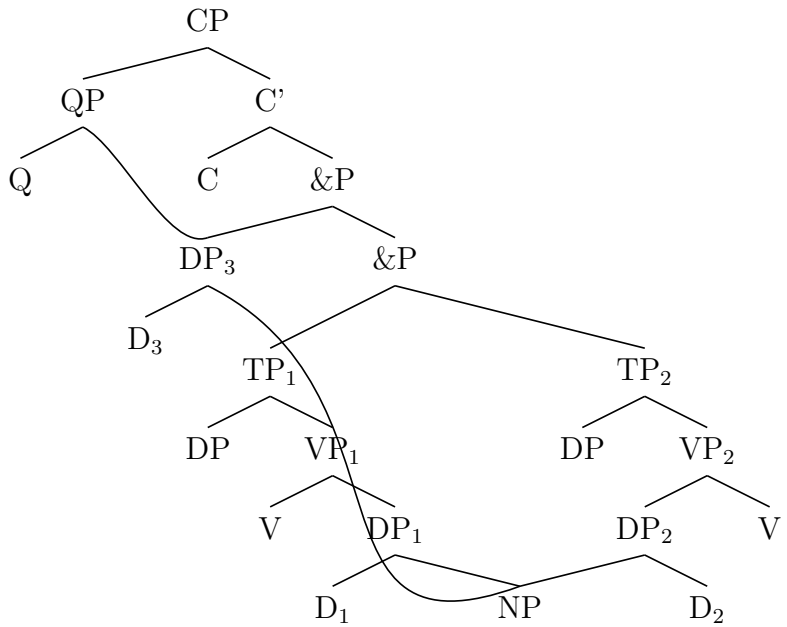
(81)



If Spellout applies to the VPs in this sort of workspace, the NP will of course be able to delay Spellout.

Now imagine that this is an ATB-derivation. Ordinary movement of the DP can take place, first with an intermediate step to the edge of the coordinate structure, and then to the final interpretable position. The resulting structure will look like (82):

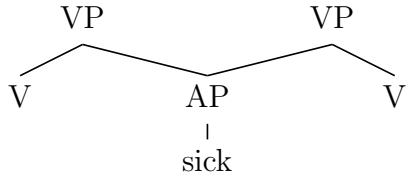
(82)



Even if there is an island inside one or the other conjunct, then movement can still take place, since Spellout of the NP is delayed. However, something is Spelled out: the two determiners. This derives another important observation of this paper: as with non-ATB cases we discussed, if a conjunct contains an island, only the D will be interpreted. This predicts anti-pronominality effects in symbiotic gap sentences for the same reason as in ordinary selective island sentences.

With movement of non-DPs, we predict a structure like the following:

(83)



As with the cases discussed earlier, if COLLECT applies early the shared phrase will simply not receive an interpretation inside of the island. This correctly predicts that symbiotic gap constructions should forbid extraction of non-DP, for the essentially same reason as with the non-sharing cases discussed in **section 1.4**. We also will predict lack of scope reconstruction in these cases, again for the same reason as with extraction of unshared material from islands.

Anti-pronominality or other selective island effects do not appear in *all* ATB-constructions because early COLLECT is *optional*. If COLLECT does not apply, then

each conjunct will be built up in a workspace in which the NP is not completely dominated by the island-forming SOD. Importantly, a full R-expression will be Spelled out in the base-generated position of the movement structure.

5 A tale of two derivations

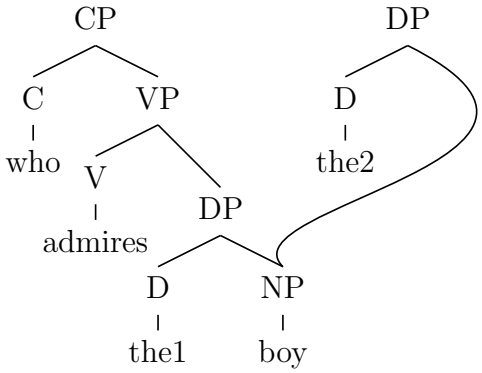
We are still left wondering why extraction of a DP from a pronoun-accepting context across an island should induce a degree of unacceptability which apparently is not found in ATB-extraction. That is, why is there a contrast between (84a) and (84b)?:

(84) a. ??Who did Gilligan talk to someone who admires _?
b. Who did Gilligan talk to someone who admires _ and Mary Ann go swimming with someone who can't stand _?

Or to put things more bluntly: why are there island effects in the first place?

To understand why this might be, consider the derivational histories we have so far outlined. For the non ATB-construction (84a), the crucial derivational stage is the island SOD, repeated below:

(85)



The important question was whether the DP formed by parallel merge is COLLECTed into the same workspace as the island SOD. If it is, and if the DP does not merge into the Specifier,CP position of the relative clause, then a delayed Spellout structure obtains.

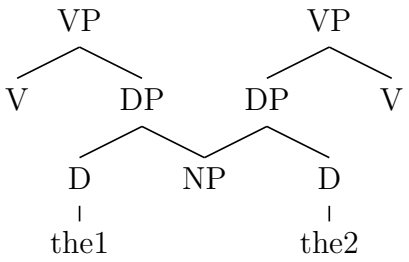
Let us suppose that, when certain phrases occupy a syntactic workspace, they are required to immediately integrate with the rest of the syntactic structure in a particular way. For DPs, this means the following:

(86) If a DP shares a workspace with another structure X, and if MERGE(DP, X) is possible, then MERGE(DP, X) must apply immediately

This constraint is not satisfied for (85): the option is available for the DP to merge into the specifier position of the CP; however, in order for delayed Spellout of the NP to take effect, Spellout must precede Merge(NP,CP).

By constrast, consider an ATB-derivation. Recall that this derivation will form a structure like the following:

(87)



In this sort of derivation, the DPs can undergo immediate MERGE, without eliminating the structural configuration which enables delayed Spellout. In other words, creation of a structure in which delayed Spellout is possible is *not* dependent upon delaying its ability to fulfill its grammatical function, so the same oddness which applies to a structure like (85) will not apply to (87).

But why should there be such a distinction between DPs on the one hand and coordination or adjunction on the other? Perhaps the distinction comes down to whether an instance of MERGE is optional. In many cases, merger of a phrase in a particular location is necessary for a derivation to converge. For instance, arguments must be merged in a phrase where they will be assigned the appropriate thematic role. This also seems to hold of successive cyclic movement operations: McCloskey (2002), for instance, argues that *wh*-phrases are required to merge at the base-position of each phase edge. The sort of constraint posited in (86) might then hold of all obligatory operations. Optional operations, like adjunction and arguably coordination (Graf, 2014), are perhaps simply not subject to this sort of constraint. One could imagine implementing this idea with a distinction between feature-driven MERGE and a non-feature-driven adjunction operation, but the details of such a proposal will be beyond the scope of this paper.

6 Some other issues

In this section, I will address two remaining issues: First, predictions that this theory makes regarding asymmetries between rightward and leftward movement; second, a problem with putative DP/non-DP asymmetries in *wh*-in-situ languages.

6.1 Selective islands and right-node raising

Like its counterpart ATB-movement, right-node raising is generally insensitive to island constraints. However, *unlike* ATB-movement, Postal (1994, 1998) shows that selective island effects typically fail hold of rightward movement:

(88) John knows someone who can speak in _, and Ted knows someone who can understand understand _, several Slavic languages.

Where Postal's analysis must stipulate this difference between leftward and rightward movement, ours naturally predicts it. The theory of cyclic linearization that we are assuming automatically predicts that rightward movement can can generally evade islands, since- as long as the appropriate conditions are met- no ordering contradiction should arise from movement (Sabbagh, 2007). Consequently, COLLECT need not be invoked to avoid islandhood here, and selective island effects should not emerge.

Ordering contradictions only arise, according to Sabbagh, in constructions which violate what he calls the *right edge restriction*. These are cases in which a phrase must move to the right over certain other phrases. This explains the contrast between the following sentences:

(89) a. John ate _ raw some clams that he had bought at Legal Seafood.
b. *John ate _ drunk some clams that he had bought at Legal Seafood.

While we will not go into the details of his analysis here, Sabbagh argues that movement over *drunk* would create an ordering contradiction. In other words, it is subject to the same sort of constraint as leftward movement out of islands. This should predict that, to the extent that the grammar can get around the *right-edge restriction*, it must invoke delayed Spellout of the rightward moved phrase. This in turn predicts that selective island effects to arise:

(90) John can speak ?(*in) _ drunk and Ted can understand _, several Slavic languages.

6.2 *wh*-in-situ

The analysis developed in this paper predicts a DP/non-DP asymmetry in extraction out of islands for languages with overt *wh*-movement, like English. What appears to be a similar contrast shows up in in-situ *wh*-constructions in languages like Chinese. Huang (1982) illustrates this for Chinese:

(91) a. ni xiang-zhidao [shei mai-le sheme]?
you wonder who bought what
'What do you wonder who bought?'
b. ni xiang-zhidao [shei weisheme da-le Zhangsan]?
you wonder who why beat Zhangsan
'Why do you wonder who beat Zhangsan?'

Example (91a) shows that a wh-DP (*sheme*) is acceptable within a *wh*-island in Mandarin. However, in (91b) a non-DP *wh*-phrase *weisheme* is trapped inside of an island. According to Huang, *wheisheme* can only be interpreted within the embedded

clause. A main clause interpretation similar to *why did you wonder who beat Zhangsan _?* is unavailable.

This appears to be a problem for the analysis presented in this paper. We have argued that a PF output constraint, by assumption Fox and Pesetsky's ban on ordering contradictions, forbids movement out of islands from preceding as normal in cases of overt movement. To the extent that overt movement out of islands is possible at all, it is because of delayed Spellout of the *wh*-phrase. Among other things, this derives a DP/non-DP asymmetry for overt movement, since delayed Spellout is only possible for DPs. But in a language where *wh*-phrase remain in-situ, like Chinese, it is unclear why movement out of an island should be problematic at all, since no ordering contradiction should arise in the first place.

I think it is likely that restrictions like those in (91) are ultimately due to the same 'weak island' effects,[16] like negative or factive islands. Abrusán (2014) has argued that 'weak islands' arise from interactions between the semantics of certain sorts of interrogatives and certain presuppositions. As an illustration, consider the islands that factive predicates form for manner adverbials:

(92) #How does Mary regret that John fixed the car?

Abrusán's crucial assumption is that the domain of manners always contains contraries. That is, for any manner X, then some manner denoting the opposite of X also exists (e.g. wisely/unwisely, fast/slow, etc.). Weak island effects arise in the case of manner adveribials because of particular presuppositions. In the case of sentences like (92), the familiar factive presupposition is the culprit. A factive like *Gilligan regrets that he ate the coconut* presupposes that Gilligan actually ate the coconut. Now take a factive in a *wh*-interrogative:

(93) What does Mary regret that John fixed?

Following a long tradition (Hamblin 1973, Kartunnen 1977, etc.), Abrusán assumes that questions denote sets of propositions. In order for a question Q to be felicitous, the presupposition of every proposition in Q must be satisfied. This predicts that for a question like (93), the presupposition introduced will be:

(94) For every contextually salient non-human entity x, John fixed x

[16]Note that weak islands and selective islands are not the same thing. As Abrusán convincingly argues, weak islands are fundamentally semantic, and include such environments as negative islands and factive islands. These environments do not show standard SI effects, such as sensitivity to anti-pronominal contexts:

(1) a. That's the river that Mary didn't have lunch on _
b. That's the river that Mary knows that Bill had lunch on _.

Both of these sentences allow the idiomatic reading which is unavailable in other island contexts.

Returning to (92) this question will presuppose that, for every Manner in the domain, John fixed his car in that manner. Or more precisely, since Abrusán assumes manners to be properties of events:

(95) for every manner X in D_M, the event of John's fixing his car was in X

This is where the assumption about contrariety becomes crucial. Because whenever *fast* is in D_M *slow* is to, this presupposition will be satisfied only if the event of John's fixing his car was simultaneously fast and slow. Since these two manners or incompatible, this presupposition will be unsatisfiable.

This theory makes a prediction which is relevant to our discussion of *wh*-adverbials in Mandarin. *How* questions are actually ambiguous. Consider:

(96) How did John climb Mt. Washington?

This sentence allows two kinds of answers. First, we might respond with a manner (well, fast, etc.). But one can also respond 'He took the Tuckerman Ravine Trail.' Following Tsai (1994), lets call this the 'means' reading. Abrusán's analysis should predict this response to be grammatical inside of a weak island, since the presupposition introduced by the 'means' reading aren't predicted to be contradictory.

For Chinese, Tsai (1994) observes a syntactic correlate to the manner/means ambiguity:

(97) a. Nimen zenmeyang da zhe chang lanqiu
you how play this CL basketball
'By what means do you play this basketball game?'

b. Lanqiu, nimen da- de zenmeyang
basketball you play DE how
'In what manner do you play basketball?'

Tsai reports that while with (6a) speakers prefer the means answer, the post-verbal *zenmeyang* in (6b) renders a manner response obligatory.

Now consider the following pair, in which *zenmeyang* is embedded within a complex NP island:

(98) a. Ni bijiao xihuan [[ta zenmeyang zhu] de cai]?
you more like he how cook PNM dish
'What is the means x s.t. you like better the dishes which he cooks by x?'

b. * Ni bijiao xihuan [[ta zhu- de zenmeyang] de cai]?
you more like he cook DE how PNM dish
'What is the manner x s.t. you like better the dishes which he cooks in x?'

Post-verbal *zenmeyang* is unacceptable, while the preverbal equivalent is grammatical on the means reading. This contrast would make sense if we assume that the unacceptability of certain adverbials is a Abrusánian weak island effect and not fundamentally syntactic. While a full analysis of weak island effects in Mandarin is beyond the scope of this paper, Tsai's observation suggests that a semantic analysis along the lines of Abrusán (2014) is compatible with Huang's data.

At this point, one might wonder whether the DP/non-DP asymmetry discussed extensively in this paper is also simply a weak island effect. I think not, for the following reason. Chomsky (1982) points out that pied-piping out of an island is absolutely forbidden:

(99) a. ?Which girl does John know a guy who talked to _.
b. *to which girl does John know a guy who talked _.

This is an instance of the familiar DP/non-DP asymmetry. It could be explained within our system, assuming that pied-piping is simply merger of Q with a constituent larger than the *wh*-DP (Cable, 2007). It is not explicable in semantic terms, however.

Conclusion

The proposal presented in this paper is ultimately a theory of what can reconstruct where. The central idea is that, to the extent that syntactic movement is possible out of an island, the moved material itself is forbidden from reconstructing. As we have seen, such anti-reconstruction effects follow directly from Bachrach and Katzir's theory of delayed Spellout, once attention is paid to the LF consequences of their proposal.

Now, under most MERGE based theories of movement, a proposal like this one would have disastrous consequences at LF. However, with Johnson's theory of DP movement, the reMERGE operation only targets the NP complement. As a consequence, only this NP fails to reconstruct. This, we have argued, allows a semantically interpretable LF to be formed in select cases, but imposes a variety of restrictions on what can move out of islands. The anti-reconstruction effects perfectly correlate with the selective island effects discussed by Cinque and Postal. We have also discovered that this analysis predicts such effects to appear in Bachrach and Katzir's symbiotic gap constructions. Finally, I have suggested that the difficulty of creating a delayed Spellout environment in DP movement is related to a restriction on merger of DPs.

As a theory of anti-reconstruction effects, my proposal has much in common with the theory of *Wholesale late merger*, as it has been employed by Takahashi and Hulsey (2009) and especially Stanton (2015). This paper's analysis of anti-pronominality is particularly indebted to Stanton's insight that anti-pronominality effects in various types of movement can be explained by restrictions on reconstruction of the NP complement into the tail of a movement chain. However, despite the

underlying similarity between my proposal and Stanton's, the empirical scope of the two proposals is quite different. My proposal covers selective island effects, Stanton's anti-pronominality in certain varieties of $\overline{A}$-movement.

We have seen that this analysis predicts that syntactic reconstruction of the NP is banned in island contexts. Along with everything else, this predicts that reconstruction for the purposes of binding should also be banned. It turns out, however, that things are not as simple as one would hope.

One of the advantages of copy and remerge theory is that they make possible a syntactic theory of reconstruction into binding domains (Chomsky, 1995; Fox, 1999). This allows us to explain why condition C effects emerge in data like (100a) and why binding of a reflexive is possible in (100b):

(100) a. *Which aspect of John$_i$ does he$_i$ despise _?
b. Which aspect of himself does John despise _?

Since our theory predicts that NP complements of moved DPs are not interpreted within islands, we should predict (i) a bleeding of condition C and (ii) the impossibility of reflexive binding into islands. This first prediction appears to be borne out:

(101) ?That's the aspect of John$_{i/j}$ that Mary was surprised because he$_i$ despises _.

This sentence is no less acceptable on the reading in which *John* co-refers with the pronoun than one in which it does not.

Curiously, however, reconstruction of reflexives into islands does not appear any worse than extraction from an island is normally:[17]

(102) a. ?That's the aspect of himself that Mary was surprised because John despises _.
b. ?That's the aspect of Fred that Mary was surprised because John despises _.

While it remains unclear why this should be so, I should note that a similar puzzle arises for Stanton's analysis of anti-pronominality. As we have already mentioned, Stanton argues for a violable constraint which prefers late merge of the NP complement to apply as late in the derivation as possible. This correctly predicts anti-pronominality effects in certain forms of A'-movement.

Stanton's proposal is in line with the predictions of *wholesale late merger* that condition C is bled by A-movement. However, in certain cases Stanton's analysis appears to incorrectly predict bleeding of Condition A, for instance in A-movement:

[17] A similar oddity arises with left dislocation structures in English:

(1) a. ?That's the aspect of John$_i$ that Mary was surprised because he$_i$ despises it.
b. That's the aspect of himself that Mary was surprised because John despises it.

(103) That aspect of himself seems to John _ to be unattractive.

If, following Takahashi & Hulsey, we assume that late merger of the NP is possible in A-movement, then Stanton's late merge constraint should require the copy left behind in the position c-commanded by *John* to be a bare D, without the NP *aspect of himself*. This predicts reflexive binding to be impossible. Instead, Stanton suggests that the preference for late merger can be overridden by semantic requirements (such as a constraint requiring a bound variable or reflexive to be c-commanded by its antecedent at LF). This forces reconstruction where it is otherwise not permitted. The violability of Stanton's late merger constraint is perfectly compatible with the optimality theoretic logic of that system. However, it is unclear how a similar proposal could be made compatible with the architecture assumed in this paper.

Finally, it is worth observing that the symbiotic gap construction discussed here bears a striking resemblance to the more familiar parasitic gap (PG) construction (Engdahl, 1983, a.o.). Both are multi-gap dependencies which permit obviation of island phenomena. Proposals to assimilate PG sentences to the across-the-board construction (Williams, 1990; Nunes, 2001) typically meet with the following objection: PG sentences are subject to a variety of restrictions- in particular selective island effects (Postal, 1993) and reconstruction asymmetries (Nissenbaum, 2000)- which fail to hold of typical ATB-sentences. However, the analysis presented in this paper can be naturally extended to PG sentences, allowing us to capture both ATB and PG sentences with the same basic mechanism (parallel MERGE). As with SG sentence, movement out of the island in PG sentences will entail anti-reconstruction effects, and therefore selective island effects. This will be desirable both conceptually, since it simplifies the grammar, and empirically, since by retaining a common mechanism between PG and ATB-sentences we are able to explain data otherwise troublesome for the alternative null operator analysis, such as similar case matching effects in PG- and ATB-sentences sentences (Franks, 1993).

Appendix: Spellout

(104) Let K be a Spellout node in some workspace W,
Spellout(K, W) initiates the following procedure

1. determine the set K' = {x: x ∈ SOD(K, W) & x is a Spellout node & x ≠ K}
2. PF Spellout:
   (a) Generate Lin(K) = {$x < y$: $x, y \in$ SOD(K, W)} via rules of linearization
   (b) Add Lin(K) to set of previously computed sets of LP statements
3. LF Spellout:

(a) If $x \in$ K', then x is assigned the semantic value assigned to **Spell-out**(x)

(b) If K dominates y & $y \notin$ SOD(K, W), $[\![y]\!]$ is treated as semantically vacuous.

(c) Using output of (a) and (b), compute $[\![$K$]\!]$ using the rules of semantic composition

4. **Spellout**(K, W) = $< Lin$(K), $[\![$K$]\!] >$

The Spellout algorithm will proceed in the following steps: (1) identify the set K' of Spellout nodes dominated by the Spellout node K; (2) compute PF Spellout by generating a set of LP statements, which are then added to the set of LP statements for the Spellout nodes in K'; (3) compute LF Spellout, which uses previously generated semantic values for Spellout nodes in K' and ignores nodes not in K's complete dominance domain; (4) finally the output of **Spellout**(K) should be the pair of LP statements and a semantic value. This definition of **Spellout** should have the effect of deriving the principles in (5) and (6) for PF and LF Spellout.

References

Abrusán, M. (2014). Weak Island Semantics. Oxford University Press.

Bachrach, A. and Katzir, R. (2009). Right-node raising and delayed spellout. In Grohmann, K., editor, InterPhases: Phase-Theoretic Investigations of Linguistic Interfaces. Oxford University Press.

Barss, A. (1986). Chains and anaphoric dependence: On reconstruction and its implications. PhD thesis, MIT, Cambridge, MA.

Beck, S. (2006). Intervention effects follow from focus interpretation. Natural Langauge Semantics, 14:1–56.

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

Chomsky, N. (1982). Some concepts and consequences of the theory of Government and Binding. MIT Press.

Chomsky, N. (1995). The minimalist program. MIT press.

Chomsky, N. (2000). Derivation by phase. In Kenstowicz, M., editor, Ken Hale: A life in language. MIT Press.

Cinque, G. (1990). Types of A'-dependencies. MIT Press, Cambridge, MA.

Citko, B. (2005). On the nature of merge: external merge, internal merge, and parallel merge. Linguistic Inquiry, 36(4):475–496.

Elbourne, P. (2002). Situations and individuals. PhD thesis, MIT, Cambridge, MA.

Elbourne, P. (2005). Situations and individuals. MIT Press, Cambridge, MA.

Engdahl, E. (1983). Parasitic gaps. Linguistics and Philosophy, 6:5–34.

Fox, D. (1999). Reconstruction, binding theory, and the interpretation of chains. Linguistic Inquiry, 30(2):157–196.

Fox, D. (2003). On logical form. In Hendrick, R., editor, Minimalist Syntax. Blackwell.

Fox, D. and Pesetsky, D. (2005). Cyclic linearization of syntactic structure. Theoretical Linguistics, 21(1-2):1–46.

Fox, D. and Pesetsky, D. (2007). Cyclic linearization of shared material. Talk presented at Bogaziçi University, April 6 2007.

Fox, D. and Pesetsky, D. (2009). Rightward movement, covert movement, and cyclic linearization. Lecture at Ben Gurion University.

Franks, S. (1993). On parallelism in across-the-board dependencies. Linguistic Inquiry, 24(3):509–529.

Graf, T. (2014). Adjuncts, islands, algebra (with a sprinkling of semantics). Invited talk, November 4, CUNY Graduate Center, New York, NY.

Heim, I. (2013). Functional readings of questions. MIT.

Huang, C.-T. J. (1982). Move wh in a language without wh-movement. The Linguistic Review, 1:369–416.

Huang, C.-T. J. (1993). Reconstruction and the structure of vp: some theoretical consequences. Linguistic Inquiry, 24:103–138.

Johnson, K. (2012). Towards deriving differences in how *wh*-movement and QR are pronounced. Lingua, 122(6):529–553.

Johnson, K. (2014). Multidominance and movement. lectures presented at the Hebrew University of Jerusalem, December 2014.

Kotek, H. (2014). Composing Questions. PhD thesis, MIT, Cambridge, MA.

Lebeaux, D. (1988). Language acquisition and the form of the grammar. PhD thesis, University of Massachusetts, Amherst.

Longobardi, G. (1987). Extraction from NP and the proper notion of head government. In Giorgi, A. and Longobardi, G., editors, The syntax of noun phrases, pages 57–112. Cambridge University Press.

McCloskey, J. (2002). Resumption, successive cyclicity, and the locality of operations. In Epstein, S. D. and Seely, T. D., editors, Derivation and Explanation in the Minimalist Program. Blackwell.

Nissenbaum, J. (2000). Investigations of Covert phrase movement. PhD thesis, MIT, Cambridge, MA.

Nunes, J. (2001). Sideward movement. Linguistic Inquiry, 32(2):303–344.

Postal, P. (1993). Parasitics gaps and the across-the-board phenomenon. Linguistic Inquiry, 24(4):735–754.

Postal, P. (1994). Parasitic and pseudoparasitic gaps. Linguistic Inquiry, 25(1):63–117.

Postal, P. (1998). Three Investigations of Extraction. MIT Press, Cambridge, MA.

Ross, J. R. (1967). Constraints on variables in syntax. PhD thesis, MIT, Cambridge, MA.

Ruys, E. (2011). Semantic reconstruction and the interpretation of chains. In Reich, I. and et al., editors, Proceedings of Sinn und Bedeutung, volume 15, pages 515–529, Saarbruücken, Germany. Universaar-Saarland University Press.

Sabbagh, J. (2007). Ordering and linearizing rightward movement. Natural Language & Linguistic Theory, 25(2):349–401.

Stanton, J. (2015). Wholesale late merger in A'-movement: evidence from preposition stranding. http://ling.auf.net/lingbuzz/002131. To appear in *Linguistic Inquiry*.

Takahashi, S. and Hulsey, S. (2009). Wholesale Late Merger: Beyond the A/A' distinction. Linguistic Inquiry, 40:487–526.

Tsai, W.-T. D. (1994). On nominal islands and LF extraction in Chinese. Natural Language & Linguistic Theory, 12(1):121–175.

Wilder, C. (1999). Right-node raising and the LCA. In Bird, S., Carnie, A., Haugen, J. D., and Norquest, P., editors, Proceedings of WCCFL 18.

Williams, E. (1990). The ATB theory of parasitic gaps. The Linguistic Review, 6:265–279.