Move Dissolves into Merge: a Theory of Locality

Michal.Starke@lettres.unige.ch

February 2001 - v2.2r4

Please check for newer versions at
http://theoling.auf.net/papers/starke_michal/
before citing, distributing, etc.

Contents

Introduction

Recent syntactic theories within the Principles & Parameters framework (i.a. Kayne (1994), Brody (1995), Chomsky (1995)) recognise at least three types of fundamental relationships between syntactic entities: the complement relationship, the specifier relationship and the movement/chain relationship.

The complement relationship takes two unrelated nodes and brings them together asymmetrically and compositionally:

(1) α
       $\diagup\diagdown$
      α β
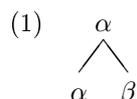
The specifier relationship brings together two nodes containing an identical feature, thereby creating a compositionally vacuous instance of concatenation:

(2) α
       $\diagup\diagdown$
      α α
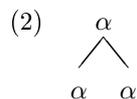
The movement/chain (cum 'trace') relationship signals that two nodes in the tree are to be interpreted as a single node:

(3) x   x

The complement relationship is the core of the grammar organ, underlying the human capacity for recursive compositionality. The specifier relationship and movement/chain relationship on the other hand are surprising constructs: as currently understood, both are interpretatively vacuous and apparently arbitrary properties of the grammar organ.

The general issue underlying the present work is to determine why the grammar organ should be such as to contain precisely the typology of relationships described above, rather than the numerous possible alternatives. In a less neutral formulation: why do the surprising vacuous relationships exist at all?

The general thrust of the answer is that they don't: Given proper scrutiny, both of them are seen to be special cases of the core relationship, (1). When properly construed, human grammar thus only contains one type of relationship between syntactic nodes: complementation.

In the case of 'specifiers', several aspects of the recent and important work on 'functional projections' based on 'spec-head agreement' occurring within them suggests that what has been taken to be a syntagm in a specifier (typically agreeing with a null projecting head) is better rephrased as a projecting syntagm (with no accompanying empty head), §9. (4a) becomes (4b):

3

(4) a.
$$\alpha P$$

YP[+α] αP

α β

b.
$$\alpha P$$

YP[+α] β
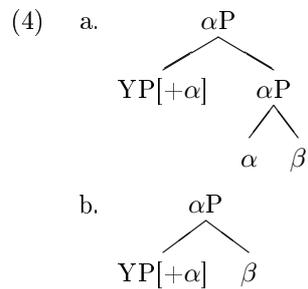
In the case of movement/chain relationship, it turns out that the properties of 'merge' and the properties of 'move' are logically identical once the hardcoded locality of merge and the asymmetric nature of move are taken into account. The two purported relations are thus two facets of a single generalized concatenation operation, §8.

The upshot is that the core 'gluing' operation - merge(α, β) compositionally - may well be the sole operation of the grammar organ.

Unifying merge and move/chain however presupposes a unified notion of move. Locality being the major property of move, this in turn presupposes a unified notion of locality. A significant amount of the effort thus necessarily consists in solving the traditional problem of unification of strong and weak islands (which I take to be the core of the locality issue).

The problem of unifying locality has been the center of intensive scrutiny by a wide array of researchers (e.g. Chomsky (1986), Cinque (1990), Manzini (1992) amongst many others), sharing one general assumption: the solution lies in a refinement of the locality principle itself. Looking at the problem from a different angle however brings a surprisingly simple solution: instead of achieving coverage by refining the locality principle, achieve coverage by refining the data-structure to which the simple and generally accepted principle applies. In practice this involves the hypothesis that syntactic features are organized into a feature-tree, rather than being an unordered set, §1-6.

Putting these results together, the overall picture is that the grammar organ is simpler than commonly assumed with respect to the operations (or principles) it contains, but more fine-grained than commonly assumed with respect to the data-structure it operates on. Encouragingly, this picture is the same picture emerging from recent work on cognitive science in general: the mind/brain involves simple processes but complex data-structures (for one illustrative line of research, see for instance Tanaka's work on visual object-recognition, e.g. Tanaka et al. (1991), Fujita et al. (1992), Kobatake and Tanaka (1994), Tanaka (1996))

1 Unifying Locality

There is a wide consensus that the locality of wh-movement is to be cut into three main generalisations: weak islands (WI), extraction out of weak islands (eWI), and strong islands (SI). Despite long-standing efforts to produce a unified theory of these phenomena, every current approach treats them with a disjunction of tools. A typical situation is that weak islands are explained by a version of *Relativised Minimality* (*Minimal Link Condition, Attract Closest*, etc.), extractions out of weak islands by some form of binding relationships, and strong islands by a version of 'barriers'. In some cases, the disjunction has arguably been brought down to two tools (address-based vs. 'categorial index'-based dependencies, Manzini (1992)), but complete unification of these major locality effects has remained elusive.

I will argue here that *Relativised Minimality* (Rizzi (1990)) is all we need to capture not only weak islands, but also strong islands and extraction out of weak islands. The key ingredient to this result is the postulation of a more refined data-structure: syntactic features are organized into a feature-tree rather than being a haphazard set (a features-tree much like the 'feature-geometry' attributed to phonological features, Clements (1985)).

1.1 Weak Islands and *Relativised Minimality*

Descriptively, WI names the fact that movement of a quantificational element is blocked if it attempts to cross another quantificational element. A wh-movement such as (5) is thus blocked if a negation (6a), a focalised element (6b), another wh-element (6c), or a quantified adverbial, (6d), intervenes (traces are either in angled brackets or as *t* throughout):

(5)   how do you think that I should cook this stuff <how>?

(6)   a.   * how do<u>n't</u> you think that I should cook this stuff <how>?
             * how do you think that I should<u>n't</u> cook this stuff <how>?

      b.   * how do you think that, <u>THIS STUFF</u>, I should cook <how>, not those eggplants over there?

      c.   * how do you wonder <u>why</u> I should cook this stuff <how>?

      d.   ?* how should I <u>often</u> cook this stuff <how>?

Similar facts obtain for focus movement, here in Hungarian (data from G. Puskas):

(7)   János FÉL ÓRÁN BELÜL akarja, hogy el-olvassuk   az   ujságot
      J   half hour   within   wants   that PRT-read-we the newspaper
      It's in a half an hour that Janos wants us to read the paper

(8) a. * János FÉL ÓRÁN BELÜL <u>nem</u> akarja, hogy el-olvassuk   az
J   half hour   within not wants that PRT-read-we the
ujságot   <fél órán belül>
newspaper

b. * János FÉL ÓRÁN BELÜL akarja, hogy <u>AZ UJSÁGOT</u>
J   half hour   within wants that the newspaper
olvassuk el   <fél órán belül>
read-we PRT

c. * János FÉL ÓRÁN BELÜL kérdezte, hogy <u>ki</u>   olvasta el   az
J   half hour   within asked   that who read   PRT the
ujságo   <fél órán belül>t
newspaper

d. * János A   LECKE ELÖTT akarja, hogy <u>gyakran</u> el-olvassuk az
J   the lesson   before   wants that often   read-we   the
ujságot   <a lecke elött>
newspaper
It's before the lesson that Janos wants us to often read the news-
paper

All the ungrammatical configurations reduce to:

(9)   * Q_1 ... Q_2 ... $<Q_1>$

where Q = {wh, neg, foc, quantificational-adverb}.

The first important fact - often missed - is that weak islands are about *classes* of features, not about features themselves: the weak island generalization states that a feature of a given *class* cannot cross *a member of the same class*. One lesson of this generalization is thus that there is at least a one-step hierarchical structure to syntactic features:

(10)

        Q   A
  wh neg foc   person number case
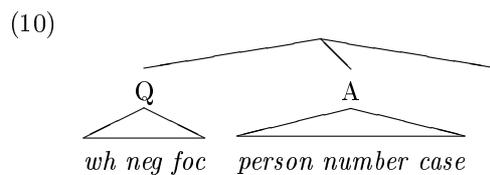
(Calling the class of elements relevant to weak islands 'quantifiers', Q, is some-what inadequate: most items traditionally referred to as quantifiers (*every, some, most, two*, etc.) appear to fall outside. I will however continue to use this term - in this restricted meaning - as no better term has yet emerged.)

The core addition I will make to the framework - from which everything else follows - is that the tree is not as flat as (10). The minimal feature-tree that we will reach will look like this:

6

(11)

```
                  ╱│╲
          Quantifier  θ   A
              │       │
          SpecificQ   iθ
```
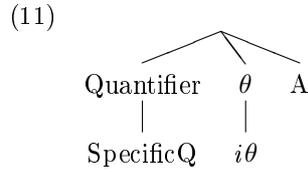
(Where $i\theta$ can be understood to stand for inherent case, and A for structural case or ϕ features.)

The second main lesson traditionally drawn from (9) is that grammar is subject to an abstract anti-identity constraint - two items of the same class cannot 'overlap'. In its simplest form:

(12) * $\alpha \ldots \gamma \ldots \alpha$
      if $\text{class}(\alpha) = \text{class}(\gamma)$ and γ intervenes between the two occurrences of α

Since a syntactic element α can be member of several classes, and the configuration (12) may be legal wrt. one class but illegal wrt another, the anti-identity constraint must be sensitive to the *type* of relation being built. This is a point with deep consequences, discussed in §8. Keeping the definition informal:

(13) X-relating two occurrences of α is legal only if $\alpha \in$ X and there is no γ, $\gamma \in$ X and γ intervenes between the two occurrences of α

('Relating two occurrences of α' can be interpreted either as 'moving from the position of one to the position of the other', or as 'creating a chain between the two'.)

This is the essence of the *Relativised Minimality* condition (Rizzi (1990)), and of all its subsequent adaptations (e.g. *Minimal Link Condition*, Chomsky (1995):296, etc.), although implementation details vary widely.

1.2   The Form of the Solution

Let us suppose that syntactic features form a feature tree, very much in the same sense as phonological features do. Each node in the tree defines a class, containing itself and all the nodes it dominates.

This creates two new situations with respect to the anti-identity principle: cases where the intervener belongs to a subclass of the elements to be related, and cases where the intervener belongs to a superclass of the elements to be related. As an illustration, take a class, C, with $\alpha \in$ C, and x$\notin$C. The classical environments in which the anti-identity principle applies are:

(14) a. $\alpha \ldots$ x $\ldots \alpha$

7

    b. $\alpha_i \ldots \alpha_j \ldots \alpha_i$

where the first is legal and the second illegal. Suppose now that C has a subclass, SC, and $\beta \in$SC. Since SC is a subclass of C, β is a member of both C and SC. To bring this fact out explicitly, let us rename β to $\alpha\beta$. The two new environments for the anti-identity principle are thus:

(15)   a. $\alpha\beta \ldots \alpha \ldots \alpha\beta$
           b. $\alpha \ldots \alpha\beta \ldots \alpha$

The anti-identity principle (13) makes clear-cut predictions for these new cases: (15a) is legal and (15b) illegal.

In (15a) $\alpha\beta$ is a member of both C and SC, and can thus chose to do either C-movement or SC-movement. If C-movement is chosen, the intervener α is of the same class as the attempted movement and triggers a 'C-over-C' configuration which the anti-identity requirement rules out. But if $\alpha\beta$ chooses to travel the SC-road, no intervener of the SC-type intervenes and the movement succeeds. At least one type of movement/chain succeeds and (15a) is thus good.

In (15b) on the other hand: α is only member of class C, and C-movement is the only option. Since the intervener $\alpha\beta$ is a member of C, this movement creates a 'C-over-C' configuration, blocked by the anti-identity principle. C-movement is thus derailed, and (15b) is illegal. (Obviously, here as throughout, the same logic can be couched in terms of elements being attracted, etc.)

The claim developed below is that the configurations (15) are not only logical possibilities, but correspond to actual linguistic constructions: (15a) depicts a successful extraction out of a weak island, and (15b) depicts a strong island.

As a first step in that direction, notice how configurations where α intervenes contrast with configurations where $\alpha\beta$ intervenes:

(16)   a.   *$\alpha_i \ldots \alpha_j \ldots \alpha_i$
           b. $\alpha\beta \ldots \alpha \ldots \alpha\beta$
(17)   a.   *$\alpha \ldots \alpha\beta \ldots \alpha$
           b.   *$\alpha\beta \ldots \alpha\beta \ldots \alpha\beta$

α is a 'selective blocker': it blocks some items but lets others through. Which is what weak islands are all about. $\alpha\beta$ on the other hand is an 'absolute blocker' in (17): nothing can go through. Which is what strong islands are all about.

To put the same thing differently, notice that (16a) is identical to the *Relativised Minimality* configuration (9). Following this trail, we know that an α-intervener is something that provokes a weak island effect. Given this, what (16b) says is that some subset of quantifiers, with special additional properties, will be able

to escape weak islands. Which precisely corresponds to the generalization about eWI.

Continuing on that trail, (17) says that if something stronger than a weak island intervenes, the game is over and we fall into strong islands: nothing will be able to extract anymore (unless a further subclass comes into play, of course).

At its core, this is it. A single principle deriving the various islands. But islands have a rich underground life, and implementation will prove to be more fine-grained than the above shows; although the core logic will remain unchanged.

What we will be doing now is to switch to an essentially descriptive mode, and bring out the generalisations underlying strong islands and extractions out of weak islands, in order to show that these generalisations essentially reduce to the logic above. Once this is over, we can take the game to the next step and bring out a deeper result: the same logic leads us to unify move and merge - or more perspicuously, dissolve move into merge (In representational terms, showing that 'chains' and 'phrase-structure' are one and the same concept).

2 SpecificQ-movement: Unifying Extractions out of Weak Islands with Standard Locality

At first sight, 'extraction out of weak islands' (eWI) is a label for a set of exceptions: configurations which should be blocked by (12) but are not, eg:[1]

(18)   what do you wonder <u>whether</u> John will cook <what>?

There is however an underlying generalization to these 'exceptions'. In its most general formulation: elements which can successfully extract out of weak islands have 'something more' than those which cannot.

While this generic formulation of the generalization is consensual, opinions diverge on what the 'something more' is: θ-role, ie. the distinction is between arguments and non-arguments (Huang (1982), Lasnik and Saito (1984), Rizzi (1988)); case or DPhood (Manzini (1992), Rizzi (2000)); 'referentiality', as opposed to purely quantificational readings (Cinque (1990), Rizzi (1990)); d-linking (Cinque (1990), Kroch (1989), Comorovski (1996)); individuation (Frampton (1991), Cresti (1995)), richness of internal semantic structure (Szabolcsi and Zwarts (1997)).

Notice however the logic of the situation: the uncontroversial generic formulation of the eWI generalization precisely corresponds to (16b). Both say that those elements which have 'something more' can extract.

To see this, let us first temporarily label β the additional property of arguments. Then, remember that (16a) corresponds to the classical cases of *Relativised Minimality*, and therefore $\alpha = Q$. (16) thus translates into:

(19)   a.   * $Q_i \ldots Q_j \ldots Q_i$
       b. $Q\beta \ldots Q \ldots Q\beta$

Now (19b) is precisely what the eWi generalisations says: wh-phrases which can extract are those that have some additional property, β.

In other words, all the ingredients to explain extraction out of weak islands are already present in standard theories: some wh-phrases (Q) have an additional

---

[1] A note of caution about judgements on eWI: there appears to be rather extensive variation across speakers in degree of degradation assigned to eWI. The variation is however restricted to the amount of the degradation: there is no variation as to the existence (or direction) of the asymmetries. All speakers who see a difference between argument and adjunct eWI report that adjuncts are less acceptable, although how much less is subject to variation (with the limiting case of the less liberal speakers who have severe difficulties parsing argument eWI and thus see no difference with adjunct eWI). Similarly, no speaker sees eWI as better than unhindered extraction: it is always worse, but how much worse is subject to variation.

Since only the asymmetries themselves are relevant to us, this variation is not of direct concern, although interesting *per se*. This does however mean that some speakers do not perceive the asymmetries discussed below, all examples being unparsable to them.

property β, and applying *Relativised Minimality* to those phrases yields the result that Qβ can jump over an intervening Q because Qβ-movement only cares about intervening Qβ elements, (19b).

But what is β?

2.1 Weak Islands and Existence

It's getting late for your little Joey, so you decide to bring him to bed and read a story to him. Part of the story involves the following:

(20) Belgamore and Belfedore lost their dog, and have been unsuccessfully looking for it for 3 days. On the fourth day, Belgamore decides to go out again and continue looking for any clue. Belfedore, tired and despaired, gives up and stays at home. In the evening, Belgamore comes back very excited and . . . .

At this point, Joey interjects:

(21)   I wonder what Belga found!

So you stop reading and ask:

(22)   a.   . . . and what do you think that Belgamore discovered?
        b.   # . . . and what do you wonder whether Belgamore discovered?

(the notation '#' indicates a sentence which is grammatical but inappropriate as a continuation of the specified context)

Why is (22a) felicitous but (22b) distinctively odd in this context? Given (21) you have good grounds to think that Joey is wondering what Belgamore found. So why is it that you can know that Joey wonders about something but you cannot ask what that something is?

The central property which sets apart (22b) from both (21) and (22a) is the presence of an extraction out of a weak island. It thus looks like eWI is triggering the inappropriateness of (22b) forbidding you to ask the relevant question.

That the inappropriateness is indeed independent of lexical semantics and due to the underlying grammatical structure is further suggested by (23): minimally tweaking the structure of the felicitous (22a) makes it become inappropriate too, without relevant changes in lexical semantics.

(23)   a.   # what is it that you think Belgamore discovered?
             # what is the stuff that you think Belgamore discovered?

b. # what is it that you wonder whether Belgamore discovered?

what is the stuff that you wonder whether Belgamore discovered?

Inserting the questions (22) into a cleft/relative makes them become uniformly odd as a continuation of the above scenario; and in fact the judgment of oddity of (23) is very similar to that of (22b). Descriptively, the cleft/relative structure imposes the 'special', inappropriate, meaning of eWI on all wh-movement. The 'special' meaning is thus not directly tied to the lexical semantics of the *wonder/think* contrast, but is rather triggered by some property of the underlying structure.

Simple, run of the mill, extractions out of weak islands therefore have a distinct 'meaning' with respect to unhindered extractions. More generally, what this tells us is that wh-movement has underlyingly access to two options: one option triggers the felicitous reading but does not survive weak islands; the other triggers the inappropriate reading and survives extraction out of weak islands.

This situation is straightforward to map onto the $Q\beta$ terminology: $Q\beta$-movement survives weak islands but triggers the inappropriate meaning while Q-movement is trapped inside weak islands and triggers the felicitous meaning. Furthermore, clefts and relatives involve $Q\beta$-movement. Going over the above examples, (22a) involves Q-movement and thus a felicitous semantics; (22b) involves $Q\beta$-movement and thus the 'special' (inappropriate) reading; the reason why *wonder* in (21) does not involve inappropriate reading is that it has Q-movement (no islands crossed); and (23) involves clefts/relatives, hence $Q\beta$-movement with the inappropriate 'special' semantics.

What is this special meaning of extractions out of weak islands? And why does it hang onto eWI?

To get a grip on the relevant requirement, consider the intuitions about what goes wrong with the eWI question above. Paraphrasing the intuitions underlying the impossibility of (22b) runs like this: while you have good grounds to think that Joey is wondering what Belgamore found, you have no reason to think that Joey has any clue about what it is that Belgamore actually found. Therefore you cannot felicitously presuppose that there is some 'thing-potentially-found-by-Belgamore' that Joey is wondering about, and ask him to give you the identity of that thing. In short, the intuitive reading of (22b) is that Joey has in mind some object-potentially-found, an illegitimate assumption in this context.

This is confirmed by this minimal pair context which renders eWI felicitous:

(24) Joey interrupts you as above but now says: 'I wonder what Belga found! Could it be... ?' and stops in the middle of the sentence, looking at you starry-eyed. Again, you stop reading and ask:

a. so? what do you think that Belgamore discovered?

so? what is it that you think that Belgamore discovered?

>
> b. so? what do you wonder whether Belgamore discovered?
>
> so? what is it that you wonder whether Belgamore discovered?

The sole difference is Joey's additional 'Could it be... ?', and somehow this makes the eWI (and cleft) questions felicitous. Again, the intuition is clear: Joey's additional utterance legitimates the assumption that he has something on his mind as an object-of-discovery.

A first stab at the generalization is thus that eWI can be used only if there are reasons to believe that there exist some entity which the interlocutor has in mind as a referent for the wh-phrase. The unhindered extraction on the other hand requires no such existential commitment: it can be asked in situations where it is clear that the interlocutor has no clue about the referent for the wh-phrase.

Let us call this asymmetry the 'prior belief' asymmetry: used with a psychological predicate, eWI is felicitous only if it ascribes prior beliefs about the referent of the wh-phrase to the clausal subject.

As an additional illustration, consider the "guess contexts":

> (25) I know that you have no clue about what Herbert will cook tonight, and that you're curious about it; so tell me:
>
> a. what do you hope that Herbert will cook?
>
> b. # what do you wonder whether Herbert will cook?

Again, unhindered extractions can be used to prompt for a guess, eWI cannot.

To make this more precise, assume that the wh-phrase extracting out of a weak island drags along an existential presupposition, while no such presupposition is present on the unhindered wh-phrase (an assumption to be refined shortly). The two configurations above can then be characterized in terms of the operator/scope technology as:

> (26) a. $\exists x \ldots$ belief $[\ldots x \ldots]$
>
> b. $\ldots$ belief $[\ldots x \ldots]$

(26a) - informally paraphrased as the fact that there exists an object such that there are beliefs about that object - is required by eWI, while (26b) - involving no existential at all - is a sufficient context for an unhindered wh-movement question.

It is important that $\exists$ have wide-scope; (27) is not what we're looking for:

> (27) belief $[\exists x \ldots x \ldots]$

This configuration - informally paraphrased as 'there is something that Belgamore discovered and I wonder what that thing is' - is not sufficient to legitimate eWI. The original Belgamore situation (22) legitimates the embedded-clause inference that 'there is something that Belgamore discovered' but (22b) is nonetheless infelicitous. eWI requires not only an existential presupposition, but an existential presupposition with wide scope.

The picture thus seems clean: the difference between Q- and Qβ-movement is that Qβ-movement carries existential presupposition and Q-movement doesn't.

Simple Clauses

The same contrast holds in simple clauses, in the absence of a psychological predicate. In the context of the Belgamore story above, simple-clause wh-movement is legitimate only with Q-movement:

(28)   a.   what did Belgamore discover?
       b.   # what is it that Belgamore discovered?
            # what is the thing that Belgamore discovered?
       c.   # what didn't Belgamore discover?

but with a richer context legitimating existential presupposition, all questions become licit:

(29)   Charles and Herbert are worried. In order to win the Tetrapon game, they need to discover two more facts about the great founder of Tetrapon Inc. They have already found the answer to the other 15 quizzes, but two questions elude them, and today is the last day. Unfortunately, Herbert is victim of his bimensual temporary sickness and cannot participate. Charles sets out to town to try and discover those two facts. His efforts are long and many, but as he comes back in the evening, he tells Herbert: 'look, I've only discovered one of the two needed facts...'. Herbert, gloomily looks at him from his bed and asks:

       a.   so, what did you discover?
       b.   so, what is it that you discovered?
            so, what is the thing that you discovered?
       c.   so, what didn't you discover?

Unsurprisingly, this context is strong enough to license the Qβ-movement in (29b) and (29c) as it provides an explicit discovered-thing to hang the existential on. (29b) is naturally interpreted as a request for the identity of the fact that Charles did find, and (29c) is natural as a request for the identity of that last painful unanswered quizz.

It is thus generally the case that Qβ-movement collapses as soon as existential presuppositions disappear. Q-movement on the other hand survives unaffected. We now have good grounds to assume this simple generalization:

(30) Qβ-movement carries existential presupposition but Q-movement doesn't

The Existential YoYo

The 'scope' effect abstractly depicted in (26-27) can be shown overtly. As a preliminary, consider the cleft/relative construction, which allows for the cleft/relative to be inserted either in the root CP or in intermediate CPs. The possibility of *wh in-situ* within the French cleft makes this entirely transparent:

(31) a. c'est **quoi** que tu crois que Jean a   cuisiné?
          it-is what that you think that John has cooked

      b. tu crois que c'est **quoi** que Jean a   cuisiné?
          you think that it-is what that John has cooked

Their meanings vary as depicted by the scope logic (26): both presuppose that 'there is something that John cooked' but (31a) additionally presupposes that 'there is something that you believe that John cooked'. (31b) does not make this additional presupposition. I.e. (31a) corresponds to (26a) and (31b) to (27).

But both are genuine questions: the scope of the question is the whole sentence. It is thus possible to disconnect the scope of the existential presupposition from the scope of the question carrying it.

The English translation shows the same meaning alternation, but absence of *wh in-situ* means that the dissociation between the scope of the question and the scope of the existential presupposition is now overt: the position of the cleft itself corresponds to the scope of the existential presupposition, and the position of the wh-phrase corresponds to the scope of the question.

(32) a. **what is it** that you think diddle cooked?
      b. **what** do you think **it is** diddle cooked?

As expected, the meaning difference can be brought out in the open by the 'prior-belief' test:

(33) I know that you have no clue about it, but I'd like your guess:
      a. # what **is it** that you think diddle cooked?
      b. what do you think **it is** diddle cooked?

The existential presupposition is thus not due to some generic filter on question-formation: it has it's own life which can be dissociated from the question as such, and it can choose to sit in various places of the syntactic structure.

How can we make sense of this? By assumption, clefts involve Qβ-movement, so there is a Qβ link at least in the root CP of (32a) and in the embedded CP of (32b). The further movement of *what* in (32b) encounters no more islands, so in principle it has the option to do Q-movement. Under that option, the relevant representations are:

(34)   a.   Qβ ... <Qβ> ... (=32a)
          b.   Q ... <Qβ> ... (=32b)

(to be precise, the nature of the intermediate link for (34a) is unspecified - and irrelevant here).

Since Q-movement does not involve existential presupposition but Qβ does, these representations make the right cut: the presupposition is in the root clause in (32a) but in the intermediate clause in (32b). In other words, the generalization (30) seems to be correct in its literal interpretation: the scope of the existential presupposition corresponds to the highest Qβ link, not to the movement of the wh-phrase *per se*.

We now have a prediction on our hands: if the lack of existential import of the upstairs movement in (32b) is an indication that this is Q-movement, this movement should be allergic to weak islands, even though the wh-phrase looks like a prototypical argument. Additionally, this should contrast with the embedded trajectory of the wh-phrase; the latter being Qβ, it should be able to jump over weak islands. Which is correct:

(35)   a.   * what don't you think it is that diddle cooked?
          b.   what do you think it is that diddle didn't cook?

Furthermore, the same negation as (35a) should become grammatical with the upstairs cleft, since upstairs cleft do Qβ-movement all along. Which is again correct:

(36)   what is it that you don't think that diddle cooked?

The correlation between sensitivity to weak islands and lack of existential import also arises with the French downstairs cleft: the *in-situ* wh-phrase is sensitive to weak islands above it. (Inserting the same negation with the cleft upstairs is fine.)

(37)   a.   * tu   crois <u>pas</u> que c'est **quoi** que Jean a   cuisiné?
                you think not that it-is what that John has cooked

b. tu crois que c'est **quoi** que Jean a <u>pas</u> cuisiné?
  you think that it-is what that John has not cooked

As elaborated in more details below, I'll take this to indicate the presence of covert Q-movement: again, sensitive to weak islands but without existential import. As a side-note, notice again that framing the issue in terms of movement and derivation is inessential to its logic, the same result is achieved if an empty Q-operator is base-generated in the higher position and forms a chain with the overt position (Brody (1995)).

(The obvious remaining question with respect to the downstairs cleft is: why is it restricted to Q-movement? What we have explained is that given Q-movement, there is a correlation between WI-sensitivity and lack of existential import; but the reason for it to be restricted to Q-movement is missing. The key observation here is made by Baltin (1992): wh-movement of predicates is always sensitive to weak islands, ie. predicates are restricted to Q-movement. See also Rizzi (1992)).

Returning to straight wh-questions, we now have a strong prediction: in a long-distance question, placing an island downstairs or upstairs should alter the scope of the existential presupposition. The following paradigm is what we are looking for:

(38) a. what don't you think that Charles discovered?
   b. what do you think that Charles didn't discover?

As expected, (38a) is only felicitous if there are reasonable grounds to suppose that 'you have something in mind, such that you don't think that Charles discovered that thing'. (38b) on the other hand makes more modest existential requests (that 'there is something, such that Charles didn't discover that thing'). In other words, (38b) can be used in a "guess context", but (38a) cannot:

(39) I know that you have no clue about it, but I'd like your guess:
   a. # what don't you think that Charles discovered?
   b. what you think that Charles didn't discover?

As a more thorough illustration consider the following two contexts. Suppose you are reading to your child the story of Charles and Herbert, (29), and imagine that the story actually says what the two questions that Charles is trying to answer are. As the story comes to the point where Charles comes home and announces his failure, you stop reading, turn to your child and ask one of (38). In this context, asking (38b) is perfectly reasonable, although neither you nor the child has any clue about the answer. The question is simply a prompt for a guess. Asking (38a) on the other hand is inappropriate; it has the out-of-place presupposition that your child has some belief about Charles' discovery and you want to tap that belief.

In short, the existential presupposition has wide scope in (38a) but not in (38b).

To explain these contrasts, the same logic as above applies: in (38b) the highest weak island is in the embedded clause, forcing Qβ-movement downstairs. But the upstairs part of the wh-movement is unencumbered by islands and is thus free to travel the Q-movement road.

(38b) thus allows configuration 'Q ... $<$Q$\beta>$ ... '. (38a) on the other hand has a weak island in the root clause and thus forces Qβ-movement all the way up: 'Qβ ... $<$Q$\beta>$ ... ':

(40)   a.   what$_{Q\beta}$ don't you think $<$what$>$ that Charles discovered?
       b.   what$_Q$ do you think $<$what$_{Q\beta}>$ that Charles didn't discover?

Again, the generalization is that the scope of existential presupposition is determined by the location of the highest Qβ, not by the final step of the wh-movement.

(Notice incidentally that the felicity of (21) in the Belgamore context shows that indirect questions also travel the Q-road, triggering no existential requirements, a fact simply assumed above.)

Cashing In

Despite the subtle character of the interpretations involved in these paradigms, the cake clearly reduces to the simple statement (30), rephrased here as (41), on the assumption that Qβ-movement can be followed by Q-movement (if no more islands or clefts force it to do otherwise).

(41)   the highest Qβ landing site must involve an existential in its interpretation

The existential presupposition faithfully follows Qβ links, and stops wherever Qβ-movement stops. Since Q-links never involve existential presuppositions, this statement reduces to:

(42)   Qβ links involve an existential presupposition

Which in turn reduces to 'β = existential presupposition'. We seem to hold our fish. But fish is a delicate thing to hold, and it turns out that this one is still slippery.

2.2 Range versus Specificity

What is the existential import of Qβ-movement? Facts have been presented above as if this was a unified phenomenon across eWI. But this is doubtful. First, eWI comes in two 'degrees' of deviance:

(43)   a.   what is it unclear whether we should repair?

         b.   what is it unclear how we should repair?

Systematically, eWI of the type (43a) are felt to be slightly less degraded than (43b). This difference disappears in eWI out of infinitive clauses, where both types of extraction pattern with (43a):

(44)   a.   what is it unclear whether to repair?

         b.   what is it unclear how to repair?

This pattern is rather subtle, but surprisingly stable across languages, in fact across families of languages with otherwise widely distinct syntactic profiles. Where does the difference come from?

Even more subtle - but equally systematic - is the accompanying interpretive intuition that (43a) and (43b) require a slightly different type of context and interpretation on the extracted wh-phrase. The raw intuition underlying this difference, is that extractions of the type (43b) require presupposition of a more 'specific' referent than (43a). In fact questions of the type (43b) are felt to go in the direction of echo-questions - the specific interpretation *par excellence* - without however falling all the way into an echo interpretation.

In this light, notice how the Belgamore and Charles stories above legitimate their eWI in two different types of contexts: in the case of Belgamore, (24), the wh-phrase has a *specific* antecedent (the entity that Joey is wondering about) whereas in the case of Charles, there is no such specific antecedent. Conversely, the Charles story provides a *range* for the wh-phrase (namely the two unanswered puzzles) while the Belgamore story doesn't. This difference can be highlighted for instance by the fact that questions such as (45) can be asked in the Charles context, but not in the (updated) Belgamore context (24):

(45)   what is it unclear whether he found?

That the subtle difference between (43a-43b) maps on a range-based versus specificity-based presupposition is confirmed by the following 'list' settings, although intuitions remain very subtle.

(46)   You are a car mechanic working in a garage; one morning, as you come late, you hear that the list of cars to be repaired might have been reshuffled. To get up to date, you ask:

which car is it now unclear whether we should repair?

This does not presuppose the existence of a 'specific' or particular car. It rather presupposes that there is a predefined set of cars, and asks for a list within those cars. Things become different when the intervener changes:

(47) You are a car mechanic working in a special garage catering to customers who are picky about how their car is repaired. To this effect, the garage keeps a list of cars paired with how their owners want them repaired. One morning as you come in, you hear that the list has been made partially unreadable by a water infiltration so that some cars have lost their how-to instructions. To get up to date on the situation you ask:

which car is it now unclear how we should repair?

In such contexts, with a range but without any specific antecedent, eWI with whether-interveners are distinctively more felicitous than eWI with how- or why-interveners. The latter trigger the odd presupposition (in this context) that there was a specific car that is under discussion and that you inquire about the identity of that specific car.

Put differently, eWI with whether-interveners can be used as generic information-request questions, given a clear range. But eWI with how-interveners require a stronger context to be licensed, such that a specific entity is available as a potential antecedent. (Notice that (47) becomes felicitous again if the wh-phrase is put in the plural - *which cars* - and this has the effect of creating a specific antecedent for it: the contextually predefined set of cars whose instructions have been affected by the water leak).

It is however not the case that the milder - 'range based' - type of presupposition is a simple existential presupposition. If the discourse is rarefied so much as to forbid any range or specific antecedent while still allowing for existential inferences, all eWI collapse, including the more modest ones:

(48) Balthazar and Sibilian are activists who share the rather banal assumption that humans tend to cheat each other in order to maintain their privileges, and the somewhat controversial assumption that people in the elite are clever enough to be unpredictable as to who they will cheat next. In short, they both believe that 'there is someone such that it is {plausible that | unclear whether} the elite will screw that someone next'. As Balthazar meets Sibilian on the street, he starts the conversation by asking:

a.   so, who do you think that the elite will cheat next?
b. ?? so, according to you, who is it plausible/probable that the elite will cheat next?
c. ?* according to you, who is it unclear whether the elite will cheat next?

From this paradigm, it does appear that the modest brand of eWI questions suffocates in arid discourse environments. Existential presupposition *per se* is not enough, some discourse-help in the form of a range or a specific entity is needed in order to survive ((48b) is a weak island induced by the predicate 'plausible/probable').

Let us thus tentatively agree that the differential degradation of (43a-43b) correlates with a 'range-based' versus 'specificity-based' presupposition requirement - as useful descriptive labels if nothing else.

(A word on terminology: 'specificity' is used here in a more restricted sense than in some of the literature on the topic. In Enç (1991) for instance, both range-based and specificity-based antecedents qualify as specific, although they are distinguished as two different types of specificity. 'Range-based' wh-phrases correspond to Enç's partitive-specifics, while specificity-based wh-phrases seem to correspond to Enç's familiar-specifics. I will use the more restrictive terminology throughout. The previous literature on eWI does not draw the distinction and alternates between the two notions: terms such as 'd-linking' (Cinque (1990)) or 'specificity' (Kiss (1993)) seem to map onto the specificity-based requirement, while Rizzi (2000) talks in terms of range. The various formulations in terms of individuation (Frampton (1991), Cresti (1995)) seem to also map onto the range-based requirement.)

This suggests an unexpected complication: there are two distinct routes out of weak islands, one triggering 'range' semantics, and the other 'specificity' semantics. Furthermore these routes are sensitive to the nature of the intervener.

Here we surprisingly land back into familiar territory: what has been called above 'whether-type' eWI versus 'how-type' eWI looks like the classical 'argument/adjunct' asymmetry. Testing the full range of wh-phrase as interveners yields the progressive degradation familiar from studies of extracted phrases and discussed in detail below. As an illustration, in (49), argument intervention causes the gentler kind of degradation, while intervention of any adjunct causes more difficulty.

(49)   a.   what is it unclear whether we should repair?
            to whom is it unclear what we should give?
       b.   what is it unclear when we should repair?
            what is it unclear how we should repair?
            what is it unclear why we should repair?

We thus have a series of new issues to confront: why does the distinction exist in the first place? why does it disappear in infinitives? why is the argument/adjunct nature of the intervener relevant? And above all, if range and specificity are two routes out of weak islands, where do their respective classes fit in the syntactic feature tree?

French wh *in situ* (and the *in situ* wh-phrase of a multiple-wh construction) gives a clear indication that the subclass of Q that we have been calling β is about specificity (in the restricted sense), not range. Range will turn out to be central, but in a different quality.

2.3 Specificity and the Locality of wh *in-situ*

At first sight, French wh *in situ* has the paradoxical property of dying under weak islands, but surviving under strong islands.[2]

The baseline case corresponding to unhindered long-distance questions is (50):

(50) a.   tu   crois   qu'elle   a   fait   quoi?
           you think that she has done what?

      b.   tu   crois   qu'elle   s'appelle comment?
           you think that she is-called how?
           what do you think her name is?

(For some speakers - to whom I belong - adjuncts sometimes require a slightly different intonation, under unclear conditions; other speakers make no difference. There is however no sharp arguments/adjunct asymmetry. See below for more on intonation and adjuncts.)

(51) illustrates the unfazed survival of the situ-wh under strong islands, here an adverbial clause, a relative clause within a noun-phrase and a coordinated syntagm ('situ-wh' is used throughout as a shortcut for 'wh-phrase in situ').

(51) a.   tu crois qu'elle a dit ça pour inciter Pierrot à séduire qui?
           you think that she said this to incite P to seduce whom?

      b.   tu crois qu'elle a dit ça pour inciter Pierrot à les inviter quand?
           you think that she said this to incite P to invite them when?

      c.   tu   crois   qu'ils   vont inviter ceux qui ont   fait   quoi?
           you think that they will invite those that have done what

      d.   tu   crois   qu'ils   vont rembourser ceux qui ont   voyagé
           you think that they will reimburse   those that have travelled
           comment?
           how

      e.   tu   crois   qu'il   a   acheté des carottes et   quoi d'autre?
           you think that he has bought carrots   and what else?

---

[2]French judgements below (and throughout) belong to informal spoken French.
This may be significant, to explain some discrepancy of judgements about French wh *in situ*. See §3.5.1 for discussion.

Inserting a weak island yields a sharp degradation, regardless of the presence of a strong island, and regardless of the exact position of the weak island above the situ-wh (this is illustrated here with negation, but any other weak-island inducer provokes the same effect, modulo intervening wh-phrases which trigger absorption, illustrated below):

(52)   a.   tu   crois   qu'elle   a   <u>pas</u> fait   quoi?
                you think that she has not done what?

        b.   tu   crois   qu'elle   veut   <u>pas</u> partir   comment?
                you think that she wants not to-leave how?

        c.   tu crois qu'elle a <u>pas</u> dit ça pour inciter Pierrot à séduire qui?
                you think that she hasn't said this to incite P to seduce whom?

        d.   tu   crois   qu'ils   vont <u>pas</u> inviter ceux   qui   ont   fait   quoi?
                you think that they will   not invite those that have done what

        e.   tu   crois   qu'ils   vont <u>pas</u> rembourser ceux   qui   ont
                you think that they will   not reimburse   those that have
                voyagé   comment?
                travelled how

        f.   tu   crois   qu'il   a <u>pas</u>   acheté   des carottes et   quoi   d'autre?
                you think that he hasn't bought carrots   and what else?

The straightforward interpretation of these paradigms is that we are looking at a 'pure' case of *Relativised Minimality*: situ-wh are entities blocked by any intervening Q, and by nothing else. Since strong islands do not involve Q, they do not block situ-wh.

The parallelism with standard cases of extraction becomes striking once we observe that examples (52) are in fact not ungrammatical, despite the description above. A more careful assessment of the facts is that (52) is sharply ungrammatical under the standard downfall intonation felicitous in (50-51). But with a slight accent on the situ-wh, (52) is acceptable.

Furthermore, this accented situ-wh triggers a presuppositional interpretation similar to the interpretation of eWI discussed above. (Of course, this is not the echo-reading. Both the intonation and the interpretation are distinct from the echo-reading).[3]

Weak islands also induce the familiar argument/adjunct asymmetry (while intervening strong islands do not, as shown above):[4]

---

[3]With situ-wh under strong islands, as in (51), it is often difficult to decide whether the wh-phrase requires the additional accentuation or not. Correspondingly, the contrast with insertion of negation is less sharp. I will work with the idealisation that SI do not induce special accent, but nothing relevant would change if they do, as will become clear immediately.

[4]To evaluate the French examples, it is important to keep in mind that the "wh est-ce que" construction does not trigger the cleft semantics, in contrast with "wh c'est que" and "c'est wh que". It rather patterns with simple wh-movement, while the latter two pattern with clefts.

(53) a. qu'est ce que tu as   pas acheté?
        what is it that you have not bought

    b. * comment t'es   pas parti?
        how   you are not left
        how didn't you leave?

(54) a. t'as   pas acheté quoi?
        you have not bought what

    b. ?? t'es   pas parti comment?
        you are not left   how

The parallelism between overt and situ-wh extends to the fact that the asymmetry becomes weak - or disappears - under modals:

(55) a. qu'est ce que tu as   pas voulu   acheter?
        what is it that you have not wanted to-buy

    b. comment t'as   pas voulu   partir?
        how   you have not wanted to-leave

(56) a. t'as   pas voulu   acheter quoi?
        you have not wanted to-buy what

    b. t'as   pas voulu   partir   comment?
        you have not wanted to-leave how

(One difference is however that the situ-adjunct under negation, (54b), is slightly less degraded than the corresponding moved wh. It rather patterns with the corresponding clefted adjunct:

(57) ?? c'est comment que t'es   pas parti?
        it is how   that you are not left

This pattern is discussed below, (62)).

The situ-wh thus not only have the same sensitivity to Q-interveners as overt wh-movement does, but they also have the same capacity of jumping over Q if they become presuppositional.

Let us take this generalization at face value and suppose that (French) covert wh-movement is a pure case of the anti-identity constraint discussed above. If so, the reason why strong islands do not block situ-wh is that strong islands contain no relevant intervener (no Q, or no $Q\beta$ intervener). The reason why any Q above the situ-wh blocks it, is the familiar *Relativised Minimality* Q-crossing-Q effect; and the reason why a presuppositional situ-wh can jump over an intervening Q is another instance of the logic developed above - the 'richer' $Q\beta$ can afford to snob the 'poorer' Q.

The relevant representations (with English words) look like this:

24

(58) a. $<\mathbf{wh}_Q>$ you think that she has done $\mathbf{what}_Q$?

b. $<\mathbf{wh}_Q>$ you think she said this to incite P to seduce $\mathbf{whom}_Q$?

c. * $<\mathbf{wh}_Q>$ you think that she has $\underline{\text{not}}_Q$ done $\mathbf{what}_Q$?

d. * $<\mathbf{wh}_Q>$ you think she has $\underline{\text{not}}_Q$ said this to incite P to seduce $\mathbf{whom}_Q$?

e. $<\mathbf{wh}_{Q\beta}>$ you think that she has $\underline{\text{not}}_Q$ done $\mathbf{what}_{Q\beta}$?

f. $<\mathbf{wh}_{Q\beta}>$ you think that she has $\underline{\text{not}}_Q$ said this to incite P to seduce $\mathbf{whom}_{Q\beta}$?

Is this presuppositionality effect about 'specificity' or 'range'? One environment where no situ-wh ever survives is specific noun phrases:

(59) * tu   aimerais   avoir   cette/ma photo   de qui?
      you would-like to-have this/my   picture of whom?

situ-wh do however survive within noun phrases with an overt range:

(60) a. tu   aimerais   avoir   une des   photos de qui?
        you would-like to-have one of-the pictures of whom

b. tu aimerais avoir une photo de qui, parmi les photos disponibles?

you would like to have a picture of whom, among the available pictures

Clearly, given the logic of subset and supersets applied to the anti-identity principle, this contrast tells us that β = specificity, and not range. The 'pure' quantifier movement can chose to involve a 'specific quantifier', in order to jump over Q, but it is helpless when a 'specific Q' intervenes.

The relevant configurations are (with 'SQ'= Specific Quantifier):

(61) a. * $<\mathbf{wh}_{SQ}>$ you would like to have $\underline{\text{my}}_{SQ}$ picture of $\mathbf{whom}_{SQ}$?

b. $<\mathbf{wh}_{SQ}>$ you would like to have $\underline{\text{one}}_Q$ of the pictures of $\mathbf{whom}_{SQ}$?

The same conclusion follows from the interpretation of the situ-wh under WI, (52). In all cases, the natural contexts for such questions are contexts in which a specific entity is presupposed, and not contexts in which the situ-wh is simply given a plausible range. There is thus a subtle contrast between (62a-b):

(62) a. qui (est-ce qu') elle voulait pas inviter?

b. elle voulait pas inviter qui?

c. c'est qui qu'elle voulait pas inviter?

There is a clear distinction between (62a) and (62c) and interpretatively the situ-wh, (62b), patterns with (62c) rather than (62a). (Boeckx (2000) makes a similar observation, albeit apparently in a different register/dialect of French with more restrictive situ-wh properties (cf §3.5.1)).

The difference between (62a) and (62c) is the difference described above: (62c) asks for the identity of a *specific* entity in previous discourse, while (62a) can be used as an information-question if a *range* is explicitly or implicitly provided. The fact that the situ-wh patterns with (62c) and not (62a) is a reflex of the fact that the situ-wh is most natural in a context where the discourse makes it clear that a specific entity is asked about, whereas overt movement (62a) is also natural in a 'list'-like context simply providing a range for the wh-phrase.

All of this follows if covert-movement is a pure Q-class movement and SpecificQ is a subclass of Q. Covert movement can thus afford to ignore strong islands, but its only means to jump over weak islands is to travel the SQ route. Overt movement by contrast has access to an additional range-based route (to which we will come back below), and therefore does not entail specificity.

From these paradigms, I thus conclude that the syntactic feature tree includes:

(63)

Quantifier   case/ϕ
   |
SpecificQ
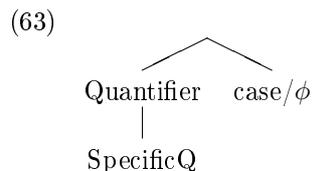
The 'Definiteness' Island

As a welcome side-effect, this model of eWI gives us an explanation of the 'definiteness'-island. It is a common observation that 'definite' noun phrases block overt wh-extraction more severely than indefinite noun-phrases:

(64)   a.   who did you want to buy a picture of?
       b. ?* who did you want to buy the picture of?

As observed *inter alia* in Enç (1991), this so-called 'definiteness' effect is in fact a 'specificity' effect as non-specific definites do not trigger the effect, and specific indefinites do. (65a) illustrates definites which are interpretatively not specific, and do not block overt wh-extraction; (65b) illustrates the converse case of indefinites which are interpreted as specific and do block extraction:

(65)   a.   who did they announce the death of?
             which film did you miss the first part of?
       b. ?* who did you want to buy a certain picture of?

26

Given the syntactic feature-tree (63), this follows automatically, in the same way as it did for covert Q-movement: wh-phrases within a noun-phrase must cross the specific determiner on their way to the root CP but the specific determiner is an SQ intervener which kills both Q-movement and SQ-movement (Qβ-movement). The extracted wh-phrase therefore has no Q-based route to travel, on its way to the root CP, whence the 'specificity' island.

(Notice that we still need to explain why a wh-phrase extracting from within a noun-phrase cannot travel the range-based route, if the latter turns out to be a genuine option to circumvent eWI. This is addressed in §4, devoted to the range-based route.)

Multiple-wh Questions

The French situ-wh paradigms which led us here are not specific to French or to situ-wh questions: the same facts obtain with multiple-wh questions in cases where the non-first wh-phrase remains *in situ*, such as:

(66) a. qui croit que Marie a fait quoi?
          who thinks that Mary did what?

In such cases, the situ-wh behaves in the same way as the French situ-wh described above. First, it doesn't fear strong islands:

(67) a. qui a invité Pierre pour attirer qui?
          who invited Peter in order to attract whom?
      b. qui a besoin d'un homme qui peut faire quoi?
          who needs a man that can do what?

Second, it degrades under weak islands, unless an accent (non-echo) is put on the situ-wh and a presuppositional reading is associated with it; regardless of the presence of a strong island (and again situ-wh under strong islands are somewhat intermediate intonation-wise).

(68) a. qui croit que Marie n'a pas fait quoi?
          who thinks that Mary didn't do what?
      b. qui a invité Pierre pour ne pas effrayer qui?
          who invited Peter in order not to scare whom?

(notice that this gives rise to a bizarre interpretive situation when a pair-list reading is attempted: the situ-wh forces a presuppositional reading which does not mix easily with the pair-list reading; the interpretation whereby there is a specific referent for the situ-wh in each pair is difficult at best. The sentences are however clearly grammatical, regardless of this interpretive quandary.)

Finally, as before the situ-wh does not survive being inside a specific DP:

(69)   a.   * qui voulait acheter ma/cette photo de qui?
            who wanted to buy my/this picture of whom?

Again, the straightforward interpretation is that the situ-wh makes a pure Q-style covert movement, and the syntactic feature tree is (63).[5]

To summarize, we now have the first part of the cake: we have good reasons to think that eWI involves configurations of the type 'SQ ... Q ... SQ', which allows us to capture both 'normal' unhindered wh-movement and extractions out of weak islands with a single locality principle variously called *Relativised Minimality*, *Minimal Link Condition* or *Attract* α. This unification of these two locality domains additionally gives us a unified account of the so-called 'definiteness' island.

Theoretically, the important point is that this unification did not necessitate any enrichment of the model: all is derived with the standard locality principle. The sole enhancement is outside of the syntactic engine itself, in the data structure on which the engine operates: the syntactic feature-tree gains one leaf.

Of course, given this state of affairs, the two prominent questions on the agenda are (i) the nature of the difference between overt and covert wh-movement, such that strong islands block overt but not covert wh-movement and (ii) the role of the range-based route out of weak islands.

---

[5]This paradigm is not particular to French. Exactly the same obtains for instance with multiple-wh questions in Swiss-German, and more generally in southern varieties of German. (see §3.6 for the difference between German varieties).

3 Strong Islands: θ-across-θ

The obvious conclusion from the situ-wh paradigms, is that overt wh-movement cannot be a pure-Q business. Overt wh-movement must involve something else on top of Q, such that this something induces sensitivity to strong islands. What is this something?

3.1 Outcasts

How Come   A fact little noted in the relevant literature is that some wh-phrases cannot be moved at all. Consider the following question:

(70)   how can this cost so much now? It was still affordable yesterday!

Any movement of this brand of wh-phrases is prohibited:

(71)   * how do you think that this can cost so much now? It was still afford-able yesterday!

Under all available theories of wh-movement, this is unexpected: 'how' is a *bona fide* wh-phrase and no island occurs anywhere; the basic Q-movement (or wh-movement) should thus be able to transport it without trouble.

The unmovability of *how* in (70-71) correlates with its interpretation: it does not receive the 'manner' or 'instrumental' interpretations typical of *how*. Rather, it receives a causal-like reading: how did the fact come about.

This reading is lexicalised in English with *how come*, which is again unmovable:

(72)   a.   how come this cost so much now?
         b.   * how come you think that <how come> this cost so much now?

The sole legal reading of (72b) connects the wh-phrase to the matrix clause; it is entirely impossible to interpret it as related to the embedded clause. (If 'come' is analyzed as a separate verb rather than part of a 'how come' idiom, then the relevant example is of the form *'how do you think that (come) this (come) costs so much now', which is a word-salad.)

This causal reading of *how* is 'fact' related rather than 'event' related: it links to the whole proposition, rather than qualifying the event/state within the proposition. More simply, it is not linked to the verb/predicate in the same way as the typical readings of *how* are.

In the French counterpart of this construction *how* cannot remain *in situ*,

(73)   a.   comment ça   se   fait,   ce   genre de chose?
                 how   this self does, this type   of thing
                 how does one construct this type of thing?

       b.   ça se fait comment, ce genre de chose?

(74)   a.   comment ça se fait, que ça coute si cher?
                 how this self does, that it costs so much
                 how can it be that this costs so much?

       b.   * ça se fait comment, que ça coute si cher?

This contrast seems to correlate with the fact that *comment* is connected to predicate in (73b) but not in (74b): the in-situ position is impossible in (74b) because the wh-phrase didn't start in that position, not being a modifier of the predicate.

The generalization that suggests itself then is that wh-phrases can move only if directly connected to the predicate. They cannot move - and are directly generated in the left periphery - if they relate to the entire proposition.

Why$_{cause}$ A confirmation of this generalization comes from another paradigm of the same family: one reading of 'why' is entirely unmovable. Here judgments are more delicate, plausibly because of the pervasive ambiguity with legal readings, but the situation seems to be the same. *Why* has two readings, one I will call 'motivation' (or 'intentional cause') reading and the other I will call the 'cause' reading for lack of a better term. The distinction between the two readings is easiest to see when both cooccur in the same clause:

(75)   a.   Sissy woke up early in order to see the sunrise because she needed some comforting

             The parliament voted this in order to prevent them from rising out of their condition because it has no interest in that happening

       b.   why did she wake up early?

             ↪ in order to see the sunrise

             ↪ because she needed some comforting

In (75a), the motivation reading is given by the 'in order to' adverbial, while the 'cause' reading is given by the 'because' adverbial ('because' can however have both readings, in principle). (75b) shows that *why* can correspond to either the motivation-phrase (*why*$_{motivation}$) or the cause-phrase (*why*$_{cause}$).

That this alternation corresponds to a structural difference is shown by the following ambiguity:

(76)   Brutus didn't go to Rome because his wheel broke

a. ↪ B didn't go to Rome and the reason for this is that his wheel broke

b. ↪ B did go to Rome and the reason for this is unrelated to the fact that his wheel broke

At first sight, this looks like a pure scopal ambiguity between negation and the reason-adjuncts, confirming that the *because* adjunct can be either high or low.

There is however a further meaning difference: (76b) must be interpreted as *why*$_{motivation}$ ('what motivated Brutus to go there is not the fact that his wheel broke'), whereas (76a) can be interpreted either as *why*$_{motivation}$ or as *why*$_{cause}$ ('what prevented Brutus from going to Rome is the breaking of the wheel, regardless of any intention in Brutus').

The low-scope position is thus necessarily interpreted as 'motivation', related to the predicate; while the high-scope position has a 'cause' reading, related to the whole clause. In pseudo-formalism:

(77)   a. (not B go Rome) causedBy: wheel broke
       b. B go Rome (not (because$_{motivation}$ wheel broke))

(Notice also how the intonation differs in the two readings: in the cause reading, the adjunct is intonationally separated from the rest of the sentence; in the motivation reading intonation doesn't separate the adjunct from the rest of the sentence. Intonation thus mimics the semantics of the two readings.)

Again then, fact-related readings are high, predicate-related readings are low.

Correspondingly, the locality of *why*$_{motivation}$ and of *why*$_{cause}$ are not equal. A first indication of this is given by their sharply contrasting reaction to the presence of an adjunct corresponding to the other:

(78)   a.   why$_{cause}$ did Sissy wake up early [$_{motivation}$ in order to see the sunrise]? Because she needed some comforting!
       b.   * why$_{motivation}$ did Sissy wake up early [$_{cause}$ because she needed some comforting]? In order to see the sunrise!

Why is this? The ban on the movement of why$_{motivation}$, (78b), seems to be an instance of a general ban on adjunct wh-movement over because-phrases:

(79)   a.   * how did Sissy behave because she likes him?
       b.   * when did Sissy wake up because she was sad? At 6:00 in the morning.
       c.   ? what did Sissy buy because she was feeling blue?

Why is the fact-related reading of *why* immune to this ban? Given the above reasoning, the reason is transparent: why_{cause} is fact-related and thus base-generated in the left periphery, while $why_{motivation}$ on the other hand originates in VP and must thus move up to the left periphery. In virtue of this, only why_{cause} circumvents whatever locality effect underlies the movement into the left periphery in (78b).[6]

The first point of relevance to our argument is that we now have an additional paradigm indicating the correlation between fact-relatedness of a wh-phrase and absence of movement (i.e. base-generation in the left-periphery).

---

[6]The ban on adjunct wh-movement over because seems to generalize well beyond *because*-phrases. The manner-of-doing reading of *how* is for instance excluded in (1), unless the temporal adverbial is dislocated:

(1) comment est-ce qu'elle l'a fait à deux heure?

    how did she do this at two o'clock?

It thus appears that temporal adverbials cause the same breakdown of adjunct wh-movement as *because*-phrases do.

The generalization is however not entirely clear, as other manner-of-doing examples do not seem to require dislocation of the temporal (although the intonational judgment is difficult):

(2) comment est-ce qu'elle s'est comportée à deux heure?

    how did she behave at two o'clock?

The delineation of this difference - if real - is unclear. Setting this difficulty temporarily aside, it is tempting to suppose that the reason for the argument/adjunct asymmetry in (78) is that arguments originate above the adverbials, while the blocked adjunct-wh originates lower than the adverbial phrase.

Support from this conjecture comes from the observation that reversing the roles makes the blocking effect vanish. No dislocation is necessary for the predicate-related interpretation of the moved wh-phrase here:

(3) quand est-ce qu'elle l'a fait avec attention?

    when did she do this with care?

(4) $why_{motivation}$ did Sissy wake up at six?

This line of thought implies that intervening adverbial phrases block wh-movement in full generality. Further support for this comes from the following paradigm, where an adverbial blocks wh-movement from a lower clause (see Kayne (1984), Cinque (1990), for discussion of similar facts):

(5)   what did he say that John would buy?

(6)   ?* what did he say loudly that John would buy?

(7)   ?* what did he shout that John would buy?

(This paradigm also shows that there is no intrinsic argument/adjunct asymmetry here: if the situation is arranged so that the argument originates from below the adverbial phrase, it is blocked too.)

The (potential) generalization that adjunct phrases systematically block wh-movement seems unexpected, it will become somewhat more understandable in the course of §4.

The second point of relevance is that why_{cause} seems to resist any movement in distance questions. The judgments are difficult since the root-clause reading is often difficult to distinguish from the embedded reading. Here is a case that seems reasonably clear:

(80)   a.   why does 2*3 = 3*2? Because of commutativity
         b.   * why do you hope that <why> 2*3 = 3*2?

In (80a) only why_{cause} is (pragmatically) plausible. When extraction is attempted, as in (80b), the embedded reading disappears: no meaning along the lines of 'I hope that the reason for 2*3 = 3*2 is commutativity and not transitivity' is available. Only a cause-of-hope reading is grammatical.

This contrasts with $why_{motivation}$ which does survive extraction, as in the classical example:

(81)   ? why do you hope that they fired him?

which can be interpreted either as asking for a hope-motivation or for a firing-motivation, albeit with some difficulty.

A more involved paradigm showing the unmovability of why_{cause} is given by the following dialogue:

(82)   Well, there's at least one reason that would never explain why the president helped this tiny minority.
          Ah!? so, why isn't it possible that he did this?
         a.   To collect more votes.
         b. # Because he is a moral man.

Here, the answer (82b) is entirely unparsable, in contrast to (82a). Both answers are however logically fine: the paraphrase 'you cannot explain his actions by supposing he wanted to collect more votes (there aren't enough votes there)' and the paraphrase 'you cannot explain his action by supposing he is a moral man (he is not)' are both perfectly plausible in this context.

The asymmetry follows if $why_{motivation}$ can move but why_{cause} cannot: the question involving a long-distance movement, it only receives the $why_{motivation}$ interpretation.[7]

---

[7]One might object that although motivation-reading is indeed more accessible than the cause-reading in the above eWI; it is often still worse than the corresponding *how* question. This is true, but plausibly due to an artifact: *why* is systematically ambiguous with a root-clause reading, while *how* is not. This is why *is possible* is used as the predicate in the text example, attempting to discourage root clause construals. In fact, if *how* is made as ambiguous as *why*, the judgments become equivalently bad, to my ear (despite a helping context):

When A third set of facts showing the frozen nature of fact-related phrases is given by *when* in (79b). It is entirely ungrammatical under the reading given above, but it does have a grammatical reading:

(83) when did Sissy wake up because she was sad?
    a.   * At 6:00 in the morning.
    b.   Last summer, her sadness used to make her wake up all the time.

In the second interpretation, *when* does not modify the predicate (it does not ask for a time of waking up); it rather modifies the whole proposition, asking when the whole fact happened. This is then another occurrence of the same pattern: non predicate-related wh-phrases are generated upstairs (thus escaping low islands), and cannot move. The following thus doesn't have any embedded reading (and forcing one only leads to the predicate-related meaning).

(84)   * when do you hope that <when> Sissy woke up because she was sad?

The generalization seems to be that wh-phrases unrelated to the predicate cannot move at all (and are base-generated in the left-periphery of the clause).

Outcasts To ease manipulation of these facts, let us introduce some terminology. First, let us agree to enlarge the use 'theta-role' to all relations between a predicate and its modifiers. The more traditional use of 'theta-role', denoting the relation between a predicate and its arguments, can be renamed to 'argumental theta-role', if need for that notion arises. In this terminology, manner-adverbs, temporals, motivations, etc. all have a theta-role (regardless of their argumental status). Second, let us dub 'outcasts' those elements that do not have a theta-role, and 'insiders' those that do.

The generalization above can now be restated as 'wh-phrases without a theta-role cannot move (regardless of islands)', or equivalently, 'outcasts cannot move'.

The fact observed by Manzini (1992) that some nominal 'adjuncts' are entirely unmovable is plausibly another instance of this generalization:

(85)   a.   * avec combien de trous est-ce que tu aimerais acheter [des jeans <avec combien de trous>] ?
          * with how many holes did you buy a pair of jeans?
    b.   * avec quels articles est-ce que tu veux acheter un journal <avec quels articles> ?

---

(1)   a.   how do you know that John decided to do it <how>?
      b.   why do you regret that John decided to do it <why>?

It is a widespread - though curious - fact that some but not all inacceptability judgments seem to be 'washed out' when an acceptable reading of the sentence is available.

           * with what articles do you want to buy a journal?

Which sharply contrasts with:

(86)   ? de qui est-ce que tu veux acheter [une photo <de qui>] ?
           of whom do you want to buy a picture

Under the hypothesis that nominal predicates differ from verbal predicates in their θ-assignment capacities (along the lines of Grimshaw (1990)) such that 'of John' receives a θ-role in 'photo of John' but 'with many holes' doesn't in 'a jeans with many holes', this is another illustration of the unmovability of outcasts.

Why are theta-less wh-phrases unmovable then? The more general underlying question, is: why should theta-roles matter to Q-mv at all?

On the one hand, the relevance of θ-roles to wh-movement is unexpected, as it goes against the grain of all current approaches to (non-island) wh-movement which view wh-movement as a pure Q-business. On the other hand, there is a striking similarity with the conclusion reached earlier on the basis of the situ-wh paradigms: the latter prompted us to believe that overt wh-movement is not pure Q-movement, but involves 'something more'. The outcast-paradigm carries us down the same path: overt wh-movement seems to be 'something more' than pure Q-movement. The additional contribution of outcasts is that they pinpoint this 'something more' to θ-roles.

Can it be that this is also what is involved in the strong island sensitivity of overt wh-movement? I.e. is the 'something more' involved in the overt/covert-movement distinction also θ-roles?

If it was, we would expect that outcasts can do covert wh-movement, even though they cannot do overt wh-movement. This prediction is unfortunately untestable with verbal outcasts, since outcasts can never be *in situ* in the VP zone (being generated in the left periphery), and wh-entities in the left periphery are never 'in situ', i.e. they always require to be in the locus of the question-scope.

Nominal outcasts might be taken to provide evidence that covert-movement is possible, as the following are fine:

(87)   a.   tu veux acheter des jeans avec combien de trous?
             you want to buy jeans with how many holes
       b.   tu veux acheter un journal avec quels articles?
             you want to buy a journal with what articles?

At this stage of the game such evidence is however not conclusive: it can be reasonably objected that covert movement is not of the outcast itself, but of a

subconstituent of it. Therefore such paradigms don't tell us whether the outcast itself is covertly movable. And since nominal outcasts are never single-word wh-phrases, this objection deflagrates onto the whole paradigm. (Later, it will become clear that this *is* in fact a good illustration of the LF/PF difference.)

In the meantime we cannot decide whether the 'something more' causing overt movement to bump into strong islands is the same 'something more' that causes outcasts to be unmovable, namely θ-roles. For evidence on this point, let us turn to strong islands themselves.

3.2   Who is a Strong Island?

The traditional approach to strong islands, since Ross (1967), is to assume that all syntactic nodes are transparent for extraction, and list the nodes that are not transparent. The resulting list looks like haphazard set of entities, and theories of locality proceed to painstakingly attribute some special property to each member of that set.

As noticed by Cinque (1977), Cattell (1978), inverting the logic provides a cleaner picture. Starting from the assumption that all syntactic nodes are opaque for extraction and listing the exceptions provides a coherent set. In fact, Cinque argued that it provides a singleton set: only objects are nodes that allow extraction from within themselves.[8]

If this is so, the goal of a theory of strong islands is to explain what makes the object node special. (It is important to keep in mind that a 'transparent node' is a node that *can* be transparent in principle, not a node which is always transparent. The object node is for instance not always transparent: it becomes opaque - or partially opaque - under various circumstances.)

The conclusion that objects are the sole transparent node is however somewhat of an idealization. Here are three other classes of transparent nodes.

Preverbal DP subjects   Many cases of extraction from within the subject are similar to extractions from within the object (e.g. Rizzi (1982):82 for Italian). There is thus only a slight difference in the following pair (in fact, some French speakers see no difference at all):

(88)   a.   de quel film est-ce que tu as raté [la première partie t]?
            which film did you miss [the first part of t]?
       b.   de quel film est-ce que tu crois que [la première partie t] va créer un scandale?

---

[8]To be precise, Cinque extends the set to all arguments respecting the branchingness-direction of the language, on the basis of clausal-subject extraction in SOV languages. Mentioning argumenthood goes in the right direction, as discussed below, but it casts the net too wide: only a subset of arguments can ever be transparent.

which film do you think that [the first part of t] will cause a big
turmoil?

This is in sharp contrast with extractions out of strong islands - which pro-
voke severe ungrammaticality - and with cases such as (85) above, which are
unparsable.

Even the otherwise fragile subextraction of *combien* survives movement out of
a subject noun-phrases. It provokes a slightly stronger degradation than seen
above, but still in clear contrast with strong island (SI) extraction:

(89)   a.   ? combien   tu   crois   que [t de films] sont arrivés?
              how-many you think that [t of films] are   arrived

        b.   combien   tu   crois qu'elle   critique [t de films] ?
             how-many you think that'she criticizes [t of films]

Two important provisos are however in order: first intonation is important with
subject subextraction. (89a) is only acceptable if the subject is part of the same
intonational phrase as the preceding complementiser. Furthermore, phonology
distinguishes the two cases via the direction of assimilation: in the sequence
'que de films', either the first or the second schwa can drop, and dropping the
first makes the subject and object sequence similar: /kdəfi/. But the sequence
comes out differently in the subject and object position: in the subject position,
there is an obligatory backwards voicing assimilation resulting in /gdəfi/, while
in the object position voicing assimilation may be either forwards, producing
/ktəfi/ or backwards, producing /gdəfi/. Although the reason for these intona-
tional/phonological differences is unclear, keeping them in mind is important
for evaluation of the grammaticality judgments.

The second quirk is that stativity of the predicate matters: extracting from
within the subject of a stative predicate is worse than extracting from within
an eventive predicate. The amount of degradation seems to vary with speakers,
and the exact boundaries of the phenomenon are still unclear to me:

(90)   a.   ? combien   tu   crois   que [t de films] sont arrivés?
              how-many you think that [t of films] are   arrived

        b. ?* combien   tu   crois   que [t de films] sont longs?
              how-many you think that [t of films] are long

        c.   de quel film est-ce que tu crois que [la première partie t] va créer
              un scandale?
              of which film do you think that the first part will create a scandal

        d. ?? de quel film est-ce que tu as dis que [la première partie t] est un
              scandale?
              of which film do you think that the first part is a scandal

Factoring these aspects out, the subject node is another node which is not opaque. It does however become entirely opaque in some constructions and some languages: clausal subjects provoke a much sharper degradation, essentially similar to SI, and languages such as Czech and Spanish seem to disallow extraction from preverbal subject entirely (English is an interesting intermediate case, with contradictory facts reported). I will come back to this important variation, and also to the source of the slight degradation in French and Italian, but let us pursue our logic first: the subject node is to be added to the set of possibly transparent nodes.

Dative-shifted DPs   A third such node is the dative-shifted English noun-phrase:

(91)   a.   (?) which department did you give this professor [the address of t]?
     b.   ? which department did you give [a professor of t] the manuscript about tax reform?

(92)   a.   ? which company did he send [an employee of his] [a nasty email about t]?
     b.   ? which company did he send [an employee of t] a nasty email?

Again, the difference with object extraction is only light (if any at all, and with some inter-speaker variation, for which see §3.5.2).

DPs in specCP   A fourth transparent node is specCP filled by a wh-DP, allowing subparts of the wh to continue with independent wh-movement:

(93)   a.   ? de qui est-ce qu'il est pas clair [combien de photos <de qui>] il veut prendre t?
     b.   ? who is it unclear [how many picture of t] he wants to shoot t?

DPs with Inherent Case   At this point, it might seem that transparent nodes are simply all DPs. But this is not the case: languages with case morphology make a split between DPs with structural and inherent case. Only noun phrases with structural case are transparent for extraction, inherent case creates a strong island. Here is the Czech paradigm:

(94)   a.   (?) ktereho   doktora   to byla chyba?
          which$_{gen}$ doctor$_{gen}$ it was fault$_{nom}$
     b.   ? ktereho   herce   by   sis rad   koupil obrazek?
         which$_{gen}$ actor$_{gen}$ would you gladly buy   picture$_{acc}$

    c.   * ktereho herce   bys   rad vynadal priteli?
        which$_{gen}$ actor$_{gen}$ would you gladly   scold   friend$_{dat}$

    d.   * ktereho herce   se bojis pritele?
        which$_{gen}$ actor$_{gen}$ you fear   friend$_{gen}$

The degradation of extraction from within datives and genitives is similar to
the degradation of extraction from within prepositional phrases (SI):

(95)   * ktereho herce si mluvil s pritelem?
          which$_{gen}$ actor$_{gen}$ self you.spoke with friend

The same is true of German topicalisation/focalisation, (Müller (1995)), where
the structural nominative and accusative pattern against the inherent genitive
and dative:[9]

(96) a.   über Scrambling hat er einem Buch über Optimalität [einen
         about scrambling has he a   book about optimality [an$_{acc}$
         Aufsatz t] hinzugefügt?
         article t] added
         he has contributed a paper about scrambling to a book on opti-
         mality

      b.   * über Optimalität hat er einen Aufsatz über Scrambling
         about optimality has he an   article about scrambling
         [einem Buch t] hinzugefügt?
         [a$_{dat}$ book t] added
         he has contributed a paper about scrambling to a book on opti-
         mality

(Notice that datives pattern with inherent case in Czech/Slovak and German,
but with structural case in the English object shift construction. Similarly,
genitive is structural in noun-phrases but inherent in clauses. See Bayer et al.
(1998) for a good overview of the structural/inherent distinction in German
together with experimental evidence confirming it.)

An updated Cattell/Cinque generalization is thus that: 'everything is a strong
island, except DPs with structural case'.

The difference between structural and inherent case is standardly taken to be
association with thematic roles: structural case is a case that is not associated
to a thematic role but merely a function of a geometric relation with a licenser,
while inherent case also involves a theta role on the case-marked phrase. As-
suming this distinction to be valid (which is not obvious by any means, as for

---

[9]The same seems to be true of Modern Greek, extracting from within accusative versus gen-
itive noun-phrases (E. Anagnostopoulou p.c.). The paradigms however involve complications
difficult to tease out.

instance Czech or Slovak inherent cases are far from being restricted to a single thematic role), let us take it in a very literal reading: structural case involves purely geometrical interpretation, and is entirely unrelated to theta-roles. I.e. syntagms with structural case do not carry any thematic role with them. Inherent case on the other hand is the spell-out of a thematic role and hence is always 'associated' to a θ-role (these assumptions are developed in more detail below).

Given this apparatus, the generalization about strong islands amounts to: only θ-less phrases are transparent. Conversely: θ-marked syntagms are strong islands.[10]

3.3   Implementing θ-movement

Putting the results from the previous sections together with this new generalization, we obtain the following picture:

(97)   a.   only θ-marked phrases can be (overtly) wh-moved

   b. θ-marked phrases are strong islands for overt wh-movement

   c.   (covert) Q-movement does not respect strong islands

---

[10]This entails that outcasts (being θ-less) should be transparent for extraction. This is mostly untestable and to the extent it can be tested the results go in the right direction, but probably for irrelevant reasons. Obviously, only multi-word outcasts are candidates for testability, and those are all prepositional or adverbial phrases. In most languages these phrases disallow movement of their complements for reasons which seem to have nothing to do with locality as understood here, but rather for reasons similar to the complementiser blocking extraction of its complement (the 'that-t' effect). Testability is thus restricted to those languages which void the P-trace effect, i.e. those that have P-stranding: Scandinavian languages and English. Those language however allow P-stranding in presence of a θ-role, as in 'what did you do it with', in which the instrumentally marked 'with which' still allows extraction of 'which'. Given this, the fact that P-stranding out of outcasts is in principle available tells us little about the influence of the θ-less status of outcasts on their transparency.

Notice that P-stranding facts with outcasts are somewhat complicated. In Swedish, P-stranding is uniformly available and sharply contrasts with attempts to move the whole outcast, which yields gibberish (data from Verner Egerland and Anna-Lena Wiklund):

(1)   a.   * med hur många knappar köpte   du   en skjorta
           with how many   buttons bought you a   shirt

   b.   hur   många knappar köpte   du   en skjorta med
       how many   buttons bought you a   shirt   with

In English, P-stranding is more restricted - in unclear ways - and only some outcasts allow P-stranding. Even then, speakers are divided. The following example was described as having a 'very clear contrast' by some speakers and 'no contrast at all' by others (the contrast seems to become clearer with contrastive focus on the subject):

(2)   a.   * with which kind of roof do you plan to build a house?

   b.   % which kind of roof do *you* plan to build a house with?

The obvious conclusion from this is that overt wh-movement involves movement of θ, i.e. θ-movement.

Under this scenario,

- the fact that only θ-bearers can move (overtly), (97a), stems from the fact that (overt) wh-movement involves θ-movement;

- the fact that overt movement is blocked by strong islands, (97b), translates into the fact that θ-movement is blocked by θ-interveners, a *Relativised Minimality* effect

- finally, the immunity of covert Q-movement to strong islands, (97c), derives from the fact that covert movement does not drag θ along.

Since treating overt wh-movement as θ-movement derives the three families of facts discussed above (situ-wh, outcasts, limitation of strong islands to inherent case), while integrating it with weak islands, eWI, and the 'definiteness effect', let us adopt this line of thought and make it more precise.

Theta Phrases. First, if θ is visible to the anti-identity principle, it must be represented structurally. I thus assume that every functional sequence is terminated by a theta-projection (put differently, the highest projection of each 'Extended Projection' is a theta projection). Incorporating the now-traditional idea that there is a KP high up in the nominal sequence, and a QP somewhere below it, a noun phrases looks like this:

(98)
```
        θP
       /  \
      θ    KP
          /  \
         K    QP
             /  \
            Q    DP
                 /\
                N
```
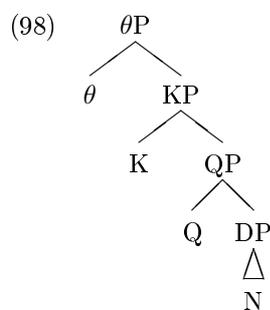
Similarly, a clause is a θP dominating a CP, etc. (In the richer representation of Rizzi (1997), it is tempting to postulate that θP *is* CP, which he takes to be the highest node of a rich sequence of projections, with wh-movement landing in a lower QP-like position). Given these representations, any θ-bearer crossing a filled θP triggers a minimality violation.

Furthermore, if θ is represented structurally, it must be 'checked' somewhere (or, in the approach adopted in §9 below, it must be 'projected' somewhere rather

than checked; both are equivalent for our present purposes). I thus assume that there is a (low) verbal functional projection associated with assignment of θ, θP.

The assumption that structural case does not require θ now comes out as (99): the θ projection is empty and the relevant relation is thus between TP and KP in the noun:

(99)

```
              TP
            /    \
          θP      VP
         /  \
        ∅    KP
            /  \
           K    DP
                /\
               N
```
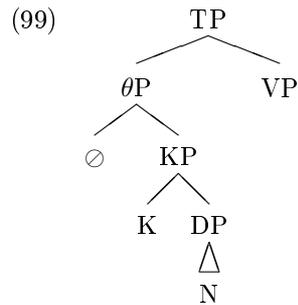
Assuming that empty projections are absent:

(100)

```
              TP
            /    \
          KP      VP
         /  \
       K_s   DP
             /\
            N
```
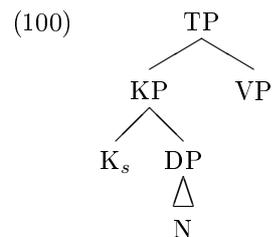
Inherent case, on the other hand, involves a θ and thus a relation between the verbal and nominal θ projections: (the 'n' versus 'v' subscripts are not significant, they simply help to disambiguate functional projections belonging to the verb from those belonging to the noun)

(101)

```
              TP
            /    \
          T      θ_vP
                /    \
              θ_nP    VP
             /   \
            θ     KP
                 /  \
                K    DP
                     /\
                    N
```
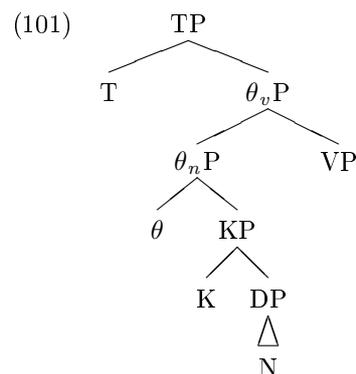
As a consequence, only (101) triggers a θ-across-θ configuration.

Pied-Piping.   The second point we need to understand is why overt Q-movement triggers θ-movement while covert movement doesn't. The obvious observation here is that the Q-morpheme is a bound morpheme in all languages under discussion: 'wh-' in English, 'k-' in Czech, French, Italian and Slovak.

Supposing that this morpheme corresponds to a projection within the noun-phrase, this projection - or its specifier - cannot move overtly: *wh ... $[_{DP}$ t]. This is because moving it overtly would make spell-out impossible, with one morpheme of the wh-word separated from the rest of the word. Overt wh-movement must thus be pied-piping of the wh-phrase by some constituent containing it (this line of thought, applied to different data-sets and with some technical variation, has been exploited by various authors, among which Watanabe (1992) is an influential example).

The same morphology-based reasoning entails that only full functional sequences can be the target of movement, assuming that Brody (2000b)'s *Mirror Theory* is on the right track. Adapting terminology, in *Mirror Theory* every head-complement sequence maps onto a morphological relation, from which it follows that the end of every morphological word corresponds to the end of a head-complement sequence.

To illustrate, an English clause with a realized complementiser consists of more than one head-complement sequences as in:

(102)
```
              CP
            /    \
         that      FP
                 /    \
               TP      F'
             /   \    /  \
          she    VP  F°   ...
                 |
               sleeps
```
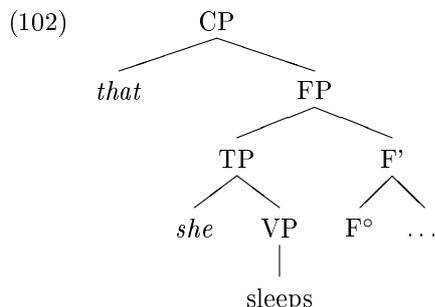
This is because there are two independent morphological words (*that* and *sleeps*) and thus two head-complement lines (with 'that' = C+F and 'sleeps'=T+V). In contrast, in a language where the complementiser is a suffix on the predicate, the underlying syntactic structure consists in a single head-complement line as in traditional representations.

Applying this line of thought to the problem at hand, it follows that pied-piping Q in (98) cannot result in overt movement of QP, since QP is a subpart of a morphological unit and moving it would result in two unpronounceable pieces. Similarly, pied-piping Q in (98) cannot result in KP-movement (i.e. A-movement as detailed below) because that would again result in the splitting of one morphological unit into two non-adjacent parts: θP and KP. It is thus a consequence of *Mirror Theory* that only the full functional sequence headed by θP can move

away (see also Wilder (1994), Wilder (1997) for a similar restriction, discussing coordination).

Overt wh-movement is thus movement of a θ-phrase in order to pied-pipe the Q-feature. Given this result, what strong islands reveal is a constraint on pied-piping of the form:

(103)   if α-movement involves pied-piping of α by γ, then locality restrictions of both α and γ must be respected

In other words, pied-piping is two movements collapsed in one (movement of the target feature, here Q, and movement of the pied-piped syntagm, here θ); and as such it is subject to dual locality restrictions.

Because of (103), every θ-marked syntagm is a trap: trying to extract a θ-bearer from inside a θ-marked syntagm will cause a θ-across-θ violation. Similarly, every θP functional projection is a trap: trying to extract a θ-bearer from below it will also create a θ-across-θ violation. Conversely, every wh-movement of a θ-bearing phrase must make an intermediate step in θP, in order to satisfy the pied-piping constraint. This entailment is what underlies §4.

Some important consequences of this line of thought include:

[i] A side-effect of (103) is that the difference between (99) and (100) becomes significant. If (100) is chosen, pied-piping of Q involves K-movement (A-movement) rather than θ-movement. This is not significant here - since A-movement is severely restricted and is irrelevant wrt. strong islands - but the asymmetry between triggering A-movement and θ-movement will underly §6.

[ii] Subextraction - such as the French *combien* split illustrated in e.g. (89a) above - cannot be a case of pure-Q movement. This is entailed both by the locality and the morphology of *combien*. Morphologically, the wh-morpheme is a proper subset of the subextracted word; moving the whole 'subextracted' word thus involves pied-piping. Locality-wise, *combien* is sensitive to strong islands and must thus involve pied-piping (of θ).

Analyzing a subextracted *combien* as pied-piping of θ is however awkward in the light of (85): if nominal modifiers such as *with many holes* do not receive any θ-role, quantifiers such as *many* shouldn't receive any either.

A more plausible line of analysis of wh 'subextractions' is that they are not subextraction but rather remnant movement, with derivations similar to those intensively explored by recent work following Kaynian guidelines. In such a derivation, the complement of *combien* makes a local movement to the left of *combien* and the θP now containing [*combien* t] is wh-moved. The existence of such structures is overtly shown by Russian examples in which wh-movement of a PP strands the noun, taking out the preposition and the wh-phrase:

(104)   skazi na kotorei ti   storone
        say   on which you side

'tell us on which side you are'

That such structures are relevant to French 'subextraction' is overtly suggested by the following data:

(105) a. tu   as   besoin de combien   de photos?
           you have need   of how-many of pictures
           'how many pictures do you need?'

       b. de combien   (est-ce que) tu   as   besoin de photos?
           of how-many (is-it   that) you have need   of pictures
           'how many pictures do you need?'

(106) a. tu   as   parlé   à combien   de photographes?
           you have spoken to how-many of photographers
           'how many photographers have you spoken to?'

       b. à combien   (est-ce que) tu   as   parlé   de photographes?
           to how-many (is-it   that) you have spoken of photographers
           'how many photographers have you spoken to?'

Developing this line of analysis of quantified noun-phrases (and prepositional phrases) would lead us far from locality concerns, so let us simply note that under such an analysis, 'subextractions' involve exactly the same pied-piping by θP as canonical wh-movement.[11]

---

[11] One difference between a remnant-extracted *combien* and movement of the full nominal phrase is agreement: *combien*-movement cannot trigger agreement, while movement of the full phrase can. This contrast surfaces both in past-participle agreement with a moved wh, and in *that-t* context, where *combien* is unable to trigger the *que-qui* alternation: (I thank Luigi Rizzi for pointing out the relevance of these facts)

(1) a.   combien   de voitures a-t-il   conduites ?
            how-many of cars   has-he driven
     b.   combien a-t-il conduit de voitures?
     c. * combien a-t-il conduites de voitures?

(2) a. ??? combien   de gens   est-ce que tu   crois qui viendront?
             how-many of people is it   that you think *qui* will-come
             how many people do you think will come?
     b.   combien est-ce que tu crois que [- de gens] viendront?
     c. * combien est-ce que tu crois qui [- de gens] viendront?

This difference is *a priori* unexpected if *combien*-extraction is essentially similar to movement of the full DP.

There is however one important difference between extraction of the whole nominal and *combien*-extraction: only *combien*-extraction involves noun-phrase internal movement of the complement: *[combien de livres]* -> *[[de livres] combien t]*. This in turn gives a solution to the agreement difference: recall that (i) agreement involves A-movement and A-movement involves the smaller KP structures (100), (ii) KP cannot subextract from θP because K and θ form a single morphological unit. Supposing that the noun-internal movement lands in θP, the following derivation is forced:

[iii] Languages which do express the wh feature as a free morpheme will be able to move it across strong islands. This is because movement of the Q-projection alone is now morphologically legal, and will be similar to the locality of French situ-wh. This suggests an intriguing alternative analysis of the generalization put forward by Cheng (1997) to the effect that languages either fill C° with a wh-particle or move a wh-phrase to specCP and of the fact that in these languages wh-phrases are also indefinite pronouns (along the lines of e.g. Cole and Hermon (1998) among others). These two facts would come together if the so-called C° particle is simply the Q projection of the wh-phrase having (overtly) moved to the complementiser projection, stranding the rest of the DP/QP: this derives the fact that the moved element looks like a particle, since it is a pure expression of wh, and also explains why the remnant looks like an indefinite pronoun, if the frequent suggestion is correct that a wh-phrase *is* the union of wh and an indefinite pronoun.

The Feature Tree. Finally, we need to know where θ lies in the syntactic feature-geometry. We know that θ does not block Q-movement because LF-movement of Q can skip θ. θ can thus not be below Q. Similarly, we know that θ-movement is not blocked by intervening A-positions, since overt wh-movement is also θ-movement. θ thus cannot be above A/ϕ/sCase. The only possibilities are for it to be either a sister to Q and A, or a daughter of A. It cannot however be a daughter of A/ϕ/sCase as this would mean that θ-roles are a subset of ϕ (or structural case) features, which is wrong. We are thus led to the following (non-final) syntactic feature tree:

- first the complement of the noun lands in specθP.

- This is however not sufficient, as the KP complement of θP cannot remnant-move due to the morphological restriction discussed above. The complement of the noun must thus move further, out of specθP and into a noun-external position.

- Second, the full θP of the form *[t combien t]* undergoes a remnant wh-movement.

(The fact that movement of the complement goes through specθP is probably an instance of the more general fact that extraction from noun-phrases proceeds through that 'escape-hatch'.)

It thus follows that remnant-*combien* is movement of a θP while simple movement of the agreeing full noun-phrase is KP-movement. The agreement puzzle thus translates into the fact that θP structures do not trigger agreement (but KP structures do).

This is however a well known fact, once it is realised that θP structures correspond to inherent case: it is well known that inherent-cased DPs do not trigger agreement (only structurally-cased DPs do). Essentially, this approach reduces the *combien*-agreement contrast to the more general and well-known contrast in agreement between structural and inherent case.

(Of course, the more general fact that KP provokes agreement while θP doesn't remains to be understood. Notice however that the present approach gives a direct solution for the related fact that agreeing noun-phrases block subextraction: if agreement forces KP and θP is a necessary landing site for subextraction-structures, subextraction will be mutually exclusive with agreement.)

(107)

Quantifier θ A [ϕ/sCase]
    |
SpecificQ
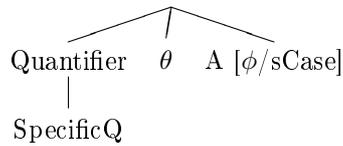
(107) together with the constraint on pied-piping derives the generalisations seen up to now:

- overt movement is Q+θ movement (in order to preserve morphological integrity). As such it is subject to both Q and θ locality, in virtue of the pied-piping constraint. This movement is thus blocked by intervening Q elements (the classical *Relativised Minimality* effect),

- this block can be circumvented if the quantifier is 'specific', making it a SQ+θ movement (the eWI and presuppositionality effect)

- if the intervener is an SQ itself, the movement is entirely killed (the 'definiteness'-island effect)

- if a θ intervenes, the movement is similarly impossible, as there is no subclass of θ to circumvent the ban (the 'strong island' effect)

- if the movement is covert, the pied-piping constraint does not apply and θ will not be taken along. The movement will then be a pure Q-movement, with the same locality effects summarized above except for insensitivity to θ-interveners, i.e. insensitivity to strong islands (the situ-wh effect)

- finally, outcasts - who are deprived of θ - cannot move at all, as the only candidate for pied-piping lacks the property that would allow it to move.

3.4 Modifier-movement and θ-movement

Our (re)definition of θ mostly coincides with Rizzi's characterization of 'modifiers', and it is thus tempting to view the two as the same movement. Rizzi (2001) adds 'modifiers' to 'quantifiers' and 'arguments' as a class relevant to *Relativised Minimality*, in order to capture anteposition facts such as those noted by Koster (1978) for Germanic, Starke (1993) for Slavic, and Cinque (1999) for Romance:

(108) a.   Het is zo dat hij helaas   waarschijnlijk ziek is
            it   is so that he unfortunately probably   sick is

       b. * Het is zo dat hij waarschijnlijk helaas   ziek is
            it   is so that he probably   unfortunately sick is

(109) a.   Helaas   is hij t waarschijnlijk ziek
              unfortunately is he   probably   sick

        b.   * Waarschijnlijk is hij helaas   t ziek
              probably   is he unfortunately sick

        c.   Waarschijnlijk is hij t ziek
              probably   is he   sick

(110) a.   On ho urcite   skoro   rozbil
              he it certainly almost broke

        b.   * On ho skoro   urcite   rozbil
              he it almost certainly broke

(111) a.   Urcite   ho skoro   rozbil
              certainly it almost broke
              he certainly almost broke it

        b.   * Skoro ho urcite   rozbil
              almost it certainly broke

        c.   * Skoro ho rozbil
              almost it broke
              he almost broke it

(112) a.   I   tecnici   hanno probabilmente risolto   rapidamente il
              the technicians have   probably   resolved quickly   the
              problema
              problem

        b.   * I   tecnici   hanno rapidamente risolto   probabilmente il
              the technicians have   quickly   resolved probably   the
              problema
              problem

(113) a.   Probabilmente, i   tecnici   hanno risolto   rapidamente il
              probably,   the technicians have   resolved quickly   the
              problema
              problem

        b.   * Rapidamente, I   tecnici   hanno probabilmente risolto   il
              quickly,   the technicians have   probably   resolved the
              problema
              problem

        c.   Rapidamente, I   tecnici   hanno risolto   il   problema
              quickly,   the technicians have   resolved the problem

Essentially, only the highest adverb can be anteposed. Preverbal subjects and other arguments are excluded from the race (a fact which is not entirely clear in Slovak). Notice that there is no need to exclude the subject from the set of modifiers in order to obtain this result: preverbal subjects in an A-position are invisible qua interveners, under the traditional assumption that only interveners 'checking' the relevant kind of property are relevant interveners. In such a characterization, modifier-movement and θ-movement appear essentially identical.

But the two cannot be collapsed: intervening modifiers do not block wh-movement (unless quantificational, of course).

(114)   how do you <u>usually</u> sleep?

A more plausible relationship between the two is that θ-roles, as understood here, are a subset of Rizzi's modifiers. This is what the actual features are: i.e. predicate-modifiers are a subset of all modifiers occurring in the clause. The grammaticality of (114) also follows: θ being a subclass of M, it can move over M by the same logic that allows SQ to move over Q.

(Identifying θ with the 'predicate-modifier subset of M' sheds some light on the generalization discussed in fn.6 above: adverbial modifiers block wh-movement. Adverbial modifiers being precisely θ-entities, this is now a further θ-across-θ effect. In fact, the contrast between those modifiers who don't block, (114), and those who do is a further reason to view θ as a subset of M.)

Rizzi's M is thus best integrated into the syntactic feature tree as the mother of θ:

(115)
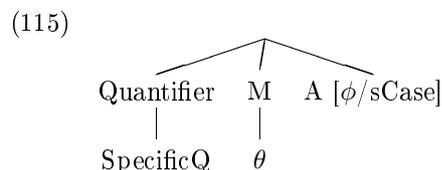
3.5   Dangling Issues

The remaining issue to address concerns the nature of the range-based route out of weak islands, and associated phenomena. Before turning to that, let us however clean up: in order to concentrate on the core logic, various points have been left open or set aside above, on which a word needs to be said.

First and foremost is the nature of the degradation in extractions from within subjects (and object-shifted datives); more generally, the degradation in what used be characterized as 'extraction from non-final noun-phrases' in earlier generative literature. A second issue left on the drawing board concerns a special restriction on coordinated situ-wh in French, and some variations of judgments reported about situ-wh.

3.5.1 Situ-wh, Coordination and Registers

The French situ-wh paradigm was characterized above as immune to strong islands. There is one case however where the situ-wh is blocked (setting aside specific DPs, discussed above, which do not belong to the SI paradigm, given the present typology). This is the first conjunct of a coordination: the situ-wh survives within the second conjunct, but not in the first:

(116) a.   tu crois qu'elle a   acheté des skis et   quoi d'autre?
            you think that she has bought skis   and what else?

       b. * tu   crois qu'elle a   acheté quoi (d'autre) et   des skis?
            you think that she has bought what (else)   and skis?

The same is true in sentential coordination:

(117) a.   tu crois que Marie a été acheter des carottes et que Jean a été
            faire quoi?

            you think M went to buy carrots and that J went to do what?

       b. * tu crois que Jean a été faire quoi et que Marie a été acheter des
            carottes?

            you think J went to do what and that M went to buy carrots?

And in multiple-wh constructions:

(118) a.   ? qui veut inviter   Jean et   qui d'autre?
            who want to-invite J   and who else?

       b. * qui veut inviter   qui et   Jean?
            who want to-invite who and J?

The underlying issue here is the status of coordination as a strong island. Characterizing strong islands as θ-bearers - within the extended definition of θ - is plausible in all cases of SI except coordination. But coordination is nevertheless standardly taken to be a strong island.

Recall however the methodological point above: transparent nodes are those nodes that can be transparent in principle, not nodes which are always transparent. We know independently that coordination *is* transparent in principle: this fact is what is descriptively called 'Across The Board' extraction.

(119)   who did Mary grade with an A and John with a B?

Instead of characterizing coordination as SI and then granting an exception to ATB cases, a more rational step is to approach the issue in the reverse: coordination is a transparent node, but some additional constraint forces ATB on it.

As has been repeatedly noted, there is such an additional constraint, dubbed 'parallelism', which constrains coordinations well beyond extraction facts (cf. e.g. Fiengo and May (1994), Fox (1995)).[12]

Given this characterization, what the above coordination asymmetry tells us - descriptively - is that at LF parallelism goes left to right. That is, whatever happens in the first conjunct must also happen in the second conjunct, but deeds happening in the second conjunct need not be reflected in the first.

The reason for the ungrammaticality of a situ-wh in the first conjunct is then traced down to the fact that if the first conjunct contains a variable bound by a wh-operator, the second must have one too. Since the converse is not true, a situ-wh can travel from the second conjunct without affecting the first.

Of course, the nature of the parallelism constraint on coordination and related constructions is mysterious, more so given the generalization that directionality at LF is a subset of that found at PF, but this is not the place to address the syntax of coordination insofar as it is orthogonal to locality (see Wilder (1996), Wilder (1997) for an enlightening discussion of the right-left versus left-right character of operations within conjunctions, deletion in his case).

The coordination asymmetry therefore does not alter the generalization that situ-wh disregard strong islands.

The Classical Description

The description of situ-wh given in some previous literature on French is significantly more restrictive than the paradigms described above. Constructions such as (120) are reported as ungrammatical, on a par with (121) and in contrast with (122).

(120)   a.   tu crois que Jean a   parlé   à qui?
             you think that John has spoken to whom

        b.   elle pensait qu'il   habitait où?
             she believed that he lived   where

(121)   a.   * je me demande Jean a   parlé   à qui
               I self wonder   John has spoken to whom

        b.   * elle ne sais   pas il habitait où
               she not knows not he lived   where

(122)   a.   Jean a   parlé   à qui?
             John has spoken to whom

---

[12]Another illustration of the potential transparency of coordination is given by the (rare) cases of (overt) extraction from within a single conjunct, which for unclear reasons escape the parallelism requirement. These are all from the *first* conjunct, to my knowledge:

(i) ? what did you eat in Taiwan as well as roasted Talama in China?

      b.   il   habitait où?
             he lived   where

Such judgments - relied on in e.g. Bošković (1997), Boeckx (2000) - are in sharp contrast with spoken French, where (120) is not only grammatical, but the unmarked way of asking such questions. (My best guess is that the more restrictive judgments correspond to 'classical' written French, as opposed to the spoken French described here. I have however been unable to find a single speaker of that dialect).[13]

A second difference concerns interpretation. Chang (1997) (as cited by Boeckx (2000)), Boeckx (2000), and related work describe French situ-wh as always requiring the rising pitch, and as always presuppositional. This does not hold in the spoken French described here.

A radical illustration of this, is that contexts which disallow presuppositional objects allow situ-wh. Utterances similar to the following are for instance frequent occurrences in informal speech:

(123)   a.   tu   fous quoi?
             you 'do' what
             what are you doing?

      b.   t'as foutu quoi pendant tout ce   temps?
             you 'did' what during   all   this time
             what did you spend all this time doing?

      c.   tu   crois qu'ils   vont foutre quoi, cet après-midi?
             you think that'they will 'do'   what this afternoon
             what do you think that they will do, this afternoon?

As discussed in §7.7, the object of *foutre* is restricted to a small set of items carrying no presuppositions (and to some idiomatic meanings). It thus cannot be the case that situ-wh requires any form of presuppositional semantics (which concords with speakers' interpretive intuitions).

Another illustration of this difference is given by possible answers to situ-wh questions. It is for instance reported that (124b) cannot be felicitously answered by 'nobody', whereas (124a) can. No such contrast exists among the speakers consulted (or for me).

(124)   a.   qui (est-ce que) tu crois qu'elle veut inviter?

---

[13]Furthermore, preliminary investigation suggests that regional variation does not play a determining role: speakers from Belgium and south-west France were found to share the (Geneva-based) judgments given in the text, modulo the noted (irrelevant) variation as to eWI acceptability.

(See also Hirschbühler and Labelle (2000) for what might be a similar discrepancy, in a different French construction.)

> who do you think that she wants to invite
>
> b.   tu crois qu'elle veut inviter qui?

The situation thus seems to be that the purported 'classical French' allows only a proper subset of what spoken French does, limiting situ-wh to single-clause movement (with obvious similarities to the locality of QR). More data is needed to draw the right lines, but I will tentatively assume that this register/dialect is in fact a proper subset of spoken French wrt situ-wh, and parametrisable along familiar lines.[14]

3.5.2   Levels of Degradations In Extractions from DPs

As it stands, the model predicts that extraction out of subjects should be equivalent to extraction out of objects (both DPs and CPs). This is not accurate, except maybe for a small set of contexts.

In order to trace the source of these degradations, some care needs to be put into the cutting of the cake: there are in fact several distinct effects here, and although the boundaries remain murky, some general lines of analyses suggest themselves, once the phenomena are described separately.

DP objects. First, there is the fact that extracting out of an object noun-phrase is different from extracting out of a (bridge) CP: slightly degraded, and with different presuppositions.

(125)   a.   qu'est-ce que tu crois que Marie veut dessiner?

              what do you think that Mary wants to draw?

       b.   de quoi est-ce que crois que Marie veut acheter un dessin?

              what do you think that Mary wants to buy a drawing of?

As discussed above, (125a) can be used in a context where the speaker attributes no belief at all to the interlocutor, a 'request for guess' context. Maybe Mary is walking around with a drawing-set, and the speaker is trying to guess what she will draw. In (125b) on the other hand, a richer - presuppositional - context is

---

[14]For completeness, it should be mentionned that situ-wh also has a subject/object asymmetry; more precisely a derived-position versus base-position asymmetry:

(1)   a.   tu crois qu'il s'est passé quoi?

             you think that 'it self' is happened what = 'what do you think happened?'

       b.   ??tu crois que quoi s'est passé?

As with other phenomena of the *that-t* family, I will leave this aside here, hoping to address it separately. (Notice however that the ungrammaticality is much less sharp here than in standard that-t structures.)

needed. It is difficult to judge whether the context needs to be 'range'-rich (i.e. a set of alternative referents must be pre-established, without the presupposition that there exist one specific referent such that the interlocutor believes that this referent is the relevant one) or whether a specific referent is required (i.e. the presupposition that the interlocutor has some entity or other in mind, and the question requests the identity of that something). Though delicate, judgments seem to lean towards specificity.

This is also suggested by the fact that the demands on contexts in (125b) appear to be essentially equivalent if a weak island is introduced (the latter being an SQ context, as discussed above):

(126)   de quoi est-ce que crois que Marie ne veut pas acheter de dessin?

          what do you think that M doesn't want to buy any drawing of?

Extractions out of wh-phrases make the same point:

(127) a.   de qui est-ce qu'il n'est pas clair [combien de portraits <de qui>] Amélie a dessiné?
       b.   who is it unclear [how many portraits of <who>] Amélie drew?

Although the extracted wh crosses another wh within the noun-phrase, the judgment is essentially identical to parallel extractions out of non-wh noun-phrases (see section on mild degradations below). Indicating once again that extracting out of noun-phrases always involves jumping over Q.

Let us then adopt the description that extraction from object DPs involves SQ-movement while extraction from object CPs involves Q-movement.

This follows from the traditional assumption that noun phrases always contain a quantifier (Thomason (1974)). In current terms, the D-node - if present - involves quantification, either directly or by association with QP. This forces the extractee to escape by the SQ road, and hence entails specificity. (Notice that the conclusion as to the Q-status of noun phrases only holds for those noun-phrases from which material can be extracted; proper-nouns and related light nominals are exempt.)

If this is so, non-specific noun-phrases are *de facto* weak islands, and the degradation of (125b) is comparable to an object eWI, which seems correct. This is particularly clear with less liberal speakers, who tend to grade eWI as '??' and give the same judgment for DP-extraction.

An interesting confirmation that DPs are weak islands comes from extraction of *combien* from inside the complement of a noun.[15] As a baseline, recall the behaviour of extraction of *de combien*:

---

[15]I owe this argument to Luigi Rizzi.

(128) a. ? de combien   est-ce que   tu   crois qu'il faudra
           of how-many is it   that you think that it   will-be-necessary
           parler   parler [_ _ de   livres]] ?
           to-speak   of   books
           how many books did you write the preface of?

    b. * de combien   ne   sais   tu pas   <u>comment</u> parler
         of how-many   have-you written the preface   of
         [_ _   de livres]] ?
         books
         how many books did you write the preface of?

The contrast (128) illustrates the classical sensitivity of *(de) combien* extraction to weak islands. This contrast is exactly reproduced in noun-phrases, where bare (determiner-less) noun phrases are transparent on a par with (128a), while DPs create an island effect on a par with (128b):

(129) a. ? de combien   as-tu   [besoin [_ _ de   livres]] ?
           of how-many have-you a-need   of   books
           how many books do you need?

    b. * de combien   as-tu   écrit   [<u>la</u> préface [_ _ de livres]] ?
         of how-many have-you written the preface   of books
         how many books did you write the preface of?

This directly follows if the determiner is a Q(uantifier) inducing a weak island effect.

Opaque Subjects.   Given this background, let us turn to contrasts between extraction out of objects and non-objects. Again, care needs to be taken to distinguish two different effects, corresponding to two levels of degradation: sharply ungrammatical extractions of the SI type versus light degradation.

Preverbal clausal subjects are typical examples of severe degradation; they contrast with non-clausal subjects in producing near-unparsable sentences:

(130) a. * quel film est-ce que tu crois que [projeter t] peut créer un scandale?
           which film do you think that to project can create a scandal

    b. (?) de quel film est-ce que tu crois que la première partie peut créer un scandale?
         which film do you think that the first part of can create a scandal?

As traditionally observed, this effect only holds for the derived position. Wherever the 'subject' can remain in the base without being focalised, it does not

trigger any degradation (unfortunately, this can only be illustrated with ergative predicates in the languages under discussion and thus involves an object-like position):[16]

(131) a. quel film est-ce que tu crois qu'il est impossible de [projeter t]?
           which film do you think that it is impossible to project
       b. * quel film est-ce que tu crois que [projeter t] est impossible?
           which film do you think that to project is impossible

Transforming the subject into a noun-phrase takes away the force of the ungrammaticality, compare (130a) with:

(132)   ?? de quel film est-ce qu'une projection peut créer un scandale?
           of which film do you think that a projection can create a scandal?

What is the difference between noun-phrases and clauses? Case.

Clauses do not bear case (or agreement), unless nominalized. Accordingly, movement of a clause into a derived subject position cannot be case-driven. I will thus assume that the only way for a clause to climb up, is for it to bear θ and move to a theta-related position (this is reminiscent of Cardinaletti (1997)'s 'SubjectOfPredication' position, and predication-driven subject-movement of Zwart (2000); it is also a version of the traditional claim that clausal subjects are in different positions than noun-phrase subjects). Conversely, the only possibility for a clause not to bear θ is for it to remain in its base position.

If this is so, clauses in derived positions are simply strong islands, qua θ-bearers.

A second class of cases where structural-cased syntagms become strong islands are preverbal subjects in Czech/Slovak (data by Lida Veselovska):

(133) a. kolik   myslis   ze   prislo dopisu?
           how-many you-think that came letters
       b. * kolik   myslis   ze   dopisu prislo?
           how-many you-think that letters came

Both the preverbal and postverbal position are in principle available for subjects, but leaving the remnant of a subextraction in a preverbal position makes the sentences unacceptable.

The same paradigm obtains with extraction of possessives. The latter are less well accepted and seem to have more speaker-variation, but the pattern is stable: extracting out of a postverbal subject is equivalent to extracting out of an object, while extracting out of a preverbal subject degrades sharply.

---

[16] The observation that sentential subjects do not trigger strong island effects in SOV languages (Grosu (1972), Cinque (1977)) is then plausibly traced down to the fact that the verb remains low in such languages and thus the 'preverbal' sentential subject can itself occupy a very low position - the base position - which never triggers degradation, as discussed below.

(134) a. ? ci   mysli ze jsi   videl matku?
          whose thinks that you-are seen mother
          whose mother does s/he think that you saw?

    b. ? ci   myslis   ze   prisel dopis?
          whose you-think that came letter
          whose letter do you think came?

    c. * ci   myslis   ze   dopis prisel?
          whose you-think that letter came
          whose letter do you think came?

This contrasts with preverbal extractions in French and Italian which incur only a very mild penalty (discussed below).

Here the model gives us two possibilities: either Czech/Slovak preverbal subjects contain θ, or they contain SQ. In other words, either their nominatives resemble inherent case, or they are specific.

Both are technical renderings of traditional claims: respectively the claim that preverbal subjects of (some) null-subject languages occupy peripheral positions (where they are not structurally licensed and thus need to bear θ) and the claim that (some) null subject languages require the subject to be specific in order to be preverbal.

Although there is some evidence for the latter claim, specificity does not seem to be a promising path to take. The evidence in favor of specificity, differentiating Czech/Slovak from French/Italian, lies in the interpretation of preverbal indefinite subjects: in Czech/Slovak a preverbal indefinite - morphologically a bare noun - unmarkedly receives a definite interpretation, whereas this is not the case in French/Italian. This can hardly be the source of the extraction difference however: it is possible to obtain an indefinite reading on Czech/Slovak preverbal subjects by adding 'some' or related modifiers, but this does not make extraction any better.

I will thus tentatively conclude that Czech/Slovak do indeed force preverbal noun-phrases to bear a θ-role, possibly because the subject position is rigidly occupied by an expletive *pro*.

Castillian Spanish possibly confirms this speculative note on a relationship between the distribution and transparency of subject: it seems to pattern with Czech/Slovak and against French/Italian both on opacity of the subject and on its distribution.

First, preverbal subjects disallow extraction (data from Marti (1999), and Marti p.c.):

(135) a. ? ¿[De qué autor] crees que han recibido [varios libros t]
         of   what author think that have received several books t
         premios internacionales?
         awards international

By what author do you think several books have received international awards?

b. ?* ¿[De qué autor] crees que [varios libros t] han recibido
   of   what author think that several books t have received
   premios internacionales?
   awards international
   By what author do you think several books have received international awards?

Second - informally put - Spanish has a 'liberal' use of the low subject position, on a par with Czech/Slovak but against Italian. (if Ordonez (1998) is correct, this is also true of the low object ('scrambling') position, where Spanish patterns again with Czech/Slovak and against French/Italian). It is thus tempting to relate the θ-nature of preverbal subjects in these languages to the liberal use of low positions. Which is traditionally expressed through the fact that an expletive *pro* rigidly occupies the preverbal subject position.

Be it as it may, opaque subjects are characterizable as subjects which bear a theta-role, both in the case of clausal subjects, and with Czech/Slovak subjects.

Mildly Degraded DP subextraction. The third set of cases are noun phrases which trigger a very mild degradation, if any degradation at all. The strength of the effect varies somewhat: the lower boundary is given by examples such as (130b) above, repeated here as (136b), which show basically no subject/object asymmetry; and the upper boundary by subject *combien* 'subextraction', repeated as (137), with a light degradation:

(136) a. de quel film est-ce que tu crois qu'elle a vu la première partie?
          which film do you think that she saw the first part of?

      b. (?) de quel film est-ce que tu crois que la première partie peut créer un scandale?
          which film do you think that the first part of can create a scandal?

(137) a. combien tu crois qu'elle a critiqué [t de films] ?
          how-many you think that'she has criticized [t of films]

      b. ? combien tu crois que [t de films] sont arrivés?
          how-many you think that [t of films] are arrived

The same light asymmetry however occurs with a situ-wh embedded deeply inside the subject/object:

(138) a. tu crois qu'elle a vu [la première partie de quel film] ?
          you think that she saw the first part of which film?

58

b. (?) tu crois que [la première partie de quel film] peut créer un scandale?

you think that the first part of which film can create a scandal?

And the same light asymmetry in fact occurs with quantified noun-phrases independently of wh altogether:

(139) a. je crois qu'elle a vu [la première partie de 3 films]
         I think that she saw the first part of 3 films

     b. (?) je crois que [la première partie de 3 films] peut créer un scandale
         I think that the first part of 3 film can create a scandal

ECM constructions show a similar very mild asymmetry, entirely absent for many speakers:

(140) a. combien tu as   vu partir de gens?
         how-many you have seen leave of people

     b. combien tu as   vu de gens   partir?
         how-many you have seen of people leave

But again, this is exactly the same independently of wh, with a postverbal (quantified) DP preferred over an ECM-moved counterpart:

(141) a. il a   vu   partir 3 personnes
         he has seen leave 3 persons

     b. il a   vu   3 personnes partir
         he has seen 3 persons   leave

Extraction from a moved wh-phrase shows the same pattern, i.e. no significant difference from object-extraction, cf. (93) repeated here:

(142) a. ? de qui est-ce qu'il est pas clair [combien de photos <de qui>] il veut prendre t?

     b. ? who is it unclear [how many picture of <who>] he wants to shoot t?

This has the same level of acceptability as a parallel eWI extraction, which does not involve DP-subextraction:

(143) a. ? à qui est-ce qu'il est pas clair [combien de photos] il veut donner <à qui> <combien de photos>?

     b. ? to whom is it unclear [how many pictures] he wants to give <to whom> <how many pictures>?

Extracting from within the wh-moved syntagm does not seem to trigger any extra penalty (see Chomsky (1986):25-27, Torrego (1986)).

The conclusion from these paradigms seems to be that extracting from a noun-phrase in derived positions does not trigger any locality violation *per se*. Whatever additional penalty is attested seems to come from independent sources, if it comes at all.

English. The status of English with respect to non-object DP subextraction is unclear, as speakers report contradictory judgments. One set of speakers behaves like French and Italian speakers, seeing no relevant subject/object asymmetry while another set behaves like Czech/Slovak in reporting a 'sharp' degradation in the familiar examples:

(144) a. which film do you think that she saw the first part of?
       b. (*) which film do you think that the first part of might create a scandal?

Interestingly, the two classes of speakers seem to pattern differently across the board (the sample is however too small for any strong conclusion). The English speakers who do not like (144b) also do not like simple eWI, in which they again pattern with Czech/Slovak against French/Italian (eWI is judged more marginal in the former). Similarly, speakers who see no additional problem with (144b) seem to pattern with French in not seeing any additional degradation in extraction from within a moved wh-phrase.

Having been unable to find more systematicity among English judgments, I will leave the question open here. (But see Hiramatsu (2000) for experimental evidence indicating that English extraction out of subject noun phrases generally patterns with object extraction and against strong islands, i.e. English speakers behave like Romance, not like Slavic.)

To summarize, the two additional complications induced by extraction from within noun-phrases are (i) noun-phrases contain a quantifier and thus wh-movement out of them must be SQ-movement, and (ii) some seemingly structurally case-marked noun-phrases bear θ and thus pattern with strong islands.

3.6 The Tensed-Complement Effect

A traditional mystery about wh-extraction is that some languages are reported to entirely disallow extractions out of indicative complements. In some German regiolects, not only is (145a) ungrammatical, but it has "no difference" with (145b).

(145) a. * was   glaubst du dass Hans kaufen will?
              what believe you that Hans buy   will

b.   * wie ist unklar   ob Hans sich benehmen wird
      how is unclear if   Hans self behave   will

This restriction seems to affect speakers from the central zone of Germany, in a band spanning Köln-Bielefeld-Berlin. Southern speakers (Southern German, Austrian German and Swiss German) and northern speakers (Hamburg, etc.) do not share these judgments. The southern varieties for instance display the same judgments as the French judgments with respect to run-of-the-mill eWI: (145a) incurs the usual penalty of argumental eWI and contrasts with (145b) which is ungrammatical.[17]

It is sometimes claimed that the ban is not general, capitalizing on the fact that these extractions become better in what looks like embedded V2, in the relevant regiolects (again, no argument/adjunct asymmetry arises):

(146)   a.   ? was   glaubst du   will Hans kaufen ?
              what believe you will Hans buy

         b.   ? wie glaubst du   hat Hans geschlafen?
              how think   you has Hans slept

Testing the relevant prediction however suggests that this is not an effect of embedded V2, but rather a parenthetical.

First, the effect is not recursive. Inserting a second level of embeddeding seriously deteriorates the structures:

(147)   ?* was   glaubst du   sagte Maria will Hans kaufen?
            what believe you said   Mary will Hans buy

Since the sequence 'believe you said Mary' is not a legal parenthetical but is a legal sequence of embedded clauses, this suggests that parentheticals are what is at work here.

Second, binding acts as if the purported 'root clause' does not bind into the 'embedded clause'. Of course, if the intervening sequence were parenthetical, we would expect lack of c-command into the lower clause. If it was a sequence of embedding, presence of c-command would be expected.

(148)   a. ?* was   glaubt jede Frau$_i$ will sie$_i$ kaufen?
             what believes every woman will she buy

         b.   ? was   glaubt   er$_i$ will Hans$_i$ kaufen?
              what believes he will Hans buy

---

[17]Central German data from Klaus Abels and Daniel Buring. Northern German from Cornelia Hamann and Renate Musan (via Gisbert Fanselow). Austrian from Winnie Lechner, Swiss-German from Eric Haeberli, Thomas Leu and Gabriel Lobos, Southern German from Uli Sauerland.

In (148a), the quantifier fails to c-command the pronoun of the lower clause, causing a contrast with (146a). In (148b) the sentence degrades, but not to the extent that principle C violation would degrade it.

It thus genuinely seems the case that Northern German indicative complements block wh-extraction. The present framework suggests that the difference between Northern and Southern German lies in case-marking: northern varieties assign an inherent case to their clausal complements while southern varieties pattern with English, French, etc. in assigning a structural case to the clausal complement. As a consequence, the northern clausal complement contains a θ, while the southern clausal complement doesn't need to. In turn, this entails that the Northern German clausal complements will act like strong islands, blocking all overt wh-movement (since the latter involves pied-piping of θ).

The distinction between Northern and Southern German is then equivalent to the distinction between extracting out of a dative and accusative noun-phrase, as illustrated in (95-96).

4 θ-movement as 'Long Scrambling'

We now seem to sit in a comfortable position, so let us profit from the view before the next storm. We've surveyed the whereabouts of two apparent bugs in the language engine - the fact that some questions have a different locality than 'normal' run-of-the-mill questions - and the fact that these questions systematically surface with a different 'meaning' than normal questions.

But that turns out to be exactly what would happen if the language organ was elegant and used only one locality principle for all types of questions, namely: '* α_i ... α_j ... α_i'. We can thus hoover two bugs in one go, while also making some progress on the theoretical front: a quantifier can only afford to skip intervening quantifiers on its path if it can claim to belong to a different class than the intervener. The only way to do this while still remaining a quantifier is to belong to a subclass of quantifiers to which the intervener does not belong. In that case *Relativised Minimality* will let it walk through undisturbed. Belonging to a subclass however has the cost of affecting meaning, and hence this trick will loosen locality but will change interpretation (making it become 'specific').

This line of thought allows us to derive eight locality effects - traditionally described as unrelated - from a single locality principle. First, three effects of Q-movement:

- the traditional *Relativised Minimality* facts (Q-across-Q violations),
- successful extractions out of weak islands (SQ-across-Q),
- and the 'specific DP'-island (SQ-across-SQ violations).

Bringing θ-roles into the mix - through morphological integrity and pied-piping - allowed us to derive five other effects from the same locality principle:

- strong islands result from the fact that θ-bearers cannot cross θ-bearers,
- the fact that outcasts cannot move results from the fact that they do not belong to any reachable class in the feature tree,
- the tensed-complement effect is a θ-across-θ effect
- the clausal-subject island is another θ-across-θ effect
- finally the liberal locality of wh-in-situ stems from the fact that they do Q- but not θ-movement

The two major assumptions required for these results were: (i) syntactic features are organised into a feature-tree, (ii) overt wh-movement involves pied-piping of Q by θ (θP).

Two Routes A typical extraction out of a weak island discussed in §2 now looks like this:

(149) [**what**$_{SQ}$ do you wonder [$_{CP}$ <u>how</u>$_Q$ John cooked <what>?

Due to the pied-piping constraint, both instances of wh-movement involve a thematic step:

(150) **what**$_{SQ}$ ... V [$_{CP}$ <u>how</u>$_Q$ ... [$_{\theta P}$ **t**$_{what}$ [$_{\theta P}$ t$_{how}$ ... [$_{VP}$ ... **t**$_{what}$ t$_{how}$

In such structures, SQ-movement of *what* proceeds in a single long step from the low thematic position in the embedded clause to the left periphery of the root clause. This is a good fit with approaches such as Cinque (1990), Rizzi (1990) where the extracted wh-phrase simply jumps over the island. Indeed, the above can be understood as a refinement of these approaches, purifying them of special assumptions about eWI ('binding chains').

Given the availability of θ-movement, nothing prevents an alternative derivation, involving successive cyclic θ-movement:

(151) **what**$_Q$ ... [$_{\theta P}$ **t**$_{what}$ ... V [$_{CP}$ <u>how</u>$_Q$ ... [$_{\theta P}$ **t**$_{what}$ [$_{\theta P}$ t$_{how}$ ...

If no θ-element intervenes *Relativised Minimality* predicts successive cyclic θ-movement to be legal (see fn.6 for cases where a θ-element does intervene). The movement from the higher θ-position to the root CP can involve simple Q-movement, since no Q-island intervenes.

This derivation amounts to a 'long-scrambling' step feeding wh-movement, the long-scrambling step being similar - or identical - to the long-scrambling type of movement studied by Mahajan (1990).

I will argue here that both (150) and (151) exist, and that they correspond to the two routes out of weak islands suggested in §2.2: specificity-based and range-based respectively.

The fact that the SQ-route (150) exists is clear from §2, particularly the situ-wh paradigm and the 'definiteness' island paradigm. That an alternative route such as (151), also exists is suggested by several paradigms discussed below: reconstruction, *tout*-movement, and the argument/adjunct asymmetry.

4.1   Reconstruction

Simple reconstruction effects in eWI - noticed by Frampton (1991) - provide evidence for an intermediate landing site in eWI-movement, closely corresponding to the higher θ-position in (151):

(152)   a.   [which picture of **herself**$_i$] does Simone$_i$ wonder <u>whether</u> Humphrey
        will dare show to his parents <which picture of herself$_i$>?

        b.   [which picture of **herself**$_i$] does Simone$_i$ wonder <u>when</u> Humphrey
        will finally throw away <which picture of herself$_i$>?

In order for it to be bound by *Simone*, the boldfaced anaphor must be c-
commanded by *Simone*. The only plausible means to achieve this configuration
is to reconstruct the entire wh-phrase *which picture of herself* under *Simone*.

Reconstructing the wh-phrase into its base-position places the anaphor too far
from its antecedent, as illustrated by:

(153)   a.   * which girl thought that Humphrey liked which picture of herself?
             * which girl wondered whether Humphrey liked which picture of
             herself?

        b.   de quelle photo d'elle-même est-ce qu'elle croit qu'il est fier?
             of which picture of herself does she believe that he is proud?

        c.   * Simone croit que Humphrey est fier de quelle photo d'elle-même?
             S believes that H is proud of which picture of herself?

        d.   * quelle fille croit qu'il est fier de quelle photo d'elle-même?
             which girl believes that he is proud of which picture of herself?

It must then be the case that eWI-movement is not a one-step movement, but
rather involves an intermediate step into which the wh-phrase in (152) recon-
structs: below the root subject, but above the embedded subject. An obvious
hypothesis is that the wh-phrase lands in or around the intermediate specCP
(Frampton (1991), Cresti (1995), Rullman (1995), Starke (1999)):

(154)   which portrait of herself does **Simone**$_i$ wonder [$_{CP}$ <which portrait of
        herself$_i$> [$_{FP}$ <u>how</u> Humphrey plans to hang?

This solution however clashes with the facts about the scope of presupposi-
tionality discussed in §2.1: eWI necessarily triggers wide-scope of the extracted
wh-phrase over the root predicate. If (154) was correct, a weak island could be
escaped by SQ-movement within the embedded clause with no scope over the
root clause.

The structure (151) on the other hand resolves the issue: if the reconstruction
paradigms involve a step of long-scrambling (successive cyclic θ-movement), the
root θ-position is adequate both for reconstruction and for scope.

Strictly speaking, what this argument tells us is that there exists an interme-
diate position between the root subject and the root predicate available for
reconstruction of the wh-phrase. It however doesn't tell us what type of posi-
tion this is. One option is (151) - repeated below - but another open option is
that the intermediate site is an SQ-related position. The two possibilities are
thus:

(155)   a. **what**$_Q$ ... [$_{\theta P}$ **t**$_{what}$ ... V [$_{CP}$ <u>how</u>$_Q$ ... [$_{\theta P}$ **t**$_{what}$ [$_{\theta P}$ t$_{how}$ ...
   b. **what**$_Q$ ... [$_{SqP}$ **t**$_{what}$ ... V [$_{CP}$ <u>how</u>$_Q$ ... [$_{\theta P}$ **t**$_{what}$ [$_{\theta P}$ t$_{how}$ ...

Postulating (155b) has the advantage of simplicity: only one type of route out of weak island is postulated. It will however be shown below that this is empirically inadequate; the intermediate position in reconstruction paradigms is indeed a successive theta-position, not an SQ-related position. Before showing this, let us however examine a radical solution to the reconstruction facts.

Logophoricity   It is sometimes suggested that the reconstruction issue does not arise because 'picture-phrases', or more generally anaphors within noun-phrases, do not obey strict locality conditions (e.g. principle A), but are rather 'logophoric', and as such, do not need to be reconstructed in order to be bound (see e.g. Pollard and Sag (1992), Heycock (1995)). This would entail that SQ-movement could be a one-step long movement as in (149), with no need for intermediate positions.

Such a solution however doesn't seem adequate. Besides simple contrasts such as (152-153), direct counter-evidence comes from the behaviour of those anaphors which are uncontroversially not logophoric: such anaphors are also capable of being bound by an antecedent below them.

Take for instance the German anaphor *sich*, the most well-known example of anti-logophoric anaphors. Its intermediate reconstruction is of course not testable in those German regiolects which disallow extraction from finite complements, but it is testable in southern variants. Here are relevant Swiss-German examples:[18]

(156)   a.   welles bild   vo sich dänkt dr Simon dass du   willsch
         which picture of self thinks the Simon that you want
         verchaufä?
         to.sell
         which picture of himself does Simon think that you want to sell?
     b.   welles bild   vo sich dänksch du dass dr Simon will
         which picture of self think   you that the Simon wants
         verchaufä?
         to.sell
         which picture of himself do you think that Simon wants to sell?
     c. * welles bild   vo sich dänksch du dass du hätsch chönnä
         which picture of self think   you that you have   could
         verchaufä?
         sell

---

[18]Data from Thomas Leu and Gabriel Lobos.

Examples (156a-b) show that reconstruction of *sich* is possible, and (156c) shows that the reconstructed site must have a local antecedent. Non-logophoric anaphors can thus access antecedents below them; and reconstruction is the sole mechanism allowing them to do that.

The same is true of the French *soi*, which is limited to ARB antecedents, and non-logophoric:

(157)   a.   quels doutes sur **soi-même**$_i$ est-ce que tu crois qu'il faut **PRO**$_i$ résoudre avant de commencer une vie de couple?

            which doubts about oneself do you think that it is needed to PRO resolve before starting a common life?

       b.   quels doutes sur **soi-même**$_i$ est-ce que tu te demandes comment il est possible de **PRO**$_i$ résoudre?

            which doubts about oneself do you wonder how it is possible to PRO solve?

       c.   * quels doutes sur **soi-même**$_i$ est-ce que tu te demandes comment il est possible que Jean ait résolu?

            which doubts about oneself do you wonder how it is possible that John solved?

       d.   quelles photos de **soi-même**$_i$ est-ce qu'il faut **PRO**$_i$ savoir comment l'accusateur a utilisé, pour bien se défendre?

            which pictures of oneself is it necessary PRO to know how the plaintiff used, in order to build a good defense?

(157a-b) show that *soi* can reconstruct into a lower clause, both with and without an island, (157c) illustrates the fact that a (lower) antecedent is required (i.e. *Jean* is not a possible antecedent for *soi*, not being ARB), and (157d) shows that it can reconstruct into an intermediate position, despite the presence of a wh-phrase occupying the intermediate comp area.

Frampton's facts thus need to be dealt with by reconstruction and eWI therefore involves successive cyclic movement of the type depicted in (155)

Reconstruction over long distance movement   An interesting pattern arises when reconstruction takes place while extracting over a wh-phrase which has itself undergone long wh-movement. This results in the two types of movement interleaving:

(158)   a. which portrait of **herself** is it unclear <u>how</u> **Simone** thinks that Humphrey should hang?

       b. pour quel portrait d'**elle-même** est-ce qu'il n'est pas clair <u>combien</u> **Simone** pense que Humphrey a donné d'argent?

for which portrait of herself is it unclear how much S thinks that H gave of money?

The interveners *how/combien* are the type of wh-phrases which require successive-cyclic Q-movement (adjuncts, 'subextraction'). The boldfaced anaphors take the boldfaced noun phrases as antecedents, and must thus reconstruct below them. In doing so, they reconstruct below *how/combien*. This entails, that *how/combien* has jumped over the reconstruction-site, ie:

(159) **wh-self** ... [$_{CP}$ <u>how</u> Simone <**wh-self**> [$_{CP}$ <<u>how</u>> ...

The availability of this structure resolves the issue left open above: is the intermediate reconstruction position an SQ-position or a θ-position?

Suppose it is an SQ-position: (159) now involves a Q-movement (of *how*) jumping over an SQ-link. This configuration is however illegal: an SQ intervener is a kind of Q, and therefore blocks all Q-movement. Suppose on the other hand that the intermediate reconstruction position is a θ-position: being of a different family of movement than Q-movement, it should allow the latter to proceed.

More generally, since the interleaving pattern shows that each involved movement can jump over the other, neither of them can be SQ-movements. We thus have evidence for successive cyclic θ-movement as a means to escape weak islands.

4.2   Long 'tout'-movement

The logic of the reconstruction paradigm is reminiscent of the traditional paradox associated with French long-*tout*, illustrated in:

(160) a. ? il a **tous** voulu que tu les invites <tous les>
           he all wanted that you them invite

       b. ? il a   **plus** voulu manger <plus> de spaghettis que toi
           he has more wante to-ead   <more> of spaghettis than you
           he wanted to eat a bigger quantity of spaghettis than the quantity
           you wanted to eat

       c. elle a **beaucoup** voulu emprunter de livres
           she many wanted to-borrow of books

       d. elle **rien** voulu que j'achète
           she nothing wanted that I buy
           There's nothing that she wanted me to buy

(It seems that this movement is also not restricted to quantifiers, as some dialects also allow this movement with *ça* ('this')).

The paradox comes from the fact that this movement can jump over both subjects (A interveners) and wh-phrases (A' interveners):

(161)   a.   ? il a **tous** voulu savoir <u>comment</u> les réparer <tous les>
                he all wanted to know how to them repare

        b.   ? elle a **tout** deviné <u>pourquoi</u> ils ont caché <tout>
                she has all guessed why they have hidden

        c.   ? ils ont **rien** avoué <u>pourquoi</u> ils ont soutenu <rien>
                they have nothing admitted why they supported

It is generally insensitive to inner islands, as the systematic contrast between *how* and *tout* indicates (with (162) as a baseline):

(162)   a.   comment est-ce qu'il aurait fallu qu'il les arrose?
                how should he have watered them?

        b.   ? il aurait tout fallu qu'il arrose
                he should have watered them all

(163)   a.   * **comment** est-ce qu'il aurait fallu que il les arrose <u>plus souvent</u>?
                how should he have watered them more frequently?

        b.   ? il aurait **tout** fallu que il arrose <u>plus souvent</u>
                he should have watered them all more frequently

(164)   a.   * **comment** est-ce qu'il aurait fallu ne <u>pas</u> les arroser?
                how should he have not watered them?

        b.   ? pour que ces plantes survivent, il aurait **toutes** fallu ne <u>pas</u> les arroser aussi fréquemment que tu l'as fait
                to keep these plants alive, it would have been necessary to not water them as frequently as you did

(165)   a.   * **comment** est-ce qu'il aurait fallu <u>beaucoup</u> les arroser?
                how should he have watered them a lot?

        b.   ? il aurait **tout** fallu <u>beaucoup</u> arroser
                he should have not watered them all a lot

Not only can long-*tout* jump over Q and A, but both Q-movement and A-movement can jump over long-*tout*. The addition of a wh-movement crossing long-*tout* for instance, (166b), doesn't add any significant degradation to (166a). The same sentences show that A-movement can cross long-*tout* since the VP-internal subject crosses it on its path to its VP-external position:

(166) a. ? elle a tous voulu qu'il les montre à Jean
           she all wanted that he them show to John
           she wanted him to show all of them to John

       b. ? **à qui** (est-ce que) elle a <u>tous</u> voulu qu'il les montre?
           to whom did she all want that he them show
           to whom did she want him to show all of them

These properties make this movement obviously similar to the θ-step of the eWI
discussed above: (i) both are able to jump over Q and A interveners, (ii) both
can be jumped over by Q- and A-movement, (iii) both land in a low position
inside the clause, (iv) both can 'scramble' across clauses. Two further similarities
confirm the parallelism between the θ-step of the eWI and *long*-tout: (v) both
can jump over adverbs ('modifiers'), (vi) both are unable to reconstruct for
scope.

θ-movement versus Modifier-movement   This approach to the locality of
long *tout*-movement in terms of θ-movement essentially follows Nicolis (1999),
modulo one node in the syntactic feature-tree: Nicolis argues that long-*tout* is
an instance of Rizzi's 'modifier'-movement (M-movement) - thus explaining its
insensitivity to A/A' interveners with the same logic as above.

Assigning long *tout*-movement to M-movement versus θ-movement does however
make one important difference with respect to locality: if long *tout*-movement
is M-movement, it should be blocked by any intervening adverb; if it is θ-
movement, it should be insensitive to intervening adverbs.

Given this difference, it is unlikely that long *tout*-movement is M-movement as
long *tout*-movement ignores intervening adverbs:

(167) a. il a   tout   voulu nettoyer
           he has everything wanted to-wash

       b. il a   tout   voulu lentement nettoyer
           he has everything wanted slowly   to-wash

Although (167b) jumps over an adverb, its acceptability is the same as the
baseline long-*tout* in (167a). Similarly:

(168) a. il a   lentement tout   nettoyé
           he has slowly   everything washed

       b. il a   tout   lentement nettoyé
           he has everything slowly   washed

The universal quantifier can perform a clause-internal left-shift above an adverb (discussed below), and doing so triggers the same acceptability as standard long-*tout* examples, e.g. (167). Adverbs *per se* thus don't create any intervention effect.[19]

As expected, the θ-step of wh-movement is also able to jump over intervening adverbs:

(169)   which picture of herself did **Mary** say $<$which picture of **herself**$>$
         that John <u>usually</u> looks at for hours

Here the wh-phrase reconstructs into the root clause, below *Mary*. To reach this reconstruction-site it has however crossed the embedded clause which contains the adverbial modifier *usually*, without incurring any locality-penalty.

Long-*tout* doesn't Reconstruct   The fifth property putting together long *tout*-movement and the successive cyclic θ-step of eWI is the absence of reconstruction for scope: scope is frozen in the surface position. Recall that a

All reconstruction paradigms discussed above for wh-movement were cases of anaphoric reconstruction, and this type of reconstruction is capable of reconstructing into a weak island:

(170)   a.   which portrait of herself do you know <u>when</u> **Maité** painted $<$which portrait of **herself**$>$

---

[19]Nicolis bases himself on (1) to argue that modifiers do in fact block long-*tout*.

(1)   * il a   tout   voulu obstinément voir en même   temps
         he has everything wanted stubbornly   to-see at the-same time

The ungrammaticality of (1) is however unrelated to *tout* and is part of the general pattern preventing many adverbials from appearing in complement infinitivals. The sentence (1) is ungrammatical irrespective of the presence of *tout*, (2), and becomes grammatical if *obstinément* is placed below *tout* in the root clause, (3c), at roughly the level of acceptability expected from the compounding of the independent degradation triggered by the mere presence of *obstinément*, (3a), and the usual degradation with long-*tout*, (3b). This paradigm thus lends further support to the fact that long-*tout* is insensitive to intervening modifiers.

(2)   * il a   voulu obstinément les   voir en même   temps
         he has wanted stubbornly   them to-see at the-same time

(3)   a.   ? il a   obstinément voulu les voir en même   temps
                he has stubbornly   wanted them to-see at the-same time

       b.   ? il a   tout voulu voir en même   temps
                he has all   wanted to-see at the-same time

       c.   ? il a   tout   obstinément voulu voir en même   temps
                he has everything stubbornly   wanted to-see at the-same time

b.   which portrait of herself do you wonder <u>why</u> **Maité** hates <which portrait of **herself**>

These cases contrast with reconstruction for scope, which is incapable of penetrating into a weak island. A recent discussion of lack of scope-reconstruction is e.g. Cresti (1995) who notes that functional readings of an eWI phrase are possible everywhere above the site of the island, but nowhere below the island:[20]

(171)   a.   [which human being]$_i$ do you think that [every person]$_i$ should love? (his neighbour)

     b.   [which human being]$_i$ does [every person]$_i$ think that you should love? (his neighbour)

(172)   a.   * [which human being]$_i$ is it unclear whether [every person]$_i$ should love? (his neighbour)

     b.   [which human being]$_i$ does [every person]$_i$ wonder whether you could love? (his neighbour)

Introducing a weak island selectively blocks scope-reconstruction under that island. Since the weak island is crossed by θ-movement, this means that Q-movement reconstructs for scope but θ-movement doesn't.

θ-movement of a wh-phrase therefore has the property of reconstructing for reference (anaphora) but not for quantity (scope) (while Q-movement has the converse property of reconstructing for scope, an interesting contrast that I will come back to in §8).

Exactly the same lack of reconstruction for scope scope holds of long *tout*-movement. This is particularly clear with *rien*:

(173)   a.   ? j'ai   rien   regretté d'avoir acheté
            I have nothing regretted to have bought
            ↪ I bought things and I don't regret it
        * ↪ I bought nothing and I regret it

     b.   j'ai   regretté d'avoir rien   acheté
            I have regretted to have nothing bought
        * ↪ I bought things and I don't regret it
            ↪ I bought nothing and I regret it

---

[20]Romero (1998) discusses one example of a cleft construction which does seem to reconstruct for scope into a weak island, but does not discuss why standard wh-movement does not do so; a judgement with which she agrees. I thus take it that the standard generalisation holds, despite her conclusions.

Notice that the text presentation presupposes that SQ-movement doesn't reconstruct for scope either, I will come back to this point.

Given these five similarities, I will take it for granted that both long *tout*-movement and the first step of wh-movement are instances of the same phenomenon, θ-movement. Because (i) it is overtly observable and (ii) it is not ambiguous with SQ-movement, long *tout*-movement provides an ideal tool to observe the general properties of θ-movement.

Inspection of the properties of long *tout*-movement does provide important cues as to the nature of θ-movement.

Range-based Presuppositions

First, long *tout*-movement reveals the presuppositional semantics of θ-movement, more precisely its range-based semantics. Given this newspaper headline:

(174)   les crues du   Nil ont commencé à tout   emporter sur
        the floods of-the Nile have started   to everything take-away on
        leur passage
        their path

The corresponding structure with long-tout is grammatical but could not be used for the same newspaper headline:

(175)   les crues du   Nil ont tout   commencé à emporter sur
        the floods of-the Nile have everything started   to take-away on
        leur passage
        their path

The difference between the two is that (174) states that the Nile has swept away everything on its way, whatever the objects that happened to be on its way were. (175) on the other hand presupposes that there is a set of objects on the Nile's path that we know about, and states that the Nile has now started to sweep out these objects. (175) is thus adequate as a newspaper headline only if there is shared public knowledge about salient objects threatened by the Nile.

The same contrast holds across the board: in the a examples below, *tout* is obligatorily anaphoric on some contextually reachable set, whereas in the b examples *tout* has no specified range beyond the usual pragmatic restrictions on universal quantification.

(176)   a.   ? il aurait   tout   voulu que tu aies acheté
             he would have everything wanted that you had bought

        b.   il aurait   voulu que tu aies tout   acheté
             he would have wanted that you had everything bought

(177) a. il a tout commencé à retaper
          he has everything began to rebuild

       b. il a commencé à tout retaper
          he has began to everything rebuild

Part of this contrast is certainly due to the different scopes of *tout*. Again, this is particularly clear with *rien*:

(178) a. il a rien commencé à retaper
          he has nothing began to rebuild

       b. # il a commencé à rien retaper
          he has began to nothing rebuild

Example (178b) is a contradiction, while (178a) is not: the low scope of *rien* in (178b) triggers a reading paraphrasable by 'he started something, and this something consists in not rebuilding anything'. While this can be made felicitous in some contexts, an 'out of the blue' reading is contradictory. No such 'out of the blue' oddity arises in (178a). This effect seems to be a simple scoping effect, caused by the fact discussed above that θ-movement does not reconstruct for scope.

Scope is however unlikely to be the whole story. First, wide-scope alone seems insufficient to trigger the presuppositional readings in the above examples: formulas such as 'for all x, he started to rebuild x' are not inherently more presuppositional than 'for all x, he rebuilt x'. If so, the presuppositionality of long-*tout* must come from something else than scope.

Second, and more importantly, the presuppositional reading is triggered independently of scoping over a predicate:

(179) a. il a déjà tout nettoyé
          he has already all washed

       b. ? il a tout déjà nettoyé
          he has all already washed

(179a) is the unmarked position for *tout*, below *déjà* in the hierarchy of adverbs (Cinque (1999)). It's reading in this position corresponds to the non-presuppositional reading of lower-clause *tout* in the b examples above.

It is however possible to place *tout* in front of *déjà*, with the characteristic slight degradation of long *tout*-movement. Doing so triggers the presuppositional meaning of long *tout*-movement. (179b) thus contrasts with (179a) in that it obligatorily refers to a set of 'things-to-clean' preexisting in the discourse ((179a) can but need not have such a reading). Since no issue of scope of *tout* over a predicate arises here, the presuppositions must be triggered independently of scope.

Following Cardinaletti and Starke (1999), Nicolis (1999), I take the short non-presuppositional movement of *tout* to be classical A-movement. The longer movement on the other hand (whether intra-clausal or cross-clausal) is an instance of θ-movement, which is the source of the presuppositional reading of *tout/rien* in these cases.

The resulting clause-structure thus looks like this (symbolising the host of A-movement by 'AgrP'):

(180)   ... $[_{\theta P}$ ... $[_{dejaP}$ ... $[_{AgrP}$ ... $[_{VP}$ ...

with all θ-related movement landing in θP.[21] These positions are all low, as indicated by the fact that negation precedes them:

(181)   a.   ils   ont   pas tous tout compris
             they have not all   all   understood

       b.   ? ils   ont   pas encore tous déjà   compris   ça
             they have not yet   all   already understood this

       c.   ... $[_{NegP}$ ... $[_{\theta P}$ ... $[_{dejaP}$ ... $[_{AP}$ ... $[_{VP}$ ...

Given these paradigms, I will from now on use the term 'high *tout*-movement' rather than 'long *tout*-movement' to indicate that this movement can be either intra-clausal or cross-clausal.

Very much the same presuppositionality effect holds in picture-anaphors, although the judgments are slightly less clear: in all examples of reconstruction discussed above, the presupposition is about a preexisting range for the wh-phrase, asking which item within the range satisfies the predicate. In no case was there an implication as to a specific item being presupposed, with the question asking for the identity of the that specific item.

This set of facts about *tout* and wh-movement under reconstruction thus clearly suggests that θ-movement maps onto range rather than specificity. If so, the θ-movement versus SQ-movement difference instantiates the contrast noted at

---

[21] Both of the following are however possible:

(1)   a.   il   a   tout   déjà   commencé à retaper
             he has everything already began   to rebuild

       b.   il   a   déjà   tout   commencé à retaper
             he has already everything began   to rebuild

   Both are presuppositional, and θ-movement can therefore optionally land in a position lower than *déjà*:

(2)   ... $[_{\theta P}$ ... $[_{dejaP}$ ... $[_{\theta P}$ ... $[_{AgrP}$ ... $[_{VP}$ ...

I will ignore this complication here.

the outset between range-based versus specificity-based readings of eWI, a point I will come back to shortly.

Notice finally that A-movement of *tout* is naturally clause-bound, since the subject A-position blocks it. There is however one case where this movement appears not to be clause-internal, structures stacking multiple infinitives:

(182) a. il   aurait voulu   pouvoir **tout**   acheter
           he has   everything wanted to-can to-buy
           he wanted to be able to buy everything

       b. il   aurait voulu   **tout**   pouvoir acheter
           he has   everything wanted to-can to-buy
           he wanted to be able to buy everything

       c. il   aurait **tout**   voulu   pouvoir acheter
           he has   everything wanted to-can   to-buy
           he wanted to be able to buy everything

Only (182c) is presuppositional. Although (182b) involves a non-local left-shift of *tout*, this movement does not trigger the presuppositional reading. The nature of this paradigm recalls the fact that extraction of wh-phrases is easier out of infinitives, in particular the fact noted in §2 about (44) that some intervention effects disappear when extracting out of infinitives. As elaborated below, I take both effects to have a common cause: infinitives allow a type of movement that tensed clause do not allow, but the nature of this movement will only become clear in §5 and the special nature of infinitivals will then be further discussed in §6.4.

The first property revealed by *tout*-movement is thus that θ-movement maps onto range.

The 'Round Robin' Nature of θ-movement

A second and important property of θ-movement revealed by *tout* is its 'round-robin' nature. High-*tout* can target either a subject, (183a), a direct object, (183b), or an indirect object, (183c).

(183) a. ? il   a   **toutes** voulu qu'**elles**   partent
             he has all   wanted that they$_{fem}$ leave

       b. ? il   a   **tous** voulu que je **les** invite
             he has all   wanted that I them invite

       c. ? il   a   **tous** voulu que je **leur**   parle
             he has all   wanted that I to-them speak

More than one can undergo θ-movement at the same time:

(184) a. ?? il a   **toutes** <u>tout</u> voulu qu'**elles**   admettent
           he has all$_{fem}$ all   wanted that they$_{fem}$ admit

       b. ?? il a   **toutes** <u>tous</u> voulu qu'**elles**   <u>leurs</u>   parlent
           he has all$_{fem}$ all$_{masc}$ wanted that they$_{fem}$ to-them speak

       c. ??? il a   **toutes**$_i$ <u>tous</u>$_k$ voulu que je <u>les</u>$_k$ **leur**$_i$ montre
            he has all   all   wanted that O them to-them show

but in such cases, a strict order 'nominative > dative > accusative' must be preserved:

(185) a. * il a   <u>tout</u> **toutes** voulu qu'**elles**   admettent
           he has all   all$_{fem}$ wanted that they$_{fem}$ admit

       b. * il a   <u>tous</u>   **toutes** voulu qu'**elles**   <u>leurs</u>   parlent
           he has all$_{masc}$ all$_{fem}$ wanted that they$_{fem}$ to-them speak

       c. * il a   <u>toutes</u>$_i$ **tous**$_k$ voulu que je <u>les</u>$_i$ **leur**$_k$ montre
           he has all   all   wanted that O them to-them show

The same ordering restriction holds with short-movement:

(186) a. ? **elles**   ont **toutes** <u>tout</u> admis
           they$_{fem}$ have all$_{fem}$ all   admitted

         * **elles**   ont <u>tout</u> **toutes** admis
           they$_{fem}$ have all   all$_{fem}$ admitted

       b. ??? **elles**   <u>leur</u>   ont **toutes** <u>tous</u> parlé
           they$_{fem}$ to-them have all$_{fem}$ all   spoken

         ?* **elles**   <u>leur</u>   ont <u>tous</u> **toutes** parlé
           they$_{fem}$ to-them have all   all$_{fem}$ sold

       c. ?? je **les**$_i$ <u>leur</u>$_k$ ai   <u>toutes</u>$_k$ **tous**$_i$ montré
           I them to-them have all   all   shown
           I showed them all to all of them

       d. * je **les**$_i$ <u>leur</u>$_k$ ai   **tous**$_i$ <u>toutes</u>$_k$ montré
           I them to-them have all   all   shown

(Combining all three in a single sentence breaks down, both with short and long movement.)

The description is thus that long θ-movement must preserve the c-command ordering of the moved constituents. This is the same property widely observed for A-like movement in some scrambling contexts (see Haegeman (1993) for discussion of such scrambling facts in West Flemish) and in some wh-movement settings (e.g. Watanabe (1992)).

One important property noted by Haegeman, is that this order-preservation constraint also holds for elements originating from distinct clauses. She for

instance notes cases in which a higher subject and a lower ECM subject move in a round-robin fashion (together with a dative and an accusative, thus giving rise to a 4-element round-robin: 'NOM > NOM > DAT > ACC').

The same effect is visible with *tout*, combining quantifiers originating in the root clause and in the embedded clause:

(187) a. ? **elles** ont **toutes** <u>tous</u> voulu qu'<u>ils</u> partent
           they$_{fem}$ have all$_{fem}$ all$_{masc}$ wanted that they$_{masc}$ leave

       b. * **elles** ont <u>tous</u> **toutes** voulu qu'<u>ils</u> partent
           they$_{fem}$ have all$_{masc}$ all$_{fem}$ wanted that they$_{masc}$ leave

(188) a. ? **elles** ont **toutes** <u>tout</u> voulu voir
           they$_{fem}$ have all$_{fem}$ all wanted to-see

       b. * **elles** ont <u>tout</u> **toutes** voulu voir
           they$_{fem}$ have all all$_{fem}$ wanted to-see

Here the structural hierarchy between the higher and the lower subject must be preserved. The same is true of a higher dative and a lower nominative, leading to a 'DAT > NOM' order instead of the usual 'NOM > DAT':

(189) a. ?? je **lui** ai <u>toutes</u> dit qu'<u>elles</u> devaient partir
           I to-him have all$_{fem}$ said that they$_{fem}$ should leave

       b. ?? je **leur** ai **tous** <u>toutes</u> dit qu'<u>elles</u> devaient partir
           I to-them have all$_{masc}$ all$_{fem}$ said that they$_{fem}$ should leave

       c. * je **leur** ai <u>toutes</u> **tous** dit qu'<u>elles</u> devaient partir
           I to-them have all$_{fem}$ all$_{masc}$ said that they$_{fem}$ should leave

The baseline is given by the somewhat marginal (189a), with (189b) giving rise to the same marginality. The crucial contrast is between (189b), marginal, versus (189c), unparsable.

Such paradigms show that the relevant constraint is not about 'NOM > DAT > ACC', but is rather a purely structural constraint enforcing preservation of c-command hierarchy; with 'NOM > DAT > ACC' an accidental byproduct due to the base order of these phrases.

The constraint also extends to quantifiers or quantifier-like elements associated to subsets of arguments:

(190) a. <u>ils</u> ont <u>tous</u> **plus** voulu acheter <plus> de livres que de
           they have all more wanted buy <more> of books than of
           gateaux
           cakes
           they all wanted to buy more books than cakes

78

    b.   * <u>ils</u> ont **plus** <u>tous</u> voulu acheter \<plus\> de livres que de gateaux

(191) a.   il a   <u>tous</u> **plus** voulu <u>leur</u>   acheter \<plus\> de livres
             he has all   more wanted to-them buy   \<more\> of books
             que de gateaux
             than of cakes
             for each one of them, he wanted wanted to buy more books than
             cakes
    b.   * il a **plus** <u>tous</u> voulu <u>leur</u> acheter \<plus\> de livres que de gateaux

An important consequence of the order-preservation constraint, capitalised on below, is that prolonging the movement is not possible in the configuration (192a), but is fine in (192b-c):

(192) a.   * XP_2 ... XP_1 XP_2 XP_3
      b. XP_1 ... XP_1 XP_2 XP_3
      c. XP_1 XP_2 ... XP_1 XP_2 XP_3

Of course, this only applies if the further movement is the same kind of movement, here θ-movement. Forking to other types of movement makes configuration (192a) acceptable.

Again, this can be directly illustrated with θ-movement of *tout*, in *tout*-splitting contexts where two quantifiers originating from the same clause land in different clauses: (193a) instantiates (192a) and (193b) instantiates (192b).

(193) a.   * J'aurais   <u>tout</u> voulu **tous leur**   montrer
             I would-have all   liked to-all to-them show
    b.   ? J'aurais   **tous** voulu <u>tout</u> **leur**   montrer
             I would-have to-all liked all   to-them show
             I would have liked to show it show all to all of them

The general form of the round-robin constraint is thus (194), a path-containment restriction.

(194)   * α_j α_i ... α_i α_j

Clearly, the standard *Relativised Minimality* configuration, (195) is a special case of (194): the case where the intervener is a degenerate one-member chain.

(195)   * α_j ... α_i ... α_j

The obvious interpretation of the round-robin constraint is thus that it reveals a generalised form of *Relativised Minimality*. The difficulty with this is that it requires a view of *Relativised Minimality* as constraining 'chains' rather than individual landing sites (such that α-chain$_1$ may not 'cross' α-chain$_2$ as a whole). This in turn entails the introduction an undesirable calculus on chains. (See Richards (1997), Haeberli (1999) for an approach in terms of the 'tucking in' mechanism, which however also requires special assumptions, both undesirable *per se* and incompatible with the general framework developed below.)

In short, the nature of the round-robin suggests an obvious general interpretation, but it is unclear how to implement that interpretation given the current state of the theory. It will be argued below in §8.3 that the generalised version of *Relativised Minimality* - and therefore the robin-constraint - become easily expressible once the relationship between merge and move is clarified.

SQ-movement and *tout*

A third property of θ-movement revealed by *tout* is its relationship to SQ-movement. The illegal 'inverted' orders, e.g. (185) improve if the first universal quantifier is interpreted as specific, and focused:

(196) a. ?? il a   TOUT **toutes** voulu qu'**elles**   admettent
           he has all   all$_{fem}$ wanted that they$_{fem}$ admit

       b. ?? il a   TOUS **toutes** voulu qu'**elles**   <u>leurs</u>   parlent
           he has all$_{masc}$ all$_{fem}$ wanted that they$_{fem}$ to-them speak

       c. ??? il a   <u>TOUTES</u>$_i$ **tous**$_k$ voulu   que je <u>les</u>$_i$ **leur**$_k$ montre
           he has all   all   wanted that I   them to-them show

       d. ??? il a   TOUT **toutes** voulu   que je **leur**   montre
           he has all   all   wanted that I to-them show
           he wanted me to show *everything* to all of them

The reverse is however not true:

(197) a. * il a   **tout** <u>TOUTES</u> voulu   qu'<u>elles</u>   admettent
           he has all   all$_{fem}$ wanted that they$_{fem}$ admit

       b. * il a   **tous** <u>TOUTES</u> voulu   qu'<u>elles</u>   **leurs**   parlent
           he has all$_{masc}$ all$_{fem}$ wanted that they$_{fem}$ to-them speak

       c. * il a   **toutes**$_i$ <u>TOUS</u>$_k$ voulu   que je **les**$_i$ **leur**$_k$ montre
           he has all   all   wanted that O them to-them show

It also seems to be the case that only one *tout*-phrase can undergo such a movement regardless of ordering, as both of the following are unparsable:

(198) a. * il a TOUT <u>TOUTES</u> voulu qu'<u>elles</u> admettent
           he has all all$_{fem}$ wanted that they$_{fem}$ admit

     b. * il a <u>TOUTES</u> TOUT voulu qu'<u>elles</u> admettent
           he has all$_{fem}$ all wanted that they$_{fem}$ admit

I take this to indicate that a specific lands in a position slightly higher than the usual θ-movement:

(199) ... [$_{NegP}$ ... [$_{SqP}$ ... [$_{\theta P}$ ... [$_{dejaP}$ ... [$_{AP}$ ... [$_{VP}$ ...

That this movement does not participate in the round-robin is also shown by the fact that it can rescue order-inversions involving quantifiers originating from distinct clauses, as in (187b-188b) above:

(200) a. ?? **elles** ont <u>TOUT</u> **toutes** voulu voir
           they$_{fem}$ have all all$_{fem}$ wanted to-see

     b. ?? **elles** ont <u>TOUS</u> **toutes** voulu qu'<u>ils</u> partent
           they$_{fem}$ have all$_{masc}$ all$_{fem}$ wanted that they$_{masc}$ leave

The SQ route out of weak islands is therefore modified from (150), repeated as (201),

(201) **what**$_{SQ}$ ... V [$_{CP}$ <u>how</u>$_Q$ ... [$_{\theta P}$ **t**$_{what}$ [$_{\theta P}$ t$_{how}$ ... [$_{VP}$ ... **t**$_{what}$ t$_{how}$

so as to take into account the SqP landing site. The result is either of:

(202) a. **what** ... [$_{SqP}$ **t**$_{what}$ ... [$_{CP}$ <u>how</u>$_Q$ ... [$_{SqP}$ **t**$_{what}$ [$_{\theta P}$ **t**$_{what}$ [$_{\theta P}$ t$_{how}$
     b. **what** ... [$_{SqP}$ **t**$_{what}$ [$_{\theta P}$ **t**$_{what}$ ... [$_{CP}$ <u>how</u>$_Q$ ... [$_{\theta P}$ **t**$_{what}$ [$_{\theta P}$ t$_{how}$
     c. **what** ... [$_{SqP}$ **t**$_{what}$ ... [$_{CP}$ <u>how</u>$_Q$ ... [$_{\theta P}$ **t**$_{what}$ [$_{\theta P}$ t$_{how}$ ...

That is, either θ-movement forks to SQ-movement in the lower clause, and the cross-clausal 'scrambling' step is an SQ-movement, (202a), or the cross-clausal 'scrambling' step is θ-movement which forks off to SQ-movement only in the higher clause, (202b). Finally, (202c) is an adaptation of the idea that SQ-movement is 'long' movement, with the lower θ-step directly feeding a jump to the higher SqP.

The percourse (202a) is however illegal, as the intermediate wh-phrase (*how* here) would cross an SQ-link of the *what*-chain. I.e. forking to SQ-movement cannot happen below the intervening Q; leaving only (202b) and (202c) as legal SQ-based routes out of weak islands (notice that the latter presupposes that SqP doesn't project in the embedded clause).

81

Negation and the two Series of Positions

The above conclusion that SqP and θP are low projections conflicts with the fact that presuppositional wh-movement is capable of jumping over negation. As it stands, this should be impossible:

(203) $[\textbf{wh} \ldots [_{NegP} \text{Q} [_{SqP} \ldots [_{\theta P} <\textbf{wh}> \ldots [_{VP} \ldots <\textbf{wh}>$

A Q-movement from θP or SqP to the Q landing site in CP would cross over Neg and triggering a Q-across-Q violation. There must therefore be θ- or SQ-landing sites above negation.

I will follow Starke (1993) here in his argument that each 'clause' is composed of at least two repeating and restructured sub-clauses, thereby giving rise to two series of θP or SqP positions. These two series of positions may correspond to what is informally referred to as the 'low' scrambling and 'high' scrambling positions (possibly, the high θP position corresponds to the TopP positions of Rizzi (1997), targeted by Romance topicalisation, thereby unifying the two instances of iterative projections):

(204) $[_{CP} \ldots [_{SqP} \ldots [_{\theta P} \ldots [_{VP=AuxP} [_{CP=NegP} \ldots [_{SqP} \ldots [_{\theta P} \ldots [_{VP} \ldots$

(Given the results of Cinque (1999), the hypothesis that there are only two subunits to a clause seems too conservative. Cinque's results, combined with those of Rizzi (1997) and Beghelli and Stowell (1997) suggest that there is a recursion of at least five or six such subunits, or sub-clauses within a complete tensed clause.)

Anti-Locality Effect

Finally, let us mention a somewhat mysterious paradigm - for which I have no explanation: *tout*-movement seems to show an inverse sensitivity to wh-interveners as compared to Q-movement.

Consider again the initial examples showing the insensitivity of long-*tout* to intervening wh-islands, or other such examples in the literature, eg:

(205)   ?il a   **tout** compris   <u>comment</u> il   faut   réparer
         he has all   understood how   one should repare
         he understood how to repare everything

A curious property of these examples is that they all involve adjunct wh-phrases *qua* interveners, and become entirely unacceptable if the intervener is replaced by a declarative complementiser:

(206)   * il a   **tout** compris   <u>qu'</u>il faut réparer
            he has all   understood that one should repare

The exact inverse pattern holds with pure wh-movement:

(207)   ? quand est-ce que tu as compris <u>qu'</u>il faut le réparer <quand>
          when did you understand that we should repair it?

(208)   * quand est-ce que tu as compris <u>comment</u>'il faut le réparer <quand>
          when did you understand how we should repair it?

Predicates such as *think* are the paradigm case of wh-extractability, but they have the opposite effect on *tout*-movement, strongly blocking it:

(209)   comment tu crois que je vais les cuire?
         how de you think that I'll cook them?

(210)   * tu crois tout que je vais cuire
         you think everything that I'll cook

Changing the predicate to a factive degrades pure wh-movement, but ameliorates *tout*-movement:

(211)   a. * comment tu as   regretté qu'il   soit parti?
             how   you have regretted that he be   left
             how do you regret that he left?
       b. ? j'ai   tout regretté qu'il   ait   jeté
            I have all   regretted that he threw away

Again the two patterns are exactly the inverse: θ-movement kicks in where Q-movement fades out.

Summary

§3 explored θ qua blocker (of overt wh-movement), concluding that θ-across-θ effects underlie strong islands. The present section examined the properties of successful θ-movement, which turn out to be:

1. can jump over A, Q and M interveners

2. allows A, Q to jump over it

3. can feed Q and SQ movement

4. reconstructs for binding but does not reconstruct for scope

5. induces range presuppositions

6. respects the round-robin constraint

The first three properties directly follow from the feature-tree in (115) above and repeated here:

(212)

Quantifier   M   A

SpecificQ θ
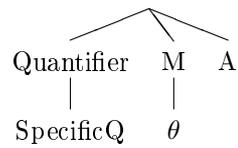
The fourth and fifth property will play an important role in that they show that each type of chain has its own interpretive characteristics; and the sixth property will allow us to refine the formulation of the locality principle.

The possible routes to escape a weak islands are now:

(213)   a. **what** ... $[_{\theta P}$ $\mathbf{t}_{what}$ ... V $[_{CP}$ $\underline{how}_Q$ ... $[_{\theta P}$ $\mathbf{t}_{what}$ $[_{\theta P}$ t_{how} ...

           b. **what** ... $[_{SqP}$ $\mathbf{t}_{what}$ $[_{\theta P}$ $\mathbf{t}_{what}$ ... $[_{CP}$ $\underline{how}_Q$ ... $[_{\theta P}$ $\mathbf{t}_{what}$ $[_{\theta P}$ t_{how}

           c. **what** ... $[_{SqP}$ $\mathbf{t}_{what}$ ... V $[_{CP}$ $\underline{how}_Q$ ... $[_{\theta P}$ $\mathbf{t}_{what}$ $[_{\theta P}$ t_{how} ...

5   The 'argument/adjunct' Asymmetry: R(θ, presuppositions)

Ever since Huang (1982), contrasts such as (214) have been on the top of the eWI agenda, and are standardly taken to suggest that the factor underlying eWI is linked to argumenthood: arguments can extract from weak islands, adjuncts cannot.

(214)   a.   what do you wonder whether to cook today?
        b.   * how do you wonder whether to cook today?

There are two issues here:

- what is the nature of the difference between (214a) and (214b)? Where is the border between elements which pattern with (214a) and those which pattern with (214b)? The answer should give an important cue about the nature of eWI.

- how does (214) relate to presuppositionality? I.e. if the presuppositional status of the extracted wh-phrase is what underlies eWI - as argued in §2 - why is there an additional 'argument/adjunct' asymmetry?

5.1   Where the Border between 'Arguments' and 'Adjuncts' Lies

Setting aside the interpretive aspects of eWI, there are two main hypotheses as to the nature of the difference in (214): either the difference is in the nature of the relation between the predicate and the moved wh-phrase, and argumenthood is the crucial underlying notion (e.g. Rizzi (1990)), or alternatively the wh-phrase being a cased DP (rather than an uncased adverbial) is the crucial underlying difference (e.g. Huang (1982), Manzini (1992), Rizzi (2000)).

The sole available argument distinguishing these two families of approaches is, to my knowledge, one negative argument:

[i] argumenthood cannot be the relevant notion *per se* because a given wh-phrase has a constant behaviour in eWI, regardless of its argumental status. The locative *where* for instance patterns with 'arguments' wrt. its extraction possibilities regardless of it being selected or not:

(215)   a.   ? where do you wonder whether he will put the cake?
        b.   ? where do you wonder whether he will eat the cake?

Conversely, *how* patterns with 'adjuncts' in being strongly degraded in eWI, irrespective of its argumental status:

(216) a.   * how do you wonder whether he will behave?

          b.   * how do you wonder whether he will sleep?

While this is a strong indication that argumenthood *per se* is not the underlying factor, this argument doesn't indicate the source of the difference. Here are then two arguments to the effect that being a cased DP rather than an uncased adverbial is what underlies (214).

[ii] consider the two readings of *how*:

(217) a. how_{instr} do you think we should open the door? With these strange keys.

          b. how_{manner} do you think we should open the door? Silently!

which both resist extraction:

(218) a.   * how_{instr} don't you know whether you should open the door? (with these strange keys)

          b.   * how_{manner} don't you know whether you should open the door? (silently)

These two readings can be expressed by two different wh-phrases in languages which have wh-phrase with instrumental case alongside with adverbials. In the Slovak examples below, *cim* is a wh DP with instrumental case and *jak* is an uncased adverbial. Both can express the instrumental reading:

(219) a.   cim   mame   otvorit tie dvere? (with a key | *silently)
             how_{instr} should-we to-open the door

          b.   jak   mame   otvorit tie dvere? (with a key)
             how_{instr} should-we to-open the door

          c.   jak   mame   otvorit tie dvere? (silently)
             how_{mnnr} should-we to-open the door

Although the instrumental reading of *jak* has the same meaning as *cim*, the two differ in eWI:

(220) a.   cim   nevies   ci mas   otvorit tie dvere?
             how_{instr} you-not-know if you-should to-open the door

          b.   * jak   nevies   ci mas   otvorit tie dvere?
             how_{instr} you-not-know if you-should to-open the door

          c.   * jak   nevies   ci mas   otvorit tie dvere?
             how_{manner} you-not-know if you-should to-open the door

Both the argumental status and the interpretation is constant between (220a) and (220b) but their extraction possibilities are nonetheless different. The sole relevant asymmetry is that *cim* is a cased DP while *jak* is an uncased adverbial. The same reasoning holds for the difference between the English (232a) and the Slovak (220a). Conversely, keeping case/DPhood constant but changing the semantic relations / argumenthood does not trigger changes in judgments: (218a-b) and (220b-c).

More generally, the eWI examples with uncased adverbial wh-phrases are ungrammatical irrespective of their reading, while extraction of the cased phrase is acceptable.

Notice in passing that the same asymmetry is found in English in preposition stranding contexts:

(221) a. which key did you (wonder whether to) open the door with?
       b. ?* which attitude did you (wonder whether to) open the door with?

[iii] Another argument showing the relevance of case is given by the nature of the cut-off point between wh-phrases which pattern with (214a) and those which pattern with (214b). The former are cased DPs, the latter uncased adverbials.

(222)   who, what > to-whom, where, with-what(instr) ≫ when > how, why
       a. who do you wonder whether I should invite?
       b. what do you wonder whether I will cook?
       c. (?) to whom do you wonder whether I will tell how to cook this?
       d. (?) where do you wonder whether they cook this kind of things?
       e. (?) *(see (220a))*
       f. ?* when do you wonder whether they cook this kind of things?
       g. * how do you wonder whether they cook this kind of things?
       h. * why do you wonder whether they cook this kind of things <why>?

While there is a constant degradation, there is a main main cut-off point is between the locative *where* and the temporal *when* (e.g. Rizzi (1990), Szabolcsi (1999)). This split does not correspond to argumenthood (locatives, can be either arguments or adjuncts, and so can manner adverbs). But it does seem to correspond to case: natural language has locative case, but it does not have temporal case (distinguishing case suffixes from postpositions). If so, the set of extractable wh-phrase again corresponds to the set of case-marked wh-phrases.

The significance of the cut between locatives and temporals is confirmed by a number of phenomena. Take for instance relative pronouns: in a number of languages, wh-pronouns can be used as relative pronouns, but only up to *where* in the scale (222). Temporals, manner, cause, etc. cannot. Such languages include French, Italian, Slovak, Czech, etc.:

(223) a. l'endroit ou je suis né
            the-place where I was born

       b. * le moment quand je suis né
            the moment when I was born

       c. * la manière comment je l'ai fait
            the manner how   I did it

       d. * la raison pourquoi je l'ai fait
            the reason why   I did it

To pick another related phenomenon, GunGbe has an 'argument/adjunct' asymmetry with respect to topic-movement in which temporals pattern with adjuncts and locatives pattern with arguments. Gun topics occur in the left periphery of the clause, and are followed by a (topic) particle, *ya*:

(224) a. Kofi-ya Assiba na de-è
            Kofi-top Assiba fut marry-him

       b. * Kofi Assiba na de-è
            Kofi Assiba fut marry-him

The *ya* particle is however only present with arguments:

(225) a. * Ene utu-ya, yokpo le   na ton
            this cause, child   plur fut go-out
            because of this, the children will leave

       b. Ene utu,   yokpo le   na ton
            this cause, child   plur fut go-out
            because of this, the children will leave

Temporal modifiers pattern with adjuncts in that they occur without the particle, while locative modifiers pattern with arguments and require the particle:

(226) a. egbe, un na ton
            today, I   fut go-out
            today, I will go out

       b. * to   lo   to, Kofi na yi flen
            river specif side, Kofi will go there
            to that riverside, Kofi will go

A similar phenomenon is found in French, but with complications that would bring us far astray: its general (and slightly imprecise) description is that a left-dislocation of a temporal survives without a resumptive clitic, (227a), whereas dislocating a locative modifier requires a resumptive clitic, (227b).

88

(227) a.   à midi, je pars
             at noon, I leave

       b. * à Genève, j'habite
             in Geneva, I live

Further phenomena distinguishing locatives from temporals - independently of
extraction and of the argument/adjunct asymmetry - include the fact that there
exist 'locative inversion' constructions but no 'temporal inversion' constructions;
the fact that locative clitics are massively reported whereas temporal clitics are
either rare or non-existent depending on the analysis of a handful of cases; the
fact that there are locative expletives but no temporal expletives.

In sum, the cut-off point between extractable and non-extractable wh-phrases
corresponds to a general distinction between two types of modifiers, and the
distinction seems to correspond to cased versus uncased modifiers.

Given that (i) extractability does not covary with argumental status, (ii) within
a given reading (e.g. instrumental) extraction is possible exactly where the wh-
phrase is a cased DP, and (iii) across readings, the difference between those wh-
phrases that extract and those that don't corresponds to the difference between
cased and uncased phrases; it seems reasonable to infer that being a cased DP
is what underlies Huang's asymmetry.

Wh-borrowing: DP versus Case

It seems possible to further isolate the source of Huang's contrast, distinguishing
between DP-hood and case.

The argument builds on the fact that (inherent) case is tightly coupled with
thematic roles. The discussion of the importance of case for Huang's contrast
was for instance almost entirely framed in terms of thematic-roles (instrumen-
tal versus manner in the *how*-argument, locative versus temporal in the scale-
argument) because of the tight relation holding between inherent case and the-
matic roles.

To ease the discussion, let us introduce some terminology, bringing out the
relationship explicitly. Recall from §3 that we are using the term θ-roles in
an extended sense, to cover all relationships between a predicate and its direct
modifiers. In this use of the term, temporal or manner adjuncts bear a θ-role.
Roles such as temporal or manner are however never associated to inherent case;
and therefore only a proper subset of θ-roles are ever associated with inherent
case. Let us then use the term $i\theta$ for those θ-roles which do correspond to
(inherent) case.

Given this distinction, consider the 'wh-borrowing' paradigm. 'Wh-borrowing'
occurs when a meaning 'borrows' a wh-morpheme canonically assigned to a
different meaning. This English example was for instance noted by Ross (1984):

(228)   what does this cost?

Here the amount meaning (canonically expressed by *how much*) 'borrowed' the wh-word *what*. This borrowing switches the wh-phrase from a cased $i\theta$-roles (e.g. theme) to an uncased θ-roles (e.g. amount).

Given the null assumption that a wh-phrase maintains its categorial status during borrowing, the question is whether the borrowed wh-phrase behaves in eWI according to its newly acquired casual status (uncased 'amount' adverbial, non-extractable), or according to its preserved categorial status (DP, extractable).

(228) is however not as clean as it should be, since amount phrases introduce their own complications wrt. eWI (see §7.3). A cleaner paradigm is given by this French instance of wh-borrowing:

(229)   a.   **ou**   est-ce que tu   as   mis mes livres?
             where is-it   that you have put my   books

        b.   **quand** est-ce que c'est arrivé?
             when   is-it   that it-is arrived
             when did it happen?

(230)   a.   l'endroit **ou**   j'ai   mis mes livres
             the-place where I-have put my books

        b.   le   moment **ou**   c'est arrivé
             the moment where it-is arrived
             the moment when it happened

Canonically, the locative *ou* contrasts with the temporal *quand*, (229). In (230) the *ou* form is however 'borrowed' by the temporal θ-role, (230b). This borrowing switches *ou* from an cased $i\theta$-role (locative) to an uncased θ-role (temporal). Again, the null assumption is that the categorial status remains unchanged: *ou* is a DP throughout.

Does eWI of a 'borrowed' *ou* pattern with its preserved categorial status (DP, extractable) or with its newly acquired casual status (uncased temporal adverbial, non-extractable)? The latter:

(231)   a.   l'endroit ou   je sais pourquoi tu   as   mis mes livres *t*
              the-place where I know why   you have put my   books
              the place where I know why you put my books

        b. ?* le   moment ou   je sais pourquoi c'est arrivé   *t*
              the moment where I know why   it-is arrived
              the moment when I know why it happened

(The same is true in (228), potentially making the same point, with complicating factors as mentioned above.)

Switching away from inherent case - i.e. switching from $i\theta$ to θ - thus causes a switch from 'argument' behaviour to 'adjunct' behaviour. That the proper formulation is in terms of case and not θ *per se* is shown by the fact that this switch occurs only in DPs. Recall for instance the instrumental paradigms, repeated here for English:

(232) a. * how$_{instr}$ don't you know whether you should open the door? (with these strange keys)

  b. * how$_{manner}$ don't you know whether you should open the door? (silently)

Even though the instrumental role is an $i\theta$-role, switching to it does not facilitate eWI. Inherent case - i.e. 'DP with $i\theta$' - is therefore the correct description of Huang's 'argument/adjunct' asymmetry: wh-DPs with $i\theta$-role extract more easily from eWI.

One potential argument in favor of DP-hood and against inherent case turns out to e a non-argument: the fact that non-$i\theta$ phrases become better in eWI if they are not adverbials, "It has been noted that when counterparts of *why* and *when* have an articulated PP structure in a language, they extract better. Korean, Japanese, and Hungarian are cases in point" (Szabolcsi (1999)). French and English show the same:

(233) a. avec quel outil est-ce qu'il n'est pas clair comment il l'a ouvert?

    with which tool is it unclear how he has opened it?

  b. pendant quelles vacances est-ce qu'il n'est pas clair comment il s'est comporté?

    during which vacations is it unclear how he has behaved?

  c. quand est-ce qu'il n'est pas clair comment il s'est comporté?

    when is it unclear how he behaved?

(233b) is a temporal PP and therefore not an $i\theta$ but it nevertheless patterns together with $i\theta$ such as (233a), and against temporal adverbials such as (233c). This might be taken to favor DP-hood as the relevant factor: PPs involve a DP but adverbials don't, for the same θ-role. This paradigm is however neutral: the preposition in (233b) assigns inherent case to its complement, while no case is present in (233c). (233b) thus plausibly involves pied-piping of inherent-case and thus patterns with cased extraction.

5.2 Licensed Theta-Roles: Presuppositions versus Case

Given that inherent case (more precisely, being a DP with $i\theta$) is what underlies Huang's asymmetry, we now seem to have two unrelated conditions on extraction out of weak islands: 'only cased wh-phrase can extract' and 'only presuppositional wh-phrases can extract'. Why should there be these two (apparently) unrelated conditions?

Bizarre though it seems at first sight, the connection between iCase (inherent case) and presuppositions is well attested. Recall for instance the discussion of specificity in Enç (1991), partly based on the optional accusative case-marking on noun-phrases in Turkish. The generalisation investigated by Enç is that morphological presence of accusative case correlates with specificity: all and only the case-marked accusative noun-phrases are specific.

The Turkish generalisation is exactly the same as the connection suggested by Huang's asymmetry, as becomes clear once misleading terminology is cleared up. First, Enç uses 'specific' in the same sense as we are using 'presuppositional', i.e. to subsume both 'partitive specificity' (our range), and 'familiar specificity' (our specificity). Rephrasing, the Turkish generalisation is that accusative morphology correlates with presuppositionality. Second, the notion of 'case' ('accusative') used in such a description is different from the traditional notion of case (such as it is used in, say, Slavic or Latin): the latter refers to a marking that all DPs have, regardless of their interpretation. Within standard terminology, a more plausible reformulation of the Turkish situation is that the accusative suffix marks *inherent* accusative case, which is optional. If so, the Turkish generalisation is that 'accusative inherent case correlates with presuppositionality', transparently revealing the link between inherent case and presuppositionality.

Very much the same obtains with the Spanish marker 'a' appearing in front of accusative noun-phrases and described as marking 'specificity' and instantiating a type of preposition often taken to spell-out case (and similarly with the Rumanian 'prepositional accusative', tied to specificity).

The connection between iCase and presuppositionality is also observable in Slovak:

(234) a.   chcel   som kupit' gitar-u
              wanted I.was to.buy guitar-ACC
              I wanted to buy a guitar

        b.   chcel   som hovorit' pan-ovi
              wanted I.was to.speak man-DAT
              I wanted to speak to the man

The unmarked interpretation of an accusative (structural) complement under a modal is that of a non-specific indefinite, whereas the unmarked interpretation of a dative (inherent) in the same context is that of a specific definite.

It thus appears that the two constraints on eWI are not independent: requiring inherent case amounts to requiring presuppositionality.

Case is not Necessary

The second observation relevant to the relation between the iCase-requirement and the presuppositionality-requirement is the fact that presuppositionality are necessary but iCase isn't.

The discussion in §2.1 showed that case alone is not sufficient: all wh-phrases used there were cased (objects), but eWI was nevertheless possible only when the wh-phrase was associated with a presuppositional reading. Presuppositions are thus a *sine qua non* for eWI.

This points to a flaw in the discussion of the 'argument/adjunct' asymmetry above: the examples used to show that uncased adverbials ('adjuncts') resist eWI did not control for presuppositionality. Since presuppositionality is a *sine qua non* in eWI, the fact that {-presuppositional; -case} adjuncts do not extract can of course not be imputed to absence of case.

Controlling for presuppositionality, it turns out that adjuncts *are* extractable: it is a known fact that 'adjuncts' fare much better in eWI when presuppositional, in fact roughly at the level of 'arguments'. The simplest test are the so-called 'checklist' contexts:

(235)   a.   when did you want to know whether we will go to the mountains: for the snow-vacations or for the cow-herding in the summer?

        b.   how didn't he want to eat the dish: with a fork or with Chinese sticks?

Specificity induced by previous discourse triggers the same amelioration:

(236)   Studying the actions of your government, you chanced upon an odd manuscript describing a certain particularly rotten manner of acting (M) that the government pledged to use at most once, on a significant date. You believe this manuscript, and you feel that if you could find out when the government acted in an M-way, you would be on the path to important discoveries. You are thus very anxious to find out when the government used M.
        If I know about your efforts, but I don't know what M is, I can ask any of:
        a.   how would you like to know when the government acted?
        b.   how do you wonder when the government acted?

When such a context is prominent, the sentences are adequate, and in fact roughly at the same level of degradation as argument-extraction out of weak islands. The same holds for the other uncased wh-phrases *when* and *why_{motivation}*. (Notice again that all these examples involve a slight intonational peak on the wh-phrase, similar to the intonational peak on some situ-wh phrases, particularly under weak islands.)

Strictly speaking, case is thus not a prerequisite; presuppositions are.

Licensed Theta-Roles

The - surprising - descriptive generalisation underlying the 'argument/adjunct' asymmetry is therefore that inherent case makes it 'easy' to build the relevant presuppositions. For wh-phrases with inherent case, building the presuppositions is effortless, while building the relevant presuppositions for other wh-phrases requires elaborate discourse settings such as (236) (and once the adequate context is found, it is sometimes reported to be 'hard to keep in mind', speakers will occasionally complain that it 'slips away' and need to make efforts to 'keep it around' while giving the judgments.)

This generalisation fits well with the fact that inherent case maps onto presuppositions: the general picture emerging is that (i) eWI requires presuppositionality on the wh-phrase, (ii) presuppositionality is facilitated by iCase, but (iii) it can be forced by an adequate discourse setting independently of case. Descriptively, what needs to be derived is:

(237)   a.   presuppositional force must be licensed
         b.   licensing takes one of two forms: inherent case or legitimation
               through previous discourse

From this it follows that uncased adverbials must ask some help from discourse in order to do eWI, while cased DPs don't. Hence the 'argument/adjunct' asymmetry.

To ease the discussion, I will refer to a licensed θ as θ_L. In these terms, the generalisation revealed by the 'argument/adjunct' asymmetry is that not all θ can do eWI (contrary to what was assumed above). Rephrasing the description:

(238)   a.   eWI requires θ_L
         b. θ_L is licensed either by inherent case or by discourse

Implementation

Since licensed θ are a subset of θ, the syntactic feature tree is now:

(239)

$$
\begin{array}{ccc}
\text{Quantifier} & \text{M} & \text{A } [\phi/\text{sCase}] \\
| & | & \\
\text{SpecificQ} & \theta & \\
& | & \\
& \theta_L &
\end{array}
$$
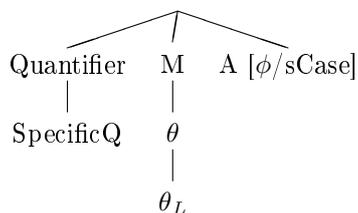
This entails that eWI-escaping θ-movement is to be rephrased as θ_L-movement, and therefore SQ-movement and θ_L-movement allow eWI, while A-movement, M-movement and θ-movement are clause-bound.

In turn, this entails that strong islands are not θ-interveners, but rather θ_L-interveners (but see McCawley (1988), Cinque (1990), Postal (1998) for some cases of wh-phrases plausibly extracting from strong islands, maybe in a θ_L-across-θ configuration). In more familiar terms, this rephrasing entails that strong islands are environments in which inherent case intervenes.

At first sight, this does not seem right: adverbial clauses are prototypical strong islands, but they do not bear inherent case. Notice however that within the strong island itself inherent case is assigned. Take typical adjunct islands:

(240)   a. What did you eat pasta [while [Mary was watching t]]?
         b. What did you eat pasta [during [the retransmission of t]]?

The temporal adjuncts *while Mary was watching t* and *during the retransmission of t* do not receive inherent case, but internally the complements *Mary was watching t* and *the retransmission of t* do receive inherent case. This is overtly visible with nominal adjuncts in cased languages such as Slovak:

(241)   a.   Vezmi si pass   pred   odchod-om
                take   self passport before leaving-case
                take your passport before leaving
         b. * Kam si vzal pass   pred   odchod-om t
                where self took passport before leaving-case
                target: *where did he take his passport before leaving for <where>

There is thus an intervening θ_L and the adverbial phrase is correctly predicted to be a strong island.

(An interesting side-effect of this refinement is that it allows movement of the entire complement of the adverbial (i.e. *Mary was watching x* and *the retransmission of* x). This is a welcome consequence as these cases do occur, in the guise of 'preposition-stranding'. Although the general form of the present approach is thus theoretically compatible with the existence of preposition stranding, I will not dwell here on the syntax of P-stranding, or head-trace phenomena in general.)

Similarly for relative clauses: as was discussed around (223), relative clauses obligatorily involve an argumental operator, and thus θ_L. As such, they will block both θ- and θ_L-movement out of them.[22] Let us thus assume that this logic generalizes and that strong islands are indeed environments with θ_L interveners.

In the updated terminology, the cases we have been looking at so far come out as:

- overt wh-movement is pied-piping of the wh-feature by θ (and therefore θ_L in cases where θ is licensed)

- strong islands are therefore cases of θ_L-across-θ_L and θ-across-θ_L

- situ-wh is only sensitive to weak islands because no θ feature is involved

- similarly, outcast cannot move overly as no θ is present and therefore no pied-piping is possible

- the special property of standard German tensed complement is that they receive inherent case and thus θ_L (and similarly for clausal subjects across languages)

- high *tout*-movement is (successive cyclic) θ_L-movement

- anaphoric reconstruction of wh-phrases is reconstruction into a θ_L intermediate trace

- weak islands are cases of Q-across-Q

- extraction out of weak islands is either θ-across-Q or SQ-across-Q

- the definiteness island is an SQ-across-SQ violation

---

[22]It is now tempting to relate the fact that English is more liberal in allowing some cases of apparent adjunct operators in relative clauses, (1b) and for many speakers (1d), to McCawley's observation that some extraction out of relative clauses are less degraded in English, (3):

(1) a.   the place where you went
     b.   the moment when you left
     c. * the way how you behaved
     d. % the reason why you left

(2)   ? what company does John have a brother who works for?

The two would connect if English has the special property of allowing relative operators to be θ rather than θ_L in some contexts.
The contrast between the English (3) and its ungrammatical French counterpart:

(3)   * pour quelle entreprise est-ce que Jean a un frère qui travaille?

follows from the fact that French requires relative operators to be θ_L, as shown by the ungrammaticality of θ relative operators in (223) above.

Given this rephrasing, what we need to understand is why θ-movement is unable to extract out of weak islands: non-presuppositional adjuncts, which bear θ but not θ_L cannot profit from θ-movement to escape weak islands.

θ-roles and base-positions

To see where this comes from, recall our literal reading of the notion 'structural case': structural case involves purely configurational semantic interpretation, with no θ feature involved at all. This reading however entails that DPs with structural case only bear structural case when they have reached appropriate structural configuration.

Only preverbal subjects will for instance be θ-less: a postverbal subject analysed as a subject-in-VP is not in the appropriate structural configuration and will therefore carry a θ-role. More generally, DP in their base-positions bear a θ-role regardless, and lose it if they move towards an sCase licenser (technically, this state of affairs can be made to follow from approaches which view checking as deletion of features: structural case features (θ) are deleted once moved, but not before moving. Furthermore, inherent case does not do checking and thus never loses its θ-feature. While appealing and compatible with everything up to now, this implementation is not compatible with the trigger for movement developed in §9).

In this context, the frequent observation that complement clauses occupy a position lower than complement DPs (e.g. Pesetsky (1995)) provides a cue as to the clause-boundedness of θ-movement. The lower position of clauses, is for instance visible in German and related languages (under and SVO analysis) where a complement clause follows but a complement DP precedes non-V2 verbs:

(242) a. $[_{CP} \ldots$ DP$_{acc}$ V
       b. $[_{CP} \ldots$ V CP$_{acc}$

Given the above assumptions about structural case, the complement clause must bear θ, since it has not moved to the relevant structural configuration with a sCase licenser. Internally, the tensed clause thus looks like:

$[_{CP} \theta \ldots$ V wh

As such, a tensed clause will thus block extraction of a θ-bearer (but will not block extraction of θ_L). θ-movement is therefore clause-bound, as desired.

(Notice that adjuncts can still skip over a preverbal subject, the latter occupying a structural-case position and thus not bearing θ).

Appendix 1 - A Tentative Note about Range and Specificity

Consider the following pair:

(243) a.   qu'est-ce que t'as   pris en photo?
               what is it that you have taken in photo
               what have you photographed?

         b.   de quoi est-ce que t'as   pris une photo?
               of what is it that you have taken a   photo
               what have you taken a picture of?

The two extractions have distinct presuppositions: The first can be asked with minimal presuppositions, maybe just seeing you walk around with a camera. The second is not felicitous in such a context. As discussed in §3.5.2, it requires a context legitimating the presupposition that there is one or more particular object that you took pictures of, and the identity of these objects is requested. In short, the extraction out of a noun-phrase requires specificity.

The contrast seems to remain under negation, although the judgments become tenuous:

(244) a.   qu'est-ce que t'as   pas pris en photo?
               what is it that you have not taken in photo
               what haven't you photographed?

         b.   de quoi est-ce que t'as   pas pris de photo?
               of what is it that you have not taken a photo
               what haven't you taken a picture of?

The presuppositions of (244a) shift to requiring a preestablished range for the wh-phrase, as usual. Wh-extraction out a noun-phrase again requires a context establishing a specific entity (or specific entities) that were not photographed and whose identity is under question. Generally then, extraction out of noun-phrase complements seem to require specificity, not range.

The difference between the two cases appears to relate to the fact that objects of verbs receive inherent case while objects of nouns don't. This in turn entails a difference in the licensing of θ_L: the noun-phrase being a weak island, §3.5.2, extraction out of a noun phrase requires θ_L; no inherent case being present, θ_L must be legitimated through discourse. Movement of the verbal complement out of a weak island also requires θ_L but θ_L is now directly legitimated by inherent case.

The contrast in presuppositions thus seems to indicate that discourse-licensed θ_L supports only specificity, not range; while a syntactically licensed θ_L is interpretable either as specific or as a range.

The same seems true of adjuncts, although judgments are more delicate:

(245) a. qu'est-ce qu'il   faut   éviter de toucher?
           what is it that one should avoid to touch

       b. qu'est-ce qu'il   faut   pas toucher?
           what is it that one should not touch

(246) a. comment est-ce qu'il   faut   éviter de le toucher?
           how   is it   that one should avoid to it to-touch
           how should one avoid touching it <how>?

       b. comment est-ce qu'il   faut   pas le toucher?
           how   is it   that one should not it to-touch
           how shouldn't one touch it <how>?

With extraction of cased phrases, a range context is sufficient: both examples
in (245) might be uttered by an experienced house-sitter preventively asking
what objects of the house should not be touched, with no specific referent for
the wh-phrase presupposed. (246) on the other hand cannot be made felicitous
in any such context; making it felicitous involves presupposing that there is (at
least) one specific manner in which the entity should not be touched and asking
for the identity of that/those manner(s).

Appendix 2 - In Defense of the Pragmatics Route

Before moving on to some further properties of θ-movement, let us pause to
consider - and defend - this view of the argument-adjunct asymmetry, which
traces it down to two different modes of licensing presuppositions: pragmatic
licensing via discourse for adjuncts vs syntactic licensing via inherent case for
arguments.

First, let us discard the classical 'kitchen-sink' charge: i.e. because pragmatics
can do anything and its opposite, invoking pragmatics is not explanatory. Notice
that according to the generalisation above, not everything can be rescued by
pragmatics, outcasts cannot for instance:

(247) a. how come it is so expensive now?
       b. how come you wonder since when it is so expensive?

Contrary to other instances of uncased-wh extraction, no amount of help from
the context can save *how come* eWI: (247b) cannot be a question built on the
basis of (247a), it can at most relate to the root clause.

Second, if pragmatics is indeed a licenser for an eWI-presuppositions, acceptabil-
ity of eWI with uncased wh-phrases should vary with the pragmatic accessibility
of the specific referent for the uncased wh-phrase. In light of this, observe how
this (factive) weak island becomes more porous as the pragmatics become easier:

(248)   a.   how do you regret having slept?
         b.   how do you regret having behaved?

While (248a) is a typical degradation of a contextless uncased-wh in eWI, (248b) is markedly better. Why? The explanation seems to lie in the uninteresting fact that regretting to have behaved in such and such way is not uncommon and thus pragmatically available, while regretting to have slept in such and such way is pragmatically more far fetched (but can be made plausible through an adequate context). That is, contextual help is easier to reach in (248b) than in (248a).

(Notice again that the natural presupposition is specificity, not range, in this context.)

The approach pursued here is emphatically *not* taking rarity of some pragmatic contexts as a given and explaining the 'argument/adjunct' asymmetry on that would-be basis. We are *not* following the line of thought sometime suggested, that 'since it is rare to think about times/manners, contexts for felicitous adjuncts are rare' (itself a version of the bizarre claim that language is the way it is because the 'external' world somehow makes it 'necessary' to be this way). This would be getting things backwards, and it would also get the facts wrong: we surely spend a fair amount of our time and energy talking about causes for events around us, but reason-phrases are nevertheless the most difficult wh-phrases to extract out of weak islands. What the present approach is doing is rather the extreme opposite: suggesting that in virtue of their syntactic make-up, some phrases make the life of the pragmatic module easier, and thus the relevant contexts more accessible. Cased-phrases make it easy to build a context because their syntax provides for a 'presuppositionality' slot that pragmatics only needs to fill-up. Uncased phrases make it difficult to create the relevant contexts because syntax is not being helpful and pragmatics needs to do a lot more work. It is only within phrases that are treated equivalently by syntax that pragmatic facilitations such as (248) can be found.

Notice that if the licensing of anaphora by pragmatics vs case is what is at stake, the 'argument/adjunct' asymmetry reduces to the more general fact that DPs have (anaphoric) determiners while adverbials don't, and 'argument/adjunct' asymmetries should thus be found independently of islands, in fact independently of movement altogether. The GunGbe asymmetry discussed in (225) is an example of the first kind, and several brands of cased-DP/adverb asymmetries are instances of the latter.

6 Clause-internal θ-movement

In this final locality section, I would like to address some less well-known locality effects, centering around the first θ-step of wh-movement and its associated robin-effects. These paradigms are somewhat more tentative than the earlier section, judgments often being very delicate, and the movements explored being less well understood. Before doing so, let us however review the pieces of the locality puzzle, as we now have ten different locality effects following from the anti-identity constraint, in sometimes intricate ways.

As usual, the baseline is a radical - or literal - reading of *Relativised Minimality* (which remains unmodified): first in that Q-movement literally ignores everything except Q interveners (and can thus jump over strong islands); and second in that the class-membership logic is followed up in its subset/superset consequences.

As is, such a literal version directly derives the liberal distribution of situ-wh (taken to be pure Q-movement). It also follows from this that Q-interveners will cause difficulties for situ-wh, regardless of where the Q-intervener is placed above the situ-wh. Given the existence of an SQ subclass of Q, Q-interveners can be skipped at the cost of a specific reading of the situ-wh. Here assumption (237) comes into play: because presuppositions are only syntactically licensed by inherent case, switching to SQ-movement causes an 'argument/adjunct' asymmetry: situ-wh with inherent case make less requirements on discourse in order to undergo SQ-movement. It is thus 'easier' for an argumental situ-wh to survive under a Q-intervener. Finally, it also follows that a situ-wh will be entirely blocked in one environment: specific noun-phrases (provoking an SQ-across-SQ violation), thereby deriving the 'definiteness island'.

By contrast, overt wh-movement is a function of the syntax of θ-movement. The difference between overt and covert wh-movement comes from pied-piping: because of morphological integrity, overt wh-movement cannot be simple Q-movement but involves pied-piping of the whole phrase, a θP or θ_LP. On the one hand, this derives strong islands (including the Germanic-like indicative complements and clausal subjects), which are environments with intervening θ_L (thus triggering θ_L-across-θ_L or θ-across-θ_L violations); and outcasts, which cannot move as they do not bear θ. On the other hand, this forces legal cases of overt wh-movement to first make a θ_L/θ-movement step. This θ-step is then what underlies:

- the argument/adjunct asymmetry (which arises from the difference between the properties of θ-movement and θ_L-movement),

- the reconstruction paradigm (which arises from the position of the θ-phrase and from the fact that θ_L-movement does not block Q-movement across itself)

- high *tout*-movement (an instance of θ_L-movement)

- the difference between extracting a complement of a noun and a complement of a verb (since the former has not access to inherent case and thus to θ_L-movement)

A canonical instance of wh-movement would now look like this:

(249) $[_{CP}$ wh ... $[_{\theta P}$ t ... $[_{VP}$ V t

and simple extraction out of weak islands have the following general outlook:

(250) $[_{CP}$ wh$_1$ ... $[_{\theta P}$ t$_1$ ... $[_{CP}$ wh$_2$ ... $[_{\theta P}$ t$_1$ $[_{\theta P}$ t$_2$ ... t$_1$ t$_2$

What we haven't looked at in any detail, is the properties induced by the low step of θ-movement. First, (249) overlooks the fact that the wh-phrase might bear structural case and thus move by A-movement rather than θ-movement (since phrases in structural case position involve no θ, the pied-piping constraint has the effect of triggering A-movement). This has important consequences yet to be explored. Second, the simple eWI (250) involves a round-robin downstairs, again a fact with important consequences.

In order to simplify the discussion, let me make some stipulations at the outset, some of which will follow from the later discussion:

(251) a. θ- and θ_L-movement land in the same type of position, θP (i.e. there is no θ_LP)

  b. robin-configurations of multiple θ_L/θ phrases, the phrases must be homogeneous: either all are θ or all are θ_LP. (A temporary way to understand this is to consider the various θ-phrases as multiple specifiers of a single head, thereby required to homogeneously agree with that head. This will follow in a more natural way from the discussion in §9

  c. θ-type movement must be uniform, it cannot start with θ_L-movement and continue with θ-movement, or vice-versa

These assumptions considerably cut down on the number of possibilities for early-steps of wh-movement. In essence, a wh-phrase must proceed by θ-steps, and these steps must be homogeneous with other θ-phrases in the structure.

6.1 A Ban On Indefinite Interveners

The above system makes a prediction about subjects under wh-movement: they cannot be θ-phrases.

If a subject is a DP with a θ-projection, it will need to do a θ-step (or θ_L-step) in virtue of the pied-piping constraint:

(252) $[_{TP}$ SU$_\theta$ ... $[_{\theta P}$ t$_\theta$ ... $[$ t$_\theta$

This has no consequence in the general case, but has unexpected effects if a wh-phrase is extracted from such a configuration. To be extractable, the wh-phrase itself must be a θ-phrase, and thus make a θ-step. In doing so it will enter a round-robin configuration with the subject. The subject being structurally higher than any other argument/modifier, it will now constitute a robin-trap for the wh-phrase. As a consequence, the wh-phrase cannot do any further θ-movement, and will therefore be unable to travel the θ-based eWI route.

Given our earlier conclusion that there are two sets of θ-related positions in the clause, one above and one below the subject, both local wh-movement and long-distance wh-movement involve a θ-step past the subject, and thus both will be affected by this predicted restriction.

To illustrate, cases of local wh-movement over a θ-subject would be forced to look like this:

(253) $[_{CP}$ wh $[_{\theta P}$ t$_{wh}$ $[_{TP}$ SU ... $[_{\theta P}$ t$_{su}$ $[_{\theta P}$ t$_{wh}$ ...

A configuration which violates the round-robin (the wh-phrase has prolongated θ-movement while the c-commanding subject hasn't).

The violation occurs regardless of the choice of θ or θ_L: in virtue of (251b) both the subject's and the wh-phrase's θ step must be of the same type. Regardless of the choice of θ or θ_L, if the object prolongates the movement, it will cause a robin-violation.

As a consequence, any intervening θ-subject will block wh-movement. In order to intervene legally, a subject must be a θ-less DP, undergoing A-movement and not creating any θ-trap.

This prediction, although surprising at first sight, seems to match a little-noticed asymmetry in the syntax of subjects under wh: they cannot be (non-specific) indefinite. Contrast (254) with (255):

(254) a. * who did a policeman see?
       b. * qui est-ce qu'un policier a vu?

(255) a. who did this policeman see?
       b. qui est-ce que ce policier a vu?

(specific indefinites pattern with (255)).

The ban is not about indefinite subjects in general, but specifically about *preverbal* indefinite subjects: passivising the sentences and thus shifting the subject into a postverbal *by*-phrase yields acceptable sentences.

(256)   a.   who was seen by a policeman?
        b.   qui est-ce qui a été vu par un policier?

Finally, 'specificity' must be understood in the stronger sense here, as simply providing a range does not help:

(257)   a.   * among us five, who did a policeman see?
        b.   * parmi nous cinq, qui est-ce que un policier a vu?

The troublesome configuration thus involves wh-movement crossing a non-specific preverbal subject.

The match between the unexpected prediction, and the empirical restriction is suggestive: both state that under wh-movement, subjects loose an option. Putting the two phenomena together now seems straightforward, but entails assigning a special interpretation to A-movement.

Suppose that θ (movement) maps onto indefiniteness (more precisely, onto non-specificity), very much like SQ (movement) maps onto specificity. As a result, the empirical observation that subjects cannot be non-specific under wh-movement is the same fact as the theoretical observation that subjects cannot be θ-phrases under wh-movement: if the subject is non-specific, its movement will be θ-movement which will create a θ-trap. Only specific subjects can thus intervene under wh-movement.[23]

(Given that θ-movement maps onto non-specificity and A-movement doesn't, it is tempting to conclude that A-movement maps onto specificity (this mapping can then be seen as the non-quantificational counterpart of SQ). As a result, each movement is associated to a particular semantics: θ-movement is non-specific, θ_L-movement is associated to range, SQ-movement is a specific referent for a quantifier, A-movement is a specific DP, and Q-movement is pure quantification. This semantic partitioning of movements/chains is an important point I will come back to.)

Extending the paradigm to non-subjects seems to confirm the robin-based explanation:

(258)   a.   who did you give a book to?

---

[23]This general reasoning entails that a wh-movement can only fork into SQ-movement from the higher θP, or else the wh-phrase could escape the subject θ-trap through SQ-movement. I have no explanation for this restriciton yet.

Similarly, the wh-phrase in these constructions must not be allowed to do A-movement, or the robin-trap would be voided. A-movement might however be a possiblity in some cases, as the acceptability of the following non-specific subject under a wh-question illustrates:

(1)   how many accidents can a man create before getting arrested?

suggesting that under some conditions wh-phrases are able to travel on the A-movement road. See next section for more evidence to this effect.

b. ?* what did you give to a man?

Again, a DP is prohibited from being non-specific just in case it's base-position is structurally higher than that of a moved wh-phrase. An indefinite DAT is a θ-phrase and is structurally higher than the object, thus creating a robin-trap for a wh-movement of the object, (258b). An indefinite object on the other hand has no effect on a DAT wh-phrase, since it is structurally lower and the robin-effect thus allows DAT to move independently of the object, (258a).

Notice how the 'NOM > DAT > ACC' scale reappears: an indefinite subject cannot dominate any wh-phrase, an indefinite DAT cannot dominate a wh-object, but other orders are fine.

The same should hold of high *tout*-movement: being an instance of θ_L-movement, an intervening θ or θ_L subject will create a robin-violation. A correct prediction:

(259) il tout deviné comment $\left\{ \begin{array}{c} \text{Jean} \\ \text{cet homme} \\ * \text{un homme} \\ * \text{quelqu}'\text{un} \end{array} \right\}$ a résolu

The relevant part of the configuration with a non-specific subject is:

(260) $[_{\theta P}$ tout $[_{CP} \ldots [_{TP}$ SU $\ldots [_{\theta P}$ t$_{su}$ $[_{\theta P}$ t$_{tout} \ldots$

Which violates the robin-constraint. In a similar vein, consider the specificity contrast in dative shift:

(261) a. I gave a book to $\left\{ \begin{array}{c} \text{a man} \\ \text{John} \end{array} \right\}$

b. I gave $\left\{ \begin{array}{c} ?*\text{a man} \\ \text{John} \end{array} \right\}$ a book

This suggests that dative-shift is θ-movement, and thus always involves specific DPs; a conjecture that will be confirmed independently below.

A speculation on null subjects

It is tempting to relate this indefiniteness restriction to another quirk affecting subjects under wh-movement: in many null-subject languages, wh-movement disallows preverbal-subjects in root clauses (including specific subjects). Here are Italian and Slovak illustrations:

(262) a. * chi Gianni ha visto?
            who G   has seen

     b.   mi chiedo   chi Gianni abbia visto
           self I.wonder who G   have seen
           I wonder who Gianni saw

(263)   a. ?* koho Pepik videl?
          who P   seen
          who did Pepik see?

     b.   neviem   koho Pepik videl
           no.I.know who P   seen

In embedded clauses the situation is the same as the English/French situation described above: a preverbal subject under wh becomes possible, but not if it is a non-specific indefinite:

(264)   a.   mi chiedo   chi Gianni abbia visto
           self I.wonder who G   have seen
           I wonder who Gianni saw

     b. * mi chiedo   chi un poliziotto abbia visto
           self I.wonder who a   policeman have seen

(265)   a.   neviem   koho Pepik videl
           no.I.know who P   seen
           I don't know who P saw

     b. * neviem   koho nejaky policajt   videl
           no.I.know who a   policeman seen

And again, the constraint against non-specific holds only for preverbal subjects: passivising the structures makes them acceptable.

Why then is there a special restriction forbidden root preverbal subjects under wh-movement in null subject languages? This above explanation of the indefiniteness quirk makes a simple suggestion: null-subject languages force a root subject to travel by θ-movement rather than A-movement. If that were true, all root preverbal subjects would create a robin-trap for wh-movement, and the two would be mutually exclusive. Whether this is a workable approach to null subject languages remains to be investigated.

6.2 Extraction under Dative-shift

Another mystery which seems to receive a natural explanation with the robin-constraint is the fact that wh-movement of an object is impossible if dative-shift has taken place:

(266)   a.   what did you give <what> to Paul

    b.   * what did you give Paul <what>

Consider first (267):

(267)   a.   ? what do you wonder who gave to Paul
          b.   * what do you wonder who gave Paul

If dative-shift is θ-movement, the unacceptability of (267b) directly follows from the round-robin: in order to escape the weak island, *what* must do successive cyclic θ-movement. But doing so violates the robin, as the hierarchically superior dative has not prolongated θ-movement:

(268)   **what** ... $[_{\theta P}$ $\mathbf{t}_{what}$ $[_{CP}$ $\underline{who}_Q$ ... $[_{\theta P}$ Paul $[_{\theta P}$ $\mathbf{t}_{what}$ ... $[_{VP}$ t_{Paul} $\mathbf{t}_{what}$

At first sight this explanation does not carry over to (266b), since the following structure seems legal:

(269)   **what**$_Q$ ... $[_{\theta P}$ Paul $[_{\theta P}$ $\mathbf{t}_{what}$ ... $[_{VP}$ ... t_{Paul} $\mathbf{t}_{what}$

But if the above conclusion that there are two series of θ-positions is correct, the structure of (266b) is:

(270)   **what**$_Q$ ... $[_{\theta P}$ $\mathbf{t}_{what}$ ... $[_{\theta P}$ Paul $[_{\theta P}$ $\mathbf{t}_{what}$ ... $[_{VP}$ ... t_{Paul} $\mathbf{t}_{what}$

Which violates the round-robin again. Plausibly then, the contrast in both (266) and (267) follows from the fact that a dative-shifted phrase undergoes θ-movement and thus robin-blocks further wh-movement. This confirms the analysis of (261) above in terms of dative-shift being θ-movement. [24]

---

[24]An interesting twist is added by:

(1)   which portrait of herself did you explain to Simone why I threw away

The portrait-wh is capable of reconstructing below the dative *to Simone* (but still above *why*), a reconstruction site substantially lower than the expected reconstruction site θP. I will however assume that *to Simone* moves at LF to reach its structural-case position - an LF dative-shift - thereby positioning itself above the θP position of portrait-wh.

This movement cannot be θ-movement however, or else PP-datives would be identical at LF with shifted-datives and the contrasts (266-267) would be lost. It must therefore be the case that this LF-movement is A-movement. The reason why LF-shift is A-movement while overt shift is θ-movement remains mysterious, but recalls the contrast between covert wh-movement being Q-movement while overt wh-movement is θ-movement.

6.3 WH-movement and Agreement

French wh-movement of an object optionally triggers agreement on the past-participle (not for all speakers), and when it does, a presuppositional reading is generated (e.g. Obenauer (1994), Adger (1994), Dobrovie-Sorin (1990), Rizzi (2000), etc.):

(271) a. combien d'atrocités   ont-ils   commises?
           how many of atrocities$_{fem}$ have-they committed$_{fem}$

           quelles atrocités   ont-ils   commises?
           which atrocities$_{fem}$ have-they committed$_{fem}$

       b. combien/quelles d'atrocités ont-ils   commis?
           how many/which of atrocities have-they committed

Contexts such as (271a) make the presupposition obvious by triggering a surprising reading in which a set of potential atrocities is presupposed, with the question inquiring about which of those atrocities were actually performed.

The agreement effect is particularly clear in regiolects in which the participle agreement is systematically spelled-out: in some regions, including regions defining 'standard' French, the agreement is only audible on past participles ending with consonants. In other areas, the agreement is systematically audible: as a glide with participles in /e/ and /i/, and as a /œ/ after participles in /y/. In the latter areas, omitting this vowel-based agreement in obligatory-agreement constructions such as passives produces strong ungrammaticality:

(272) a. l'indépendance   a   été   déclaré[:j]
           the independence has been declared
       b. * l'indépendance a été déclaré

(273) a. l'indépendance   a   été   établ[i:j]
           the independence has been established
       b. * l'indépendance a été établi

(274) a. l'indépendance   a   été   obtenu[:œ]
           the independence has been obtained
       b. * l'indépendance a été obtenu

In these vowel-agreement varieties (to which Geneva-French belongs), the wh-agreement paradigm extends throughout wh-movement of objects:

(275) a. combien   d'atrocités   ont-ils   nié[:j]?
           how many of atrocities$_{fem}$ have-they denied$_{fem}$

b. combien d'atrocités   ont-ils subi[:j]?
      how many of atrocities$_{fem}$ have-they lived through$_{fem}$

  c. combien d'atrocités   ont-ils vu[:œ]?
      how many of atrocities$_{fem}$ have-they seen$_{fem}$

Following traditional analyses, let us suppose that agreement indicates A-movement. The structures corresponding to the agreeing versus non-agreeing examples then naturally come out as presence versus absence of A-movement of the structurally case-marked object:

(276) a. $[_{CP}$ wh ... $[_{\theta P}$ t$_{wh}$ ... $[_{AgrP}$ t$_{wh}$ ...
       b. $[_{CP}$ wh ... $[_{\theta P}$ t$_{wh}$ ...

Given the conclusion of §6.1, the fact that A-movement maps onto presuppositions but θ-movement does not is the expected outcome.

Similarly, the fact that agreement is not possible on the past participle of a higher clause falls out from the fact that A-movement is clause-bound:

(277) * combien d'atrocités   est-ce qu'il a dites   qu'ils   ont
         how many of atricities$_{fem}$ is it   that he has said$_{fem}$ that they have
         commis?
         committed

A less known fact also falls out: agreement is possible with an object-wh undergoing long distance movement, but not if the long-distance movement is eWI-movement:

(278) a. ? laquelle de ces voitures est-ce que tu crois qu'elle a conduite
             which one of these cars$_{fem}$ do you think that she drove$_{fem}$
       b. ?* laquelle de ces voitures est-ce que tu te demandes quand elle a
             conduite
             which one of these cars$_{fem}$ do you wonder when she drove$_{fem}$
       c. laquelle de ces voitures est-ce que tu te demandes quand elle a
             conduit
             which one of these cars$_{fem}$ do you wonder when she drove

Since eWI involves (successive cyclic) θ-movement rather than A-movement, this is again the expected outcome.

The legal long distance structure is the traditional successive-cyclic Q-movement:

(279) $[_{CP}$ wh ... $[_{CP}$ t$_{wh}$ ... $[_{AgrP}$ t$_{wh}$ ... t$_{wh}$ ...

But no corresponding eWI structure exists: the presence of agreement ensures that the weak island cannot be escaped by θ_L-movement since A-movement and any brand of θ-movement are mutually exclusive. Similarly, according to the reasoning above SQ is the quantificational counterpart of A-movement, and is thus mutually exclusive with A-movement (a phrase cannot be both quantificational and non-quantificational). If so, both the θ_L-road and the SQ-road are blocked by an initial step of A-movement.

(This raises an interesting point: the θ-step of a wh-movement is necessary in the originating clause of the wh-movement, but not in higher clauses crossed by the movement. Put differently, θP must project in the clause where the movement originates, but not necessarily in the higher clauses.)

Agreement is however possible with the eWI-inducer itself, as expected:

(280)   a.   à quels invités est-ce que tu sais quelles maison ils ont montré[:j]?
              to which guests do you know which houses$_{fem}$ they have shown$_{fem}$

Finally, a puzzling fact is that subject and object agreement have different effects: wh-movement of an agreeing subject does not trigger the same presuppositional reading that an agreeing object does.

(281)   combien de personnes ont été admises aux urgences?
           how many persons were admitted to the urgency-room?

One view is to see this as indicating that subject-agreement intrinsically differs from object-agreement (Rizzi (2000)). Another possible interpretation is that optional agreement is different from obligatory agreement (subject agreement is obligatory in (281), while object wh-agreement is optional).

There is some evidence for the latter interpretation: some verbs allow optional agreement in object-to-subject constructions, and in these cases the presuppositional reading patterns with agreement again:

(282)   les femmes se sont dispersé / dispersé[:j]
           the women self are   dispersed / dispersed$_{fem}$

In (282), a subtle difference is perceptible between the agreeing and non-agreeing versions: the agreeing version favors a reading in which the women have been mentioned previously, as opposed to a group-like (quantity) reading when agreement is absent. This paradigm is very elusive as few verbs allow optionality of agreement, with no clear pattern; and the same verb behaves differently in different contexts. Here are some more examples:

(283)   a.   une flaque d'eau s'est ?répandu / répandu[:œ] dans le
             a   water puddle self is dispersed / dispersed$_{fem}$ in   the
             couloir
             corridor

    b.   une candidate s'est présenté / présenté[:j]
          a   candidate self is presented / presented$_{fem}$

    c.   une catastrophe s'est produit   / produite   en montagne
          a   catastrophy self is produced / produced$_{fem}$ in mountain

    d.   Marie s'est saoulé   / saoulé[:j]
          Mary self is drunken / drunken$_{fem}$
          Mary got drunk

This light asymmetry gives us evidence that presuppositional agreement is optional-agreement, rather than object-agreement. Again, this seems characterizable in θ versus A terms: verbs allowing optional agreement allow both θ- and A-movement, with the latter triggering presuppositional readings as above.[25]

(Associating A-movement with specificity leads to a situation reminiscent of the elsewhere-principle (and therefore of economy-like principles): verbs which do not have optional agreement but rather require it, do not force presuppositionality on their subjects. In those cases, A-movement is thus allowed to map onto both specific and non-specific readings. The generalization would thus be that choosing A-movement over the non-specific θ-movement results in specificity, while applying A-movement where no choice is present leaves the semantics free. The elsewhere/economy situation is however a property of semantics in this case, as it involves the procedure to interpret structures. This is in turn reminiscent of Sperber and Wilson (1986)'s *Relevance* theory, which contends that the core of interpretation involves choice-based effects.)

6.4 A-movement and Intervener-effects

Coming back to a paradigm left open in §2, what underlies (6.4)?

(284)   a.   what is it unclear whether we should repair?
        b.   what is it unclear how we should repair?

And why does this difference disappear with infinitival complements (while the double-superiority effect does not disappear with infinitival complements):

(285)   a.   what is it unclear whether to repair?
           to whom is it unclear what to give?
        b.   what is it unclear how to repair?

---

[25]The fact that some verbs allow both kinds of movement while other restrict their subject to only one type lends some plausibility to the above speculation that null subject languages have the property of restricting their root subjects to only one type of movement, θ-movement.

(Here judgments are subtle but remarkably stable across languages: English, French, Hebrew, Slovak, Southern German for instance all have the same asymmetry.)

Assumption (251b) about the homogeneity of θ-movement is central again: the intervener in (b) is an uncased adverbial (an 'adjunct') and its first step is thus θ-movement, not θ_L-movement. As a result, the extracted *what* cannot do θ_L-movement and thus cannot extract via a θ-based route. Both of the following are illegal:

(286)   a. $[_{CP}$ wh$_1$ ... $[_{\theta P}$ t$_{\theta_L 1}$ ... $[_{CP}$ how$_2$ ... $[_{\theta LP}$ t$_{\theta_L 1}$ $[_{\theta P}$ t$_2$ ... t$_1$ t$_2$
       b. $[_{CP}$ wh$_1$ ... $[_{\theta P}$ t$_1$ ... $[_{CP}$ how$_2$ ... $[_{\theta P}$ t$_1$ $[_{\theta P}$ t$_2$ ... t$_1$ t$_2$

The first because it mixes θ with θ_L, the second because θ-movement is clause-bound. The only way for the argument to skip the adjunct is thus to fork to the SQ road:

(287) $[_{CP}$ wh$_{\theta/S}$ ... $[_{SqP}$ t$_{\theta/S}$ $[_{CP}$ wh$_\theta$ ... $[_{\theta P}$ t$_{\theta/S}$ $[_{\theta P}$ t$_\theta$ ...

If an interrogative complementiser intervenes on the other hand, θ_L becomes an option:

(288) $[_{CP}$ wh ... $[_{\theta P}$ t ... $[_{CP}$ if ... $[_{\theta P}$ t ... t

In other words, the presence of an adjunct forces the robin to be of the θ-type, and thus blocks the θ_L road. By blocking θ_L-movement, it forces SQ-based eWI which has stronger contextual requirements and thus provokes a stronger degradation than range-movement, in out of the blue contexts. This now explains (6.4).

Why does this adjunct-intervention effect disappear in infinitivals, (285)? Because infinitivals allow θ-movement out of them.

Several paradigms show that complement infinitivals have the option of moving higher than complement tensed clauses, plausibly to the same position as object DPs. If so, they occupy a structural case position in which they need not bear θ. As a consequence, θ-movement can extract from the infinitival, and the following derivation is legal:

(289) $[_{CP}$ wh$_1$ ... $[_{\theta P}$ t$_{\theta 1}$ ... $[_{CP}$ wh$_2$ ... $[_{\theta P}$ t$_{\theta 1}$ $[_{\theta P}$ t$_{\theta 2}$ ...

(As a confirmation, notice that those Germanic regiolects in which tensed complements are strong islands allow extraction out of infinitival complements. This follows from the present view of infinitivals as not involving θ).

The fact that θ-movement can extract from infinitivals doesn't however help with double-superiority: the robin constraint still prevents the lower member of the θ-robin from prolonging the movement independently on its own. The asymmetry between double-superiority and type-of-intervener effects is thus correctly derived.

Conclusion

It thus seems that we have a reasonable case to the effect that the various pieces of locality stem from a unique underlying locality principle: the anti-identity constraint (*Relativised Minimality*, *Attract Closest*, etc.). Not only does this allow us to unify the major pieces of locality, but it also guides us into a series of less well known and more involved paradigms, apparently giving us a productive tool for the job.

Having such a unified notion of locality allows us to ask deeper questions about how locality (or rather the 'move' / 'chain' operation which gives rise to it) relates to other tools in the syntactic toolbox. Before going there, let us however show that many paradigms which have been thought to provide evidence about the underlying nature of eWI in fact reduce to its presuppositional nature.

6.5 Appendix - Double Superiority

Judgments being extremely delicate in this construction and their interpretation often extremely difficult (and in need of stabilisation), let us examine the relative order of extraction of two wh-phrases in a very tentative vein. Call this the 'double-superiority' paradigm:

(290)   a.   à qui   est-ce que tu   voulais savoir   de quoi il   faut
             to whom is it   that you wanted to-know of   what one should
             éviter de parler?
             avoid to speak
             to whom did you want to know what one should avoid talking about?

       b.   de   quoi est-ce que tu   voulais savoir   à qui   il
             about what is it   that you wanted to-know to whom one
             faut   éviter de parler?
             should avoid to speak
             about what did you want to know who one should avoid talking to?

Both extractions in (290) are standard eWI, but (290b) seems slightly more degraded than (290a). Why?

There are two facts to notice about (290) in order to see the generalisation: first, the dispreferred ordering (290b) is in fact possible if the extracted wh-phrase is made strongly specific. In other words, 'about what' extracted over 'to whom' behaves somewhat like adjunct eWI; whereas extracting 'to whom' over 'about what' behaves like standard argument extraction. Second, systematically comparing pairs of wh-phrases in the double-superiority constructions produces a 'preferred' ordering which corresponds to the ordering of situ-wh phrases and to the round-robin ordering:

(291) LOC > INSTR

    a.   + dans quelle salle il n'est pas clair avec quelle clé il a réussi à entrer?

         into which room is it unclear with which key he managed to enter?

    b.   - avec quelle clé il n'est pas clair dans quelle salle il a réussi à entrer?

         to whom did you want to know what he wants to present?

(292) ACC > LOC

    a.   + qu'est-ce que tu voulais savoir où il va présenter?

         what did you want to know where he will present?

    b.   (-) où est-ce que tu voulais savoir qu'est-ce qu'il doit présenter?

         where did you want to know what he will present?

(293) DAT > ACC

    a.   + à qui est-ce que tu voulais savoir qu'est-ce qu'il va présenter?

         to whom did you want to know what he will present?

    b.   (-) qu'est-ce que tu voulais savoir à qui il va présenter?

         what did you want to know to whom he will present?

(294) DAT > LOC

    a.   + à qui est-ce que tu voulais savoir où il a parlé?

         to whom did you want to know wher he spoke?

    b.   - où est-ce que tu voulais savoir à qui il a parlé?

         where did you want to know to whom he spoke?

The preferred order is thus 'DAT > ACC > LOC > INSTR', with the semantic asymmetry described above: the preferred order allows a 'list-like' reading within a predefined range, whereas the inverted order requires enquiring about the identity of a specific entity (the asymmetry is somewhat milder in paradigms involving ACC, (292-293). I come back to this asymmetry immediately below). Other pairs confirm this, e.g.:

(295) DAT > LOC

    a.   + à qui est-ce que tu voulais savoir dans quelle salle il doit parler?

         to whom did you want to know in which room he is supposed to speak?

    b.   - dans quelle salle est-ce que tu voulais savoir à qui il doit parler?

         in which room did you want to know to whom he is supposed to speak?

The ordering of subjects within the sequence is unclear due to the independent difficulties with their extraction, including ECM contexts:

(296)   a.   who is it unclear what they caught t stealing t?

         b.   what is it unclear who they caught t stealing t?

(297)   a.   who did you want to know to whom they saw t speak t?

         b.   to whom did you want to know who they saw t speak t?

Both in French and in English the judgment is difficult, but speakers seem to prefer the 'subject>object' and 'subject>dative' order when they can parse the sentences at all.

Similarly, in null-subject languages allowing *that-t* configurations, the 'NOM > DAT' and 'NOM > ACC' orders are preferred:

(298)   a.   + kto   nevies   komu   hovoril?
               who   you.not.know   who$_{dat}$ talked
               who$_i$ don't you know to whom he$_i$ talked

         b.   - komu   nevies   kto   hovoril?
               who$_{dat}$ you.not.know   who talked

(299)   a.   + kto   nevies   co   povedal?
               who   you.not.know   what said

         b.   - co   nevies   kto povedal?
               what you.not.know   who said

The ordering giving rise to simple eWI-effects is thus 'NOM > DAT > ACC > LOC > INSTR'. This ordering is the same as the round-robin ordering (where testable), i.e. 'NOM > DAT > ACC'.

It is also the same ordering given by situ-wh, modulo one complication:

(300)   a.   * t'as   été   avec cette voiture ou   ?
               you have been with this   car   where

         b.   ? t'as   été   en vacance comment?
               you have been on holiday   how

Given that wh-phrases cannot be dislocated, the fact that an instrumental-wh can follow a locative modifier indicates the availability of LOC > INSTR. Since the locative-wh may not follow an instrumental, this is the sole available order (i.e. locatives need to be dislocated in order to follow instrumentals).

This technique of inferring the underlying order of situ-wh is more complicated than simply putting several situ-wh together, but produces more reliable results as the latter often yields unclear judgments. Also, putting a non-interrogative PP after the wh is always possible due to right-dislocation, and is thus irrelevant.

Similarly:

(301) a. t'as   parlé   à   Jean dans quelle salle?
          you have spoken with John in   which room

      b. *t'as   parlé   dans cette salle à   qui?
          you have spoken in   this   room with whom

and therefore 'DAT > LOC > INSTR'.

The complication comes from the pairing of an accusative with a dative, which yields 'ACC > DAT' instead of 'DAT > ACC':

(302) a. ?? t'as   montré à   Jean quelle chaise?
          you have shown to John which chair

      b. t'as   montré cette chaise à   qui?
          you have shown this   chair to whom

This however correlates with the fact that the accusative is a DP while the dative is a PP. I will thus assume that the unexpected 'ACC > DAT' order is a consequence of the DP moving overtly to a position to which the PP does not move (a difference parallel to the fact that accusative DPs occupy a position higher than accusative CPs, as observed above. See also McCloskey (2000) on the presence of such a short object movement in English, and Ordonez (1998) for what is possibly the same movement in Spanish).

Modulo this complication, situ-wh, the round-robin and double superiority all point to the same base order. Why should this hold in the case of double-superiority? And why does the inverse order require specificity?

The core logic of this effect follows from the round-robin constraint. Consider the situation where two wh-phrases with inherent case are to be extracted:

(303) $[_{CP}$ ... $[_{CP}$ ... wh_1 wh_2

In virtue of the pied-piping constraint, both must do a θ-step - in fact an θ_L-step given the presence of inherent case - and this step must preserve ordering:

(304) $[_{CP}$ ... $[_{CP}$ ... $[_{\theta P}$ wh_1 $[_{\theta P}$ wh_2 ... t_1 t_2

Now an asymmetry arises: because of the round-robin, wh_1 can proceed with θ_L-movement independently of wh_2, while wh_2 cannot proceed without wh_1:

(305) a. $[_{CP}$ ... $[_{\theta P}$ wh_1 ... $[_{CP}$ ... $[_{\theta P}$ t_1 $[_{\theta P}$ wh_2 ... t_1 t_2
      b. * $[_{CP}$ ... $[_{\theta P}$ wh_2 ... $[_{CP}$ ... $[_{\theta P}$ wh_1 $[_{\theta P}$ t_2 ... t_1 t_2

This means that wh_2 can move into the lower CP by Q-movement, with wh_1 skipping it by $i\theta$-movement into the higher clause. This is the canonical range-based escape from a weak-island illustrated above:

(306)　$[_{CP}$ wh$_1$ ... $[_{\theta P}$ t$_1$ ... $[_{CP}$ wh$_2$ ... $[_{\theta P}$ t$_1$ $[_{\theta P}$ t$_2$ ... t$_1$ t$_2$

If however wh$_1$ moves to the embedded CP, wh$_2$ is trapped. Because of the round robin it cannot move into the higher clause by $i\theta$-movement; the following is illegal:

(307)　*$[_{CP}$ wh$_2$ ... $[_{\theta P}$ t$_2$ ... $[_{CP}$ wh$_1$ ... $[_{\theta P}$ t$_1$ $[_{\theta P}$ t$_2$ ... t$_1$ t$_2$

Supposing it is possible to fork to SQ-movement at that stage, the only way for wh$_2$ to reach the root clause and still comply with the round-robin is to drop out of the robin by forking into SQ-movement:

(308)　$[_{CP}$ wh$_2$ ... $[_{SqP}$ t$_2$ ... $[_{CP}$ wh$_1$ ... $[_{\theta P}$ t$_1$ $[_{\theta P}$ t$_2$ ... t$_1$ t$_2$

Doing SQ-movement of course entails that wh$_2$ is specific, thus deriving the observed effect of inversion. It thus generally follows that inverting the order of two wh-phrases is only possible if the invertee travels on the SQ-road.

Under this analysis, the root of the double-superiority paradigm is the round-robin: a lower wh-phrase is 'trapped' by the higher ones, unless it forks onto the SQ-road. This is an example of the logic illustrated in (192).

One striking aspect of the double-superiority paradigm - to the extent that judgments are to be trusted - is that it is the reverse of what simple superiority would lead us to expect: simple superiority would require the highest wh-phrase to land in the embedded CP, thus leaving only the lower wh-phrase for eWI. Of course, the reason for this reversal is clear: in cases of simple superiority, the robin is irrelevant, and only the highest wh-phrase is attracted. It is only when attempting an additional eWI that the 'trap' effect appears and causes a reversal of preferred ordering.

What of the fact that ACC-interveners yield lighter or no double-superiority effect (292-293)? ACC being a structural case, it has the option of not bearing θ and proceeding through A-movement. If it takes advantage of that option, the derivations become:

(309)　a. $[_{CP}$ wh$_1$ ... $[_{\theta P}$ t$_1$ ... $[_{CP}$ wh$_2$ ... $[_{\theta P}$ t$_1$ $[_{AP}$ t$_2$ ... t$_1$ t$_2$
　　　 b. $[_{CP}$ wh$_2$ ... $[_{\theta P}$ t$_2$ ... $[_{CP}$ wh$_1$ ... $[_{\theta P}$ t$_2$ $[_{AP}$ t$_1$ ... t$_1$ t$_2$

with no round-robin violation in either case, thus explaining the different status of sCase-interveners (the respective ordering of the A-position and the θ-position irrelevant in these structures).

This naturally leads to another oft-discussed paradigm: the fact that sentences involving two *which*-phrases lead to lesser or no simple superiority effects (e.g. Pesetsky (1987)). (Again, judgments are somewhat unstable on this construction, with high variation across speakers. I will follow the literature in idealizing the effect and supposing that superiority *is* alleviated).

(310) a. tell me which corporation sponsored which senator
       b. tell me which senator which corporation sponsored

Since *which*-phrases are also phrases which yield more acceptable eWI and optimal eWI is θ-movement, the obvious redescription of these facts is that simple superiority effects are alleviated (or absent) when two θ-movements are involved. This translates into the fact that simple superiority fades out when a round-robin is involved. If so, a natural explanation suggests itself, in the general spirit of 'equidistance': if the round-robin involves a configuration similar to 'multiple specifiers' (see §8.3 for a different and more precise characterization), then extracting from a round-robin is the only case in which the two attractees are in the same functional projection, and thus plausibly equally distant from the landing site; hence the alleviation.

A Systematic View

Since the derivations are involved and five different routes are available (A, Q, SQ, θ, θ_L), let us make a systematic review of the possibilities. To cut down the number of possibilities, remember that (i) A-movement and θ/θ_L-movement are mutually exclusive - a given wh-phrase either bears theta or it doesn't - and (ii) both A-movement and θ-movement are clause-bound and thus useless for eWI (infinitivals aside).

Restricting ourselves to two wh-phrases, and allowing the wh-phrases to be iCased arguments, sCased arguments or adjuncts, 18 possible base combinations arise (ordered and differentiating the extractee). Of these, 12 involve eWI-extraction of either A or θ and are thus illegal. The six remaining base-orders are (with the eWI extractee underlined): '$\underline{\theta_L}$ A', '$\underline{\theta_L}$ θ', '$\underline{\theta_L}$ θ_L', 'A $\underline{\theta_L}$', 'θ $\underline{\theta_L}$', 'θ_L $\underline{\theta_L}$'. Let us explore them in turn.

First, consider the pairs involving two θ_L-phrases: '$\underline{\theta_L}$ θ_L' and 'θ_L $\underline{\theta_L}$'. These correspond to canonical cases of double-superiority such as (290) and their derivations are thus (306) and (308). In short, the former can proceed with θ_L-movement but the round-robin forces the latter to fork to SQ-movement.

The two pairs mixing θ_L and θ ('$\underline{\theta_L}$ θ' and 'θ $\underline{\theta_L}$') fall prey to the restriction against mixing the two within a single robin, and thus force one of the two to fork to SQ-movement, which corresponds to the fact that adjunct-intervention is degraded over argument-intervention or intervention of an interrogative complementiser.

Finally, the pairs involving A-movement, '$\underline{\theta_L}$ A' and 'A $\underline{\theta_L}$' correspond to the two cases in (293). As noted, the double-superiority effect is lighter in these cases, which derives from the fact that A-movement of either wh-phrase does not robin-block the other, (309). (Although the residue of a double-superiority effect, if reliable, is unexplained.)

7   Interlude: Some Interpretive Illusions

Over the past few decades, a number of interpretive contrasts linked to eWI have been uncovered, and it is natural to think that these contrasts reveal the underlying nature of eWI. It turns out however that these contrasts are artifacts of the (independent) presuppositional nature of eWI: as a general description, the presuppositional reading makes some meanings less reachable, independently of eWI. Observing that these meanings are not reachable in eWI contexts then tells us nothing about eWI beyond the fact that it is presuppositional. The asymmetries to be discussed are:

- eWI disallows 'new guess' readings §7.1
- eWI is favored by providing explicit values for the wh-phrase, §7.2
- eWI disallows quantity readings, §7.3
- eWI disallows event readings, §7.4
- eWI disallows unicity readings, §7.5
- eWI disallows group readings, §7.6

7.1   The Prior Belief Asymmetry

For completeness, let us include the 'prior belief' asymmetry, although it is to my knowledge not discussed in the literature. One might hold that an additional semantic/pragmatic restriction on eWI is that eWI questions with psychological predicates can only be used to tap prior beliefs of the interlocutor, and can never be used as requests for guesses or on-the-spot belief-formation:

(311)   I know that you wonder what John cooked and that you have no clue about it, but I'd like your guess:
    a.   what do you think that John cooked?
    b. # what do you wonder whether John cooked?

A slightly more elaborate method of bringing out the same point:

(312)   a.   *Bill*: what do you think that John cooked?
        John: well, actually, I'm sorry but you have the wrong idea here; I never thought about that and I have no opinion on the matter.
        Bill: I *know* that, but I wanted to have your guess.
    b.   *Bill*: what do you wonder why John cooked?
        John: well, actually, I'm sorry but you have the wrong idea here; I never thought about that and I have no opinion on the matter.

Bill: # I *know* that, but I wanted to have your guess.

Again, if a question is misunderstood as a request for prior beliefs, an unhindered-movement question can be corrected to a 'new guess' interpretation, but an eWI cannot.

Such paradigms follow from the fact that eWI requires the extracted wh-phrase to carry existential presupposition: the very meaning of an existential with wide scope over a psychological predicate is that 'there is something that is believed'. A question with such a presupposition simply entails prior beliefs.

7.2   The Checkbook Asymmetry

The fact that extraction of uncased wh-phrases can be rescued has long been noted with respects to contexts such as (313), (Kroch (1989), Szabolcsi and Zwarts (1997)):

(313)   a.   how does John wonder whether to cook today: with his new vapor-stove, or with my stone-oven?

        b.   [pointing to three pictures] given these three possibilities, tell me: how didn't John know whether to behave?

Is the checklist-improvement independent of the need for a wide-scope $\exists$? Is it enough that the wh-phrase has a referent, independently of scope? Notice the contrast between (313) and:

(314)   a.   how do you think that John should cook today: with his new vapor-stove, or with my stone-oven?

        b.   They wonder how John should cook today: with his new vapor-stove, or with my stone-oven?

In both (313) and (314) a specific pair of individuals are available as referents, but the presuppositions nevertheless differ: in (313) both examples are questions about John's prior beliefs (or 'wonders'), and cannot be understood as requests for a guess/new-opinion. The (indirect) questions (314) on the other hand do not presuppose any prior opinion and can be felicitously used in novel-opinion contexts. In short, the presence of a checklist-context does not obviate the need for a wide scope existential in the eWI wh-phrase.

Checklists in fact provide a new and elegant way to see the contrast between Q-movement and presuppositional movement:

(315)   a.   what do you think that John should buy: a piano or a guitar?

        b.   what do you wonder when John should buy: a piano or a guitar?

The indefinites in the checklist are interpreted differently in (315a) and (315b). The indefinites in (315a) can be interpreted as non-specific, i.e. 'some piano/guitar or other' whereas in (315b) the indefinites become specific. In other words, the indefinite is linked to a wide-scope existential in eWI, (315b) but not in Q-movement, (315a).

A striking instance of this contrast is the following asymmetry in contradictions:

(316) I know you don't believe in unicorns and never think about them. I also know there are some things that you are wondering whether John will draw, so I ask:
    a.   what do you hope that John will draw: a unicorn or an elephant?
    b. # what do you wonder whether John will draw: a unicorn or an elephant?

The question (316b) is contradictory with the prior assumption that you never think about unicorns, whereas (316a) is not. Given the wide-scope requirement, the reason is obvious: (316b) presupposes that 'there is an object on your mind which could be a unicorn or an elephant, such that...' which is contradictory with that assumption. (316a) on the other hand has no such presupposition.

In sum, checklists facilitate eWI only to the extent that they facilitate a wide-scope $\exists$.

7.3   The 'Quantity Reading' Asymmetry

Cases such as (317) are apparently simple examples of the eWI contrast: (317b) requires ascription of beliefs about specific entities (e.g. people being invited by John), while (317a) does not.

(317)   a.   how many people do you think (that) John invited?
            b.   how many people do you wonder why John invited?

The traditional phrasing of the observation, is that (317a) has a 'pure quantity' reading whereby the answer can be simply a number, with no link to any actual people invited (as in 'seven, but I have no clue who: I only heard John saying that he wants to invite 7 persons'), whereas (317b) does not have this reading. It only has a reading whereby you have a specific set of people in mind, and the answer is the cardinality of that set (as in 'there are seven persons that I wonder about, and they are: Doodle, Google, Poodle, Diddle, Moodle, Suddle and Boodle').

Does the quantity-asymmetry reduce to the existential presupposition of eWI? Apparently not. Part of it does: the fact that (317a) has a reading which does not involve any beliefs about specific entities, while (317b) does not have such

a reading follows. But the number/non-number contrast does not follow: as far as eWI is concerned, the specific entities ascribed to the interlocutor's beliefs could as well be numbers. There is no reason why (317b) shouldn't be able to mean 'what is the number N that you have in mind such that you wonder why John invited N-many people'.

But the claim is that it doesn't have this reading. Why? As Szabolcsi and Zwarts (1997) note, the fact is that this reading does exist in eWI. It only requires some context in order to pop up. Here is one:

(318)   you have found your grand-mother's old diary, where she documents her young partying years. She happens to be somewhat of a scribble-maniac, and writes down the people she invites to her parties, together with the reason why she wants to invite them. Unfortunately, time has somewhat destroyed the diary, and there is a party for which the names and the reasons are unreadable, although it is easy for you to tell how many name/reason pairs are out of reach. Given that I know all of this, but I don't know what the number is, I ask you any of:

   a.   how many people is it unclear why grandma invited to that party?

   b.   how many people do you wonder why grandma invited to that party?

   c.   how many people don't you know why grandma invited to that party?

These are all eWI - exactly parallel to (317b) - with a quantity reading: in this context 'people' does not refer to specific persons; what is rather attributed to beliefs of the interlocutor is a number (the number of unreadable lines in the diary).

A quantity-reading *is* thus available in extraction of quantity phrases out of weak islands, it only needs to comply with the independent requirements on eWI: justifying presuppositions about a specific quantity.[26]

---

[26]Variations on the following minimal pair are often used to illustrate the quantity-reading effect (Koopman and Sportiche (1990), Rizzi (1990)):

(1)   a.   what do you think that you could weigh?

      b.   what do you wonder how you could weigh?

The relevant fact is that (1a) is ambiguous but (1b) is not ambiguous: (1a) has an agentive reading in which there is an action of weighing some object, and a stative reading in which the interlocutor's wheight is under consideration, but (1b) only has the agentive reading.

But again, once the eWI generalisation tells us where to look, it is easy to find contexts that make the stative reading of (1b) natural. The target reading is 'there is a number of kilos, such that you wonder how to weigh that number of kilo'. First, simply making the example pragmatically more plausible makes this reading easier to find:

(2)   how many kilos do you wonder when Mom weighed?

In a context in which we know that mom weighed 55 kilos when she was 20, and steadily progressed to 80 kilos now that she is 65 years old, and you want to know when she reached

A trivial way of showing that the ban is not on quantities *per se* are examples which explicitly mention numbers, making this reading entirely natural:

(319) a. what number do you wonder whether I saw come out at the lottery?
       b. what is the number of people that you wonder why John invited?

Szabolcsi and Zwarts (1997) note that there is a third quantity reading, an 'amount reading', which does involve quantities but does not involve measures or counting. This reading, according to them, is truly eWI-phobic.

Although they give no indication of what this reading would mean, a reasonable guess is that 'amount'-paraphrases and mass terms should qualify:

(320) a. how much anger do you wonder whether the treaty will generate?
       b. how many people do you wonder why John invited?
           Answer, pointing at a large group: this much.

Since 'anger' cannot be measured or counted in any way (given current technology) and the indexical answer to (320b) involves a non-numeric amount, these are non-measurable amounts. Is there a ban on eWI with such expressions?

Again, it turns out that once the adequate presupposition is legitimated (e.g. 'what is the amount such that you wonder whether the treaty will generate this amount of anger'), the sentences drop back to the usual levels of eWI extraction. Suppose for instance that Félicien and Maximilien have been comparing a new treaty to two older examples of treaties that had generated respectively mild and violent anger. Maximilien is wondering whether the new treaty will generate a comparable amount of anger to one of the two older ones and Félicien didn't understand which one of the two Maximilien has in mind. Félicien can now felicitously ask (320a). Idem for (320b).[27]

The same point is illustrated by the fact that inserting a partitive structure makes the mass-terms acceptable again (Kiss (1993), Rizzi (2000)):

---

some intermediate amount of kilos that I don't know about; it is natural to ask the above question.

Exactly the same holds for (1b): if you are overweight and you know how to go down to, say, 80 kilos, but there is some lower amount that you don't know how to go down to, I can ask about it using (1b).

[27]Szabolcsi and Zwarts (1997) give (1) as an example of an impossible amount eWI:

(1) * combien   de cercles as-tu   beaucoup dessiné?
       how-many of circles have-you much   drawn

This is a somewhat unfortunate example as it is ruled out independently of weak islands:

(2) * tu as   beaucoup dessiné 5 cercles.
       you have much   drawn 5 circles

It might be tempting to view (2) as a weak island induced by *beaucoup* which would block an LF-movement of '5 circles'. But that will not work, as it would wrongly rule out (3):

(321)   how much of his anger do you wonder whether Bartholomy will be able
        to channel healthily?

Partitives being prime inducers of range-readings, the reason is now clear: quantity readings survive eWI once presuppositions about quantities are legitimated.

Again, a trivial way of inducing an acceptable amount reading is to ask about amounts in the question itself:

(322)   what amount of anger do you doubt that/if Bartholomy will be able to
        channel healthily?

Notice that the same is true of examples such as (323) which otherwise seem to resist eWI strongly: comparing them to other amounts or mentioning amounts explicitly rescues them.

(323)   how many mutilated children do you wonder whether it will take to stop
        the sweatshops?

The (descriptive) moral of this story is simple: there is an interaction between presuppositions and quantity readings. Whenever a noun-phrase of the form 'how many/much N' is presuppositional, the presuppositions is unmarkedly about specific Ns, not about specific quantities (numbers or unmeasurable amounts).

Making this precise and understanding its underlying cause is an interesting topic concerning noun-phrases and their presuppositional readings, but it is hardly relevant to the theory of weak islands *per se*. As far as islands are concerned, the only relevant fact here is that eWI requires presuppositional readings.

7.4   The 'Event Reading' Asymmetry

It has been claimed that event-related readings are killed by eWI (Doetjes and Honcoop (1997)). This is based on the fact that (324a) has two readings: one

---

(3)   tu   n'as pas dessiné 5 cercles.
        you have not drawn 5 circles

Once the example is purged of the independent source of ill-formedness, the number/amount reading does pop out, as expected:

(4)   how many circles have you often drawn?

Here *often* is a WI-inducer, and accordingly provokes an odd reading whith presuppositions about particular circles being drawn more than once. As usual, the example becomes fine if a context is set up so that the presupposition is about numbers/amounts rather than circles: 'what is the number/amount such that you have often drawn that number/amount of circles'.

whereby we are counting ship-passing events (the same ship passing several times counts several times) and one whereby we are counting actual ships (the same ship passing several times counts only once). But (324b) allegedly only has one reading: the 'actual ship' reading, not the 'ship-passing event' reading. The proposed generalization being that event readings are generally blocked in eWI.

(324)   a.   how many ships do you think that the guard saw/let pass through the canal?

         b.   how many ships do you wonder whether the guard saw/let pass through the canal?

As with quantities however, the 'impossible' reading doesn't hide very far. Let us start by the trivial test: changing the examples to explicitly mention events makes the issue disappear:

(325)   a.   how many ship-passings do you think that the guard saw this year?

         b.   how many ship-passings do you wonder whether the guard saw this year?

Both of these have the reading that counts events of ship passing, given moderate help from context. Whatever the generalization is, it is thus not about 'events' *per se*, but rather about events associated to noun-phrases headed by 'ship' as opposed to 'ship-passings'.

As expected, inserting the undoctored examples into an adequate context also makes the issue disappear. Suppose for instance that you have a particular interest in the ship-passings. Maybe you are the owner of the canal, so that each traversal brings you money. One day you come to know that there were three traversals made at night with all lights off, in the hope that the guard wouldn't notice them. I know all this, and knowing how greedy you are, I know that you must be worried whether the guard caught them or not. If I don't know how many traversals were involved, I can ask (324b). Obviously, in this context the question is about traversals, not individual ships, since only traversals correlate with the size of your wallet: the intended (grammatical) meaning is 'what is the number of ship-traversing-the-canal events that you have in mind, such that you wonder whether the guard saw/let that many traversals happen'.

The (descriptive) moral is again simple: there is an interaction between presuppositions on noun-phrases and event-readings. Whenever a noun-phrase of the form 'how many/much N' is assigned an existential presupposition, the presupposition is unmarkedly about specific Ns, not about specific events in which N took part. Again, this is an interesting interaction, but hardly a topic for locality. As far as locality goes, the relevant fact is that eWI requires a presuppositional reading of the wh-phrase.

Notice that in this case, we have independent evidence that the source of the issue is an interaction between presuppositions and event-readings: partitives. Transforming the wh-phrase into a partitive - independently of islandhood - triggers exactly the same effect:

(326) a. how many of the ships from Babylon did the guard saw/let pass through the canal?

  b. how many of the ships from Babylon passed through the canal?

In both cases the event reading recedes, just as in the eWI (324b), but there is no island in sight. Simply put, eWI and partitives are two different ways of triggering presuppositions, and the latter wreak havoc with event-readings.

7.5 The Uniqueness Asymmetry

Some more intricate paradigms involving the interaction of uniqueness requirements and eWI are noted by Comorovski (1996), attributing (327b) to B. Partee, and by Szabolcsi and Zwarts (1997):

(327) a. who does Sue $\left\{ \begin{array}{c} \text{believe} \\ \text{know} \end{array} \right\}$ that John is fond of?

  b. who does Sue $\left\{ \begin{array}{c} \text{believe} \\ \text{\#know} \end{array} \right\}$ that John is married to?

(328) a. who do you $\left\{ \begin{array}{c} \text{think that} \\ \text{wonder whether} \end{array} \right\}$ I showed the letter$_{token}$ to?

  b. who do you $\left\{ \begin{array}{c} \text{think that} \\ \text{\#wonder whether} \end{array} \right\}$ I got this letter$_{token}$ from?

In each case, the difference between the a and b examples is that the predicate in b needs uniqueness: the conventions are that John can only be married to one person although he can to be fond of many, and the world is such that a particular letter token can be received only once (barring resending the same token), although it can be shown many times.

Within the b examples (requiring uniqueness), eWI is degraded: in (327b) - Comorovski's ex. (22-23), p.174 - the factive *know* induces a non-felicitous reading whereby John is married to several different persons, whereas no such reading is induced by *believe*. Similarly, in (328b) - adapted from S&Z - eWI induces a reading whereby the letter was received multiple times.

Is eWI uniqueness-phobic then? Here are four arguments that the real problem is not related to extraction of singletons - in fact not related to extraction at all - although it is related to weak islands and partitivity (range).

[i] eWI with uniqueness predicates *is* in fact possible: putting partitivity away makes the oddity disappear. Assigning a 'specific' reading rather than a 'range' reading makes the eWI examples acceptable:

(329) There was this wonderful singer in ancient Egypt who was adulated and about whom many rumors were flying around. One of the rumors is that he was secretly married to Nefertiti. Archibald is a historian studying this singer and is pretty certain that he has uncovered the secret of the marriage. His colleague Slimey wants to know about the Nefertiti-marriage controversy, and knows that Archibald uncovered it, but doesn't know who the singer is. To discover the identity of the guy he asks a typical slimey question:

who do you now know whether Nefertiti was married to?

In such a context, no multiple-marriages are evoked by eWI. Exactly the same happens with the temporal-uniqueness examples. To illustrate, imagine that it is important for you to know who I received that letter from, but I won't tell you. Knowing that you've finally found one possible sender but you're still wondering whether it's really the right one, I'm curious whether you're on the right track and I ask you: 'so, who do you wonder whether I received that letter from?', a felicitous question with no implications about multiple senders.

The generalization is then that the unicity issue only arises with range-based eWI. But range, or partitivity, is not the whole story, as the next argument shows.

[ii] eWI with uniqueness predicates is possible even in a partitive reading if the pragmatics is set right. In a situation which legitimates beliefs about several alternative referents of a unique slot, the questions stop invoking multiple referents of the wh-phrase:

(330) We all know that Magdalena is married to one of the ten guys over there, but we don't know who. You claim to already know about 6 of them that they are not the husband, but you're still unsure who among the remaining 4 is the right one. So I ask you:

who (among those guys) do you still wonder whether Magdalena is married to?

Again, no multiple-marriages are evoked by eWI, despite it being a range-eWI. Exactly the same holds for the letter-sending example (substitute letter-sending for marrying in the above scenario). Uniqueness is thus not phobic of eWI *per se*. (The - intermediate - generalization is rather that the uniqueness-oddity arises in 'contextless range/partitive eWI'.)

[iii] the same uniqueness-oddity arises in the absence of eWI, if (contextless) partitivity is present:

(331) a. # who$_i$ (amongst you) regrets PRO$_i$ having killed Slimey?

  b. # who$_i$ (amongst you) regrets the way in which he$_i$ has killed Slimey?

(332) a. # who$_i$ (amongst you) wonders whether he$_i$ killed Slimey?

  b. # who$_i$ (amongst you) wonders why he$_i$ has killed Slimey?

Here 'kill Slimey' is a uniqueness-requiring predicate as people only die once. No extraction has taken place, the wh-phrase is the subject of the root clause, and binds a pronoun in the lower clause. Nevertheless, the unicity-oddity is triggered again: the reading of these questions is that several people could have independently killed Slimey, and the speaker wants to know which subset actually did; an absurd reading (the same judgments hold for the other uniqueness-predicates above). (Notice that the effect is much lighter with *whether*, (332a), on which see below).

The crucial, and surprising, point is that the asymmetry between weak islands and normal extraction is also reproduced, despite the absence of extraction:

(333) who$_i$ (amongst you) thinks that he$_i$ has killed Slimey?

(333) does not create a uniqueness-clash. The systematic difference between questions which pattern with (331-332) and with (333) is that the former are weak island contexts and the latter are not. The descriptive generalization is thus that the uniqueness-oddity arises only where a partitive takes scope over a WI-structure, regardless of extraction.

[iv] That we are along the correct path is suggested by the fact that any distributivity over a unicity-requiring WI-context leads to the same violation, even where no question is involved at all:

(334) a. # every crook wonders how he killed Slimey

      # three boys wonder how they killed Slimey

  b. every crook thinks that he killed Slimey

      three boys think that they killed Slimey

Put aside the 'group' reading whereby the three boy have collectively killed Slimey (this reading is discussed in the next section), and consider the interpretation where each boy/crook thinks/wonders separately. Again, the weak-island sentences are odd, (334a), while island-free contexts are not, (334b).

The generalization is thus simple: distributing over a WI-structure destroys uniqueness, regardless of movement or questions. The uniqueness-effect is thus not a property of movement out of a weak island. Why does distributing over a WI trigger uniqueness?

Without diving too far afield, we can now reduce it to the following paradigm: 'spelling out' the wide-scope partitive reading, so to speak, provokes the same

uniqueness-oddity (this also constitutes an additional argument against linking this phenomenon directly to eWI). That is, asking the corresponding yes/no question more than once is odd exactly in those contexts where eWI-uniqueness is odd:

(335)   Tell me: do you think that I received this letter from John? And Monika, do you think that I received this letter from Monika? What about Marcello, do you think that I received this letter from Marcello?

(336)   a.   # Tell me: do you wonder whether I received this letter from John? And Monika, do you wonder whether I received this letter from Monika? What about Marcello, do wonder whether I received this letter from Marcello?

        b.   # Tell me: do you wonder why I received this letter from John? And Monika, do you wonder why I received this letter from Monika? What about Marcello, do wonder why I received this letter from Marcello?

(337)   # Tell me: do you regret that I received this letter from John? And Monika, do you regret that I received this letter from Monika? What about Marcello, do regret that I received this letter from Marcello?

Notice that none of these questions involves extraction out of weak islands, but the symptoms are exactly the same as in the other paradigms above: degradation only in WI contexts, and rescuable by an adequate discourse-context.

This transparently suggests that the root of the issue is that a range-based question is equivalent to asking the question multiple times. More precisely, range induces distributivity over the whole question. Put differently, the population in the range consists of individuals potentially satisfying the questioned predicate. To illustrate, the relevant gloss for a partitive structure should be (338b) rather than (338a):

(338)   a.   among John, Monika and Marcello, which one do you wonder whether you received this letter from?

        b.   given John, Monika and Marcello,
              do you wonder whether you received this letter from John?
              and do you wonder whether you received this letter from Monika?
              and do you wonder whether you received this letter from Marcello?

If that is what range does, the rest (descriptively) follows.

Of course, this only pushes the question further down: why is distributivity sensitive to scoping over eWI (independently of any extraction taking place)?

This is no doubt an important question, but again tangential to our current concern (locality). Let me thus simply point at a plausible path for further

inquiry: it is a little noted fact that indirect questions are factive contexts - i.e. transparent propositional attitude predicates:

(339)   a.   Ulrich wonders when the parliament passed the Bertelsman-sponsored bill
             ↪ *the parliament passed the Bertelsman-sponsored bill*
        b.   we all regret that the senate passed the RIAA-sponsored DMCA
             ↪ *the senate passed the RIAA-sponsored DMCA*
        c.   Slimey thinks that abusing colleagues for one's own comfort is ok
             $\neq>$ *abusing colleagues for one's own comfort is ok*

A natural road to travel is thus to derive the above paradigms from the putative fact that distributing over factives implies the existence of multiple facts/events while distributing over non-factives doesn't. If true, this entails that unicity-requiring predicates will shy away from environments distributing over factive contexts.

(Interestingly, the factivity effect does not hold of yes/no indirect questions, and yes/no questions are precisely contexts which provoke a lighter unicity effect, cf. the remark above about (332a).

Coming back to our concerns, the unicity paradigm again tells us nothing about eWI *per se*: the fact that (some) eWI wreak havoc with unicity is a side-effect of their being presuppositional, or more precisely, range-based.

7.6   The 'Group Reading' Asymmetry

A closely related paradigm is the putative interaction between eWI and pluralities. Szabolcsi and Zwarts (1997) hold that 'group readings' are washed away by eWI, and only distributed readings of pluralities survive. In (340a), John can have been helped several times by several different people, or can have been helped just once by a group of people. The eWI (340b) on the other hand makes the 'single event of help by an entire group' difficult or impossible to reach.

(340)   a.   which people do you think that John was helped by?
        b.   which people do you wonder how John was helped by?

The asymmetry is particularly visible in uniqueness-requiring (341): the distributed reading is impossible as seen above and therefore only the group reading remains. But the group reading is washed away by eWI, so no reading remains at all:

(341)   a.   which trolls do you think that John was killed by?
        b.   # which trolls do you wonder how John was killed by?

The judgments are subtle, but let us grant the asymmetry. Does this tell us that eWI are group-phobic?

It doesn't look like it. As with quantity readings and event readings, what seems to be the case is that a presuppositional reading of plural noun-phrases or noun-phrases of the type 'which N' lead by default to presuppositions about Ns, not about groups of N. But making the group reading more available cures the problem. Let us go over three types of facts indicating this.

The first fact we already know about from the preceding section: there is an independent factor promoting the distributed reading in (340b) and (341b), namely the presence of a range-reading. Recall that the unmarked reading for a context-less eWI is the range-reading, which involves distributivity. The fact that (340b) and (341b) prefer a distributed reading over a group-reading therefore doesn't allow us to conclude that the group-reading is handicapped by eWI. Maybe the presence of presuppositionality is forcing/favoring a distributive reading and thus giving the illusion that group-readings are absent.

Here is an illustration of presuppositionality forcing distributivity and thus making the group reading unavailable independently of islandhood:

(342)   which trolls among those over there (do you think) John was killed by?

(342) is a minimal variation on (341a) but is just as odd as (341b), making the group-reading fade-away.

To ascertain whether eWI is sensitive to group readings, what is needed is a context which forces such a reading and pushes the distributive presupposition out of the way. If group-readings are still unavailable in these conditions, eWI is group-phobic. If the group-reading is available under such contexts, the group-reading effect is an artifact of the deeper presuppositionality requirement.

The second relevant data-set is thus examples such as the following, which show that pragmatics cures the problem: the group reading *is* available if the context makes it reachable. Providing plausible groups is for instance sufficient:

(343)   by which soldiers do you wonder whether John was killed, the Americans or the Burmese?

Similarly, mentioning groups explicitly also makes the group-reading natural in eWI:

(344)   by which group of teenagers do you wonder whether John was killed?

It is thus indeed not the case that eWI is group-phobic. The group problem only appears when eWI forces range, and range is left to itself without context. In such a case, default presuppositions for the noun-phrase are built, and these

happen to be about Ns and not about 'groups of N', giving the illusion that group-readings are handicapped by eWI.

Again this is an interesting topic for the interaction between existential readings and noun-phrase structure, but it is not directly relevant to extractions out of weak islands. As far as eWI is concerned, there is nothing here beyond the fact that eWI is presuppositional.

7.7 Idiom Chunks and Inherently Unspecific Objects

Many chunks of idioms have a special sensitivity to eWI: they can be wh-moved, but they cannot cross weak islands. This turns out to correlate with the fact that these chunks do not support presuppositions.

Here is an extreme case: the French (345a) and (345b) have the same meaning, *modulo* style (i.e. (345b) is 'popular' - or 'vulgar', depending on context):

(345)   a.   qu'est-ce que tu crois qu'on   va faire cet après-midi?
             what is-it that you think that we will do   this afternoon

        b.   qu'est-ce que tu crois qu'on   va foutre cet après-midi?
             what is-it that you think that we will do   this afternoon

but (345b) collapses in weak islands, while (345a) survives:

(346)   a.   qu'est-ce que tu aimerais savoir comment faire?
        b. * qu'est-ce que tu   aimerais   savoir   comment foutre?
             what is-it that you would-like to-know how   to-do

similarly with negation: in a context where activities were planned, but some of them were canceled, (347a) is acceptable, but (347b) is impossible:

(347)   a.   qu'est-ce qu'on fait plus?
        b. * qu'est-ce qu'on   fout plus?
             what is-it that-we do   not-anymore
             what don't we do anymore?

Similarly in 'what is the x' structures (independently of islandhood):

(348)   quelle est la chose que tu crois qu'on va $\left\{ \begin{array}{c} *\text{foutre} \\ \text{faire} \end{array} \right\}$ cet après-midi

This series of restrictions correlates with the fact that *foutre* only takes a drastically limited range of objects: it only occurs with the non-specific 'something' and 'nothing' (putting aside some idiomatic readings such as *foutre le feu* (put on fire), etc.). *Faire* on the other hand takes the standard range of objects:

$$(349) \quad \text{je vais} \left\{ \begin{array}{c} *\text{foutre} \\ \text{faire} \end{array} \right\} \left\{ \begin{array}{c} \text{mes devoirs} \\ \text{une promenade} \\ \text{la vaisselle} \\ \text{un gateau} \end{array} \right\}$$

In other words, the idiomatic version of the construction can never be presuppo-sitional, independently of movement, and can thus not occur in $\exists$Q-movement.

The same holds for the standard cases of idiom-chunk extraction. Here are the cases discussed by Rizzi (1990) (translated into French, where judgments seem identical):

(350) a. ? le   parti   que je compte tirer   de   la   situation
           the advantage that I intend to-take from the situation

       b. * le   parti   que je sais   comment tirer   de   la   situation
           the advantage that I know how   to-take from the situation

       c. * c'est ce   parti   que je compte tirer   de   la   situation
           it'is this advantage that I intend to-take from the situation

       d. * il y a   du   parti   que je compte tirer   de   cela
           there is some advantage that I intend to-take from this

(351) a. l'attention   que je pense que tu   portes à ce   projet
           the'attention that I think that you give   to this project

       b. * l'attention   que je sais   pourquoi tu   portes à ce   projet
           the'attention that I know why   you give   to this project

       c. * ceci est l'attention   que tu   portes à ce   projet
           this is   the'attention that you give   to this project

       d. * il y a   de l'attention que tu   portes à ce   projet
           there is some attention that you give   to this project

(352) a. the credit that I think that you are giving to the UN's good faith

       b. * the credit that I know why you are giving to the UN's good faith

       c. * this is the credit that I am giving to the UN's good faith

       d. * there is credit that I am giving to the UN's good faith

(353) a. the headway that I think that the movement is making

       b. * the headway that I wonder how the movement is making

       c. * this is the headway that the movement made

       d. * there exists a headway that the movement made

In each case, the idiom-chunk movement degrades with eWI (the b-examples), but also with specific readings of the wh-phrase, regardless of islands (the c-

examples). Even an existential (non-specific) constructions is unbearable for the idiom-chunk (the d-examples).[28]

Again, iChunks do not like presuppositions and thus refuse to partake in eWI-movement.

---

[28]For reasons that I do not understand, some counterparts of the c examples are markedly better, sometimes perfect when involving questions on clefts. Here are some minimal pairs with the text examples above:

(1)   C'est tout le parti que tu comptes tirer de la situation?

         is this all the advantage that you intend to take from the situation?

(2)   C'est toute l'attention que tu portes à ce projet?

         is this all the attention that you give to this project?

(3)   Is this all the credit that you're giving to the UN's good faith?

(4)   Is this the only headway that the movement managed to make?

This paradigm recalls Romero (1998)'s observation that some clefted eWI do reconstruct for scope. Both cases involve SQ-movement which exceptionally behave like Q-movement.

8   Unifying Merge and Move

Given our current understanding, concatenation and movement are two funda-
mental operations of syntax; variations in terminology ('merge', 'group', 'xbar',
'syntagm', etc. for concatenation; 'chain', 'displace', 'remerge', etc. for move-
ment) and in implementation (derivational, representational, top-down, bottom-
up, percolation-based, etc.) aside. Theoretical progress over the last decade has
accentuated the centrality of these two core operations: the substance of recent
progress lies in the attempt to strip the syntactic toolbox down to these two
tools, ideally relegating other properties of syntax to mere byproducts of the
idiosyncratic nature of concatenation and movement.

A striking aspect of these two core operations is their similarity: the essence of
both operations is to 'group' two entities into one abstract 'unit'. In the case of
concatenation, the two input units are distinct and the result is an asymmetric
composition of the two; in the case of chain/move, the link expresses an identity
between two otherwise unrelated nodes in a structure.

The fact that both primitive operations of syntax are 'grouping' operations is
an encouraging result: it is the very nature of syntax to be a 'grouping engine',
taking a sequence of tokens as input, and recursively grouping them into a tree
or a graph. Achieving substantial descriptive adequacy while limiting the theory
to little beyond the essential grouping operation is a remarkable achievement.

The relevance of this achievement is particularly obvious if current versions of
the theory are compared to earlier or different brands. In the latter, it is by far
not the case that the syntactic toolbox reduces to grouping operations. Most
of the constraints in GB have for instance little or nothing in common with the
core concatenative procedure (the 'case filter', the theta-principle, and more
importantly, the notion of government itself). This is still more radical with
approaches pre-dating the 'Principle & Parameters' framework: apart from the
trivial observation that a rewrite rule is essentially a grouping operation, the
substance of the empirical research did not map onto grouping in any direct
way. The empirical research mapped onto the *form* of each individual rewrite
rule, together with some generic filters on those rules. For all their potential
explanatory merits, both of these notions are unrelated to the core concatenation
operation.

The current theoretical success is however incomplete: the similarity of the two
core operations is not expressed, leading to duplication. As it stands, the theory
states twice that syntax is a concatenation engine, missing their common core.

What is more, the two basic operations not only have a common core, but
there is also a clear source for their difference: their input is different. One
operation operates on pairs of identical items ('chain') while the other operates
on pairs of distinct items ('merge'). A true explanatory account would thus not
limit itself to factoring out their common core, but would seek to unify the two
operations entirely, deriving their apparent difference from the differing nature
of the elements they are applied to.

(If binding is thought of as a syntactic operation, the overall picture becomes more striking still: binding is a third variant of the abstract grouping operation. It dictates the conditions under which two overt elements are considered to be only one underlying unit by the grammatical module; again, the conditions under which two items can be grouped into one. Bringing binding into the picture magnifies the fact that all principles of modern syntax are about the core 'grouping' function of syntax; but it also augments the redundancy: the theory is now repeating three times that 'grouping' is what syntax is all about. To simplify the discussion, I will follow Chomsky and Lasnik (1993) in considering BT to be an outside filter on the syntactic engine, and will thus leave BT aside. See however Starke (1997) for an approach to binding along the general lines pursued below).

At first sight, it appears that some effort has been put into reducing this duplication by factoring the 'move' operation into two subparts: a merge step and a detach (or copy) step, thus dispensing with the 'move' construct (let us call this the 'cut & paste', or 'copy & paste' approach). Such efforts have the merit of simplifying the theory by reducing part of its content to more primitive and sometimes independently needed components, but they do not address the above issue of redundancy. A 'copy & paste' world still needs to relate the landing site (the 'paste' position) to the departure site (the position of the 'copy' operation), in order to state their identity and various other properties. To do this it needs to supplement 'merge' with a tool equivalent to 'chains'; thereby reproducing the core redundancy.

In fact, the very idea that 'move' ('chain') and 'merge' ('syntagm') might be two facets of the same underlying concept appears preposterous within the traditional understanding of these terms in linguistic research. It is standardly thought that the properties of move/chain and the properties of merge/syntagm are disjoint, and hence not directly comparable.

More precisely: the major independent property of chain/move is 'locality' while the major independent property of merge/result-of-merge is 'labelling' (or 'projection'). If chain/move has no interesting labelling property and merge/result-of-merge has no interesting locality property, the two operations have distinct properties and are thus distinct concepts.

This traditional view however rests on tacit assumptions - or hidden stipulations - obscuring the relevant symmetry. The notion of 'merge' (or xBar) *does* have an interesting locality property: the sisterhood property hard-coded into the operation itself. Conversely, the 'chain' ('move') construct *does* have non-trivial 'labelling' properties, obscured by tacit - and unnecessary - assumptions about the nature of labelling.

On closer inspection, it thus turns out that the *kind* of properties of both constructs are the same: they both have a 'locality' property and they both have a 'labelling' property. We can thus meaningfully ask whether the properties of move and the properties of merge are the same or not.

136

Examining each of these properties in turn will reveal that they are in fact the same across operations: the locality of merge is a special case of the locality of move, and the 'labelling' property of chains has the same characteristics as the labelling property of syntagms.

If so, the two grouping operations share all their properties, and are thus one and the same concept. Seeing this allows us to significantly simplify current syntactic theories, reducing both 'move' and 'merge' to a single generalized grouping operation. In more traditional terms, 'move' is unneeded: it dissolves into 'merge'.

8.1 Locality and 'Labels'

Merge always occurs under sisterhood. Why? At its root, this encodes the assumption that compositionality always occurs under strict sisterhood (in fact under linear adjacency, given current assumptions). This is the underlying cause of the fact that syntactic phrase-structure theories systematically have the built-in property of sisterhood, and of the fact that function-application is tacitly stipulated to apply to sisters.

But there is nothing intrinsic to this sisterhood property. One can easily imagine and construct (compositional) formal grammars in which compositionality does not respect any kind of adjacency or sisterhood. To put the point in more familiar syntactic terminology, nothing principled prevents a merge operation from relating two distant nodes in a preexisting tree, thereby creating a graph. Such configurations are stipulated to be inexistent, by building a sisterhood restriction into the merger operation itself. So why is sisterhood required?

As it happens, the sisterhood constraint follows from standard *Relativised Minimality*. Remember that it is implicit in Rizzi's *Relativised Minimality* that features classes form a (trivial) tree:

(354)

```
                        R
               Q               A
           wh neg foc     person number case
```
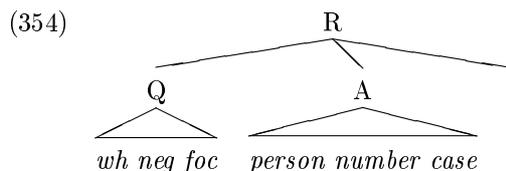
Locality is then simply expressed as the impossibility of 'crossing' a different token of the same class.

Consider then what happens in a typical instance of merger, or compositionality: two items α and β from different classes are grouped into a new unit, γ. At first sight, *Relativised Minimality* has nothing to say here, since the two items appear not to belong to a common class. But as (354) makes explicit, for any two tokens α, β there is always at least one class including both of them: the root class R.

137

If *Relativised Minimality* looks at a canonical merge operation, it will thus see an 'R-grouping', or an 'R-chain'. As a result, it will make a strict prediction for the locality of this merge operation: any token from class R is predicted to be a deadly intervener.

Since every syntactic feature is a member of R, it follows that any syntactic intervener, regardless of its nature, will prevent the merge operation from succeeding. Sisterhood being the sole structural relationship where no syntactic node intervenes, *Relativised Minimality* forces 'R-groupings' to apply under sisterhood. Compositionality is thus locked into sisterhood relationships. As a result, the locality of 'merge' is a special case of a more general - and independently needed - locality principle.

(Notice that this conclusion is entirely independent of the refinements of *Relativised Minimality* explored in the preceding chapters: it follows from the standard unmodified version once the tacit existence of a root class is brought out.)

It was the goal of the preceding chapters to show that this very locality principle exhausts the locality of 'move'. If so, the locality of 'move' and the locality of 'merge' are literally one and the same locality: an updated version of *Relativised Minimality*.

The apparent difference between the two locality requirements (±sisterhood) now follows from the differing nature of the input, as desired: with two identical elements as input, the common class is a subclass of R and locality is accordingly looser; with unrelated elements as input, the common class is R and locality is strict.

Whatever the right implementation of 'merge' - and thus compositionality - we want it to express the fact that its sisterhood property is a consequence of a general principle, and not hardcode it into the operation itself.

Asymmetric Inheritance

What of the projection - or labelling - property of merge, standardly taken to be absent in move/chain? What labels express are three facts: (i) non-terminal nodes are typed, i.e. they are not all equal, (ii) the type of a non-terminal node is inherited, (iii) the inheritance is asymmetric.

Once labelling is understood as a mere name for this cluster of properties, it is obvious that move/chains has it too. The most trivial illustration of this is the fact that the locality property of a given chain is a function of the *type* of chain it is: A' chains have a different locality (are sensitive to different interveners) than A-chains. Chains thus come in different types, just like nodes do. Furthermore, the type of a given chain is a function of the links that constitute the chain, the type is thus inherited. Finally, the determination of the type of the chain is asymmetric: a given departure position will create different types of chains depending on the landing site. The type of the chain is thus a function of

the landing site, not of the departure point. (Notice the similarity with head-complement groupings: the leftmost element decides the type of the resulting entity, a point I come back to in §9).

The major independent property of merge - asymmetric inheritance - is thus shared by move/chain; both are subject to the same general algorithm to determine their type.

A more interesting illustration of the fact that chains are typed ('labelled', in a now misleading terminology) comes from the reconstruction properties and interpretive properties of chains discussed above. Recall for instance that SQ-movements do not reconstruct for scope, but Q-movements do (expressing the fact that there is no scope reconstruction into weak islands). Taking this fact seriously entails that each movement/chain has its own interpretive properties and chains are thus typed. Again, the type of the movement/chain is a function of its landing site. Similarly for the various other interpretive properties of chains discussed above (range, specificity, etc.).

Clearly, whatever the precise labelling algorithm turns out to be (§9), the 'labelling' property by itself does not provide us with evidence that the two operations 'move' ('chain') and 'merge' ('syntagm', 'xBar') are distinct. To the opposite, the similarity between determining the type of a chain and determining the type of a projection suggests a unified account.

8.2 Move as 'Distant' ReMerge

Putting both results together, the 'move' operation is characterizable as a "grouping operation subject to *Relativised Minimality* whose type is asymmetrically determined from its input entities" and the 'merge' operation is a "grouping operation subject to *Relativised Minimality* whose type is asymmetrically determined from its input entities".

How can we express this deep similarity? Since merge is more basic than move (move presupposes merge but merge does not presuppose move), a unification of the two operations necessarily results in a version of merge and therefore the lesson from the above reasoning is that move is a form of 'distance merge'.

In its most simplistic - and least palatable - implementation, instances of move/chain are mechanically substituted by instances of merge.

First, each move-step/chain-link inherits a 'label' from its landing site: since move/chain is typed in the same general way as merge is typed, syntactic representations should express either both or neither. Second, the graphical hack of drawing the two relations with different types of connectors is unwarranted. A mechanical translation thus results in (355b) in place of (355a): (irrelevant layers of T/V and irrelevant functional projections omitted for space)

(355)   a.   Q
                  Q
                     T
                        V
                           DPwh

b.   Q
      DPwh   Q
               T
                  V
                     DPwh
                  Q

Or equivalently:

(356)   Q
                Q
            DPwh   Q
                      T
                         V
                            DPwh
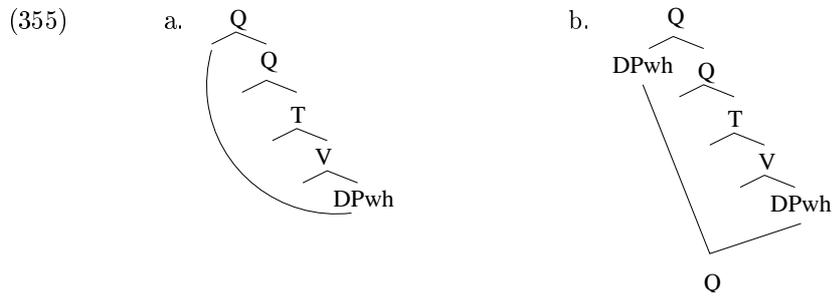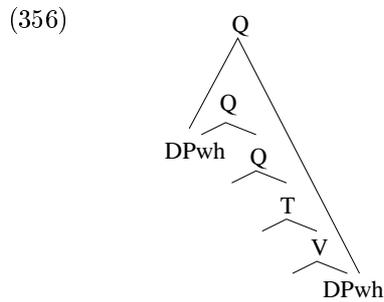
This representation encodes the fact that what we (mis)took for a move operation is nothing beyond a merge operation operating on two non-adjacent nodes.

But it still needs to be simplified. Consider the wh-phrase in specQP. According to (356), it has the bizarre property of being merged with itself, before being merged with an independent Q node: i.e. to construct (356), one first makes a copy of DP_{wh}, then merges that copy with the original, and only then proceeds to merge with the top node of the tree, the Q projection. An absurd sequence of operations.

One way to see the absurdity of this is to realise that such a self-merger could be iterated *ad infinitum*, with the same result, eg:

(357)   Q
                   Q
                      DPwh   Q
                  Q
              DPwh   Q   DPwh
                        T
                           V
                              DPwh

140
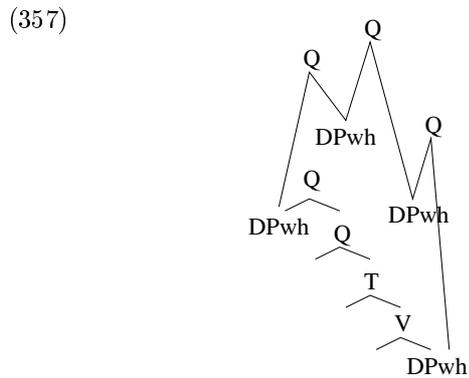
This undesirable result is simply a holdover of the stipulation that merge is subject to linear adjacency: this assumption caused the creation of several copies ('traces') of each moved phrase, so that each copy could individually comply with adjacency. But we have now eliminated the stipulation and allowed merge to operate 'at a distance', thereby unwittingly revealing the awkwardness of self-merger. Eliminating it is trivial and produces the following structure:

(358)
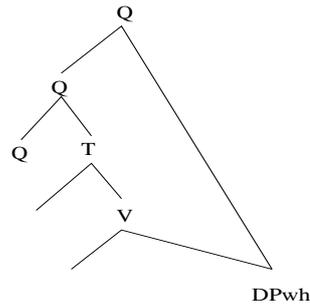
This structure expresses the same relationships as (355-356) and also captures the core similarity between local and distance groupings. To take one example, in (358), the structural relationship between the DP node and its 'long distance' sister is exactly that of a traditional specifier, with the same c-command relationships established, as desired.

The node corresponding to the old notion of 'move', or 'chain', participates in the phrase structure as any other node, eg:

(359)
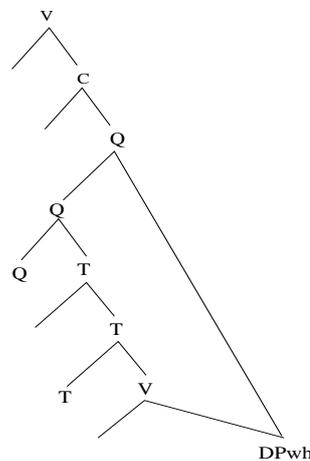
The two immediate results of unifying merge and move are thus that syntactic representations are generalized to graphs rather than trees - albeit highly constrained graphs - and the notion of 'traces' ('copies') becomes superfluous.

141

Instead of creating traces or copies, 'movement' is now a literal 'remerger' of a node already merged into the structure.

The results based on traditional trees are however preserved since each traditional tree can be algorithmically mapped onto its corresponding new form (and vice-versa). The new representations, such as (358), are theoretically more austere than traditional trees, but equivalent to the corresponding traditional representations for the central empirical cases handled by traditional trees (see however the discussion of order-preservation below for a case where the new representations seem empirically superior to traditional tree-representations).

To illustrate with one central case, take locality again. The locality of a simple xBar layer comes out exactly as needed: a head-complement relationship, say Q-TP in (358), cannot be a 'long distance sister' because it is an R-grouping and any intervening node would disrupt it. The relation between the spec and its sister (between DP_{wh} and Q in (358)) on the other hand is a Q-Q relation and is thus allowed to be a long-distance sister, only sensitive to Q interveners.

Locality of eWI is also preserved in the new structures: in a representation such as (358), it is still the case that the intermediate specQP intervenes between the departure point and the landing site (many irrelevant intermediate nodes omitted):

(360)   *
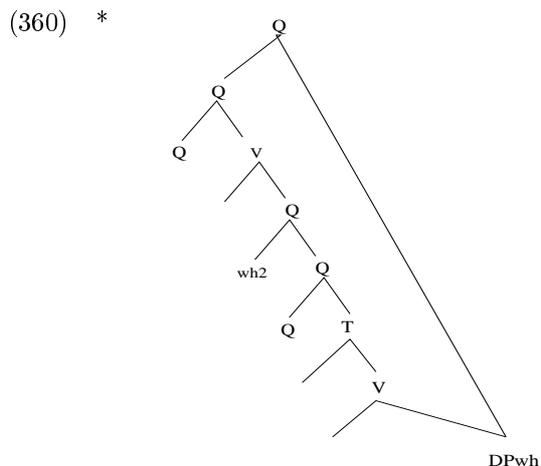
Locality results are also preserved in more convoluted situations. Take for instance the case where *tout* has remerged in a higher clause, and a wh indirect-object remerges in the higher CP, in an eWI configuration:

(361) *
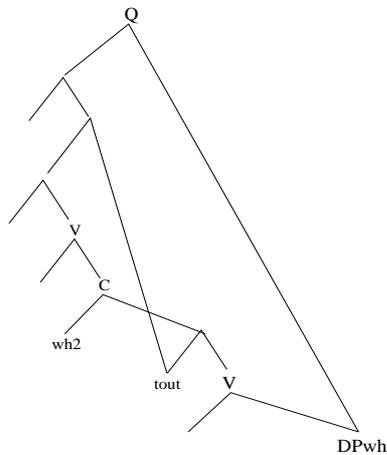
The 'shortcut' introduced by the remerger of *tout* does not affect locality: the intermediate wh still intervenes between the root CP and the *situ* position of the indirect object.

Generally then, the traditional and new structures are empirically similar, but the traditional approach is unnecessarily rich in its theoretical assumptions.

Preliminary implementation of 'merge'

No relevant modification to the logic of the anti-identity requirement is needed to obtain these results, although one change is needed in the formalization: c-command must be substituted by dominance.

(362) X-merge(α, β) iff
       α, $\beta \in$ X, α dominates β, and $\neg \exists \gamma$, $\gamma \in$ X, α dominates γ and γ is closer to α than β is.

This is a direct result of the fact that locality doesn't hold of a specifier and its trace, but of the topmost node of the to-be-merged tree with the in-situ element. It is also a welcome result: domination is a more primitive notion than c-command and locality is thereby simplified.

(Notice that 'closer' can be substituted by 'γ c-commands β'. More below on the formalization of closeness)

There is one aspect of concatenation which is however not adequately covered by (362): its asymmetric nature. Since head-complement relations, for instance, are cases of R-relationships, the entailment is that their label is uniformly 'R', an unexpected result. The relationship between the spec and its sister on the other hand gets the usual label. Eg:

143

(363)

```
              QP
             /  \
        DP_{+Q}   R_{+Q}
                 /    \
               Q°     TP
```
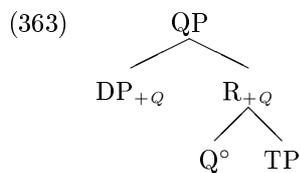
This bizarre (but not incoherent) result is the consequence of a deeper issue: our unified merge inherits from 'move' the assumption about identity between the departure point and the landing site. The asymmetric nature of the relationship will however come out clearly from the discussion in §9; and discussion of this point is thus postponed to that section.

8.3 Representations, Derivations and Imperfections

We have now attained the goal of expressing the common core of the two basic grouping operations: 'move' or 'chain' is literally an instance of 'merge'; and syntactic theory contains little beyond the core merge operation.

A simple declarative clause, with the subject starting in VP and ending in TP, thus comes out as:

(364)

```
                    T
                  /   \
                 T     \
               /  \     \
              T    V
                  / \
                SU   \
                      OB
```
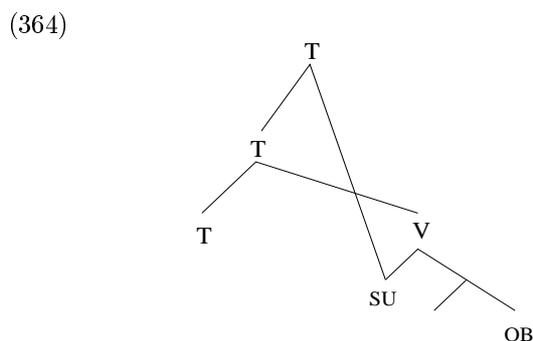
Let us then go through some consequences of this unification, starting with some non-issues.

First, there appears to be a new issue of spell-out: if there are no traces or copies and syntagms are simply remerged from their base position, how does the interface know where to spell the phrase out? This question is however the exact equivalent of the traditional question: which copy is spelled out? Instead of asking which copy is spelled out, we now ask which merger operation is taken into account by spell-out, the same question, giving rise to some same issues.

Similarly for split-reconstructions and other splittings of a single phrase into several interpretive pieces: to the extent that such operations affect a single phrase (as opposed to various subcomponents of a phrase), the issues arise identically

144

in the traditional and in the new approaches: questions about multiple traces map onto questions about multiple mergers.

Reduplication paradigms are another instance of this logic: to the extent that they are adequately analyzed in terms of spell-out of a trace (i.e. spell-out of multiple 'copies'), they are now reanalysed as spell-out of multiple merger operations (instead of the default case: spell-out of a single merger).

Throughout, questions about multiple traces map onto questions about multiple mergers. No novelty introduced there.

Representations versus Derivations

There is however one massive consequence to the unification of merge and move: a complete rephrasing of the 'representations versus derivations' issue. More precisely, the one and only strong argument in the debate (favoring representations) now disappears.

The heart of the representation/derivation issue concerns traces, or more importantly, the need to keep the landing site and departure point of a movement connected long after the movement has landed. The movement operation itself has no means of making this connection, and another connecting device thus needs to be postulated ('chains'). This device however duplicates the functionality of the move operation itself, giving rise to a redundant theory.

Redundancy can be cut by throwing out movement and keeping chains, but cannot be cut by throwing out chains and keeping movement; i.e. movement is a proper subset of chains in its capacities. This is the logic of the traditional - and decisive - argument, making all current derivational theories redundant (careful elaborations of this point and its consequences is found across a series of Brody's recent work: Brody (1992), Brody (1995), Brody (1998), Brody (2000c)).

This argument however disappears once both 'move' and 'chain' are dissolved into the more primitive notions of 'merge' and 'syntagm': if neither 'move' nor 'chain' exists, no issue arises about their redundancy. This leaves the debate with no solid argument either way, and therefore brings derivational approaches back into the game as true contenders.

What remains of the debate is a different representation/derivation issue: the question as to whether the concatenation itself is to be viewed as representational or derivational.

Here however, no redundancy arises: both notions entail the other, as two sides of the same coin. A merge operation entails the existence of a 'result of merge' (a syntagm), and the presence of a syntagm entails the existence of an operation creating it. Both the representational and the derivational approach are possible and the 'representations versus derivations' issue now genuinely arises, possibly for the first time: which side of reality is linguistically relevant?

Brody (2000a) suggests that there is one remaining argument, favouring a representational view: a representational view is inherently more restrictive since it does not allow itself to use 'ordering' of operations as an explanatory tool. It is thus still a proper subset of the corresponding derivational toolbox.

Taking the psychological status of linguistic theories seriously - as all parties of the debate are - makes this argument however somewhat weaker than seems at first sight, and in fact possibly reverses it.

As observed above, the presence of a syntagm ('result of merge') entails the presence of a merge operation. This would not be true in an abstract mathematico-logic theory, but is a necessary consequence of the embodiment of the theory in speakers. There is thus no doubt that some form of derivation - in the sense of merger steps - exists. The issue is therefore not about the existence of one side or other, but rather about the location of the linguistic principles: do they hold of each derivational step or of the final result of the derivation (the final representation). Seen in this light, a representational theory is equivalent to a derivational theory with one additional constraint: the constraint that all linguistically relevant principles apply after the final step. (This of course presupposes that 'derivations' are taken in a psychologically literal sense. This is often denied in theory - for irrelevant historical reasons - but corresponds to the actual practice of derivationalists.)

On theoretical grounds the two approaches are thus difficult to tease apart: granting Brody's point that limiting linguistic principles to the last derivational step leads to restricted expressive power; a representational theory gains on parsimony of expressive power what it loses on additional stipulations: the stipulation restricting linguistic principles to the last step is dispensed with in a derivational approaches.

This issue can now in principle be translated onto empirical grounds: the issue is whether linguistic principles do in fact need to make any use of the expressive power of the notion of 'ordering' of operations; i.e. (counter-)bleeding and (counter-)feeding relationships.

In sum, recognizing the underlying similarity of *merge* and *move* has the unexpected result of cutting the traditional simplicity argument in favour of a representational view, thereby opening up the issue.

'Imperfections'

Another important theoretical consequence of the unification of move with merge is the disappearance of a central question in the minimalist programme (e.g. Chomsky (1995)). Minimalist approaches emphasize one important question: why is grammar not as austere as it could be if it's sole function is to relate sound to meaning? In particular, why does it contain paraphernalia which is unnecessary to achieve this goal, such as uninterpretable features or movement

(chains)? Explaining away these unnecessary devices, 'imperfections', is the driving force of minimalist research.

Of course, in such a terminology 'merge' is not an unnecessary device, it is the core syntactic operation, necessary to create a single LF from a multitude of input tokens. The present result of reducing move to merge thus suggests that grammar is substantially more 'minimal' than standard minimalist views assume it to be; and in particular there is no such thing as a move-imperfection whose existence needs to be justified.

This presents a challenge to the minimalist: a grammar based on dissolving move into merge is substantially more minimal than any current minimalist model. It suggests that the very core of minimalism - the attempt to justify imperfections - is the wrong strategy, as the 'imperfect operation' does not exist. What does need justification is the assumption that there is such an imperfection.

Reconstruction

Although dissolving move/chains into merge/syntagm preserves all major empirical results, there is one domain which it does affect: the properties of chains themselves. One such case is the phrasing of which move/chain reconstructs and which does not.

The issue arises from the different rendering of successive cyclic chains, and in particular their differing departure points. A simple case of successive cyclic 'movement' (ignoring the θ-step):

(365)

is now mapped onto:

(366)

Q
Q
Q   V
         Q
         Q
         Q   T
         Q   V
                        DPwh

There is no more direct link between the intermediate position and the final landing site. The links all originate from the only position of the remerged element: the base position.

At first sight this wreaks havoc with reconstruction: there is no obvious way to state the fact that the topmost link 'reconstructs' into the intermediate link. This is particularly troublesome with respect to the clean generalization that SQ-chains do not reconstruct for scope while Q-chains do, a natural result given the independent interpretive nature of the chains.

In the new structures, the 'reconstruction' of a Q-grouping (into an intermediate SQ position) seems to amount to the stipulation that the set of remergers of α is interpreted in the position immediately <u>below</u> the lowest Q-merging of α, an unpalatable state of affairs.

The issue cannot be avoided by simply linking the intermediate position to the topmost landing site:

(367)

Q
Q
Q V
Q
Q T
Q
V
DPwh

This would represent a link of the entire intermediate CP with the topmost landing site, an analysis maybe plausible for cases of partial wh-movement (McDaniel (1989)), but not for cases of interest here.

The problem is however only apparent: the relevant generalization is straightforwardly restated once the generalization is inverted and no 'reconstruction' is brought into the picture. The relevant statement, suggested by traditional phrasings of the scopal nature of *tout*-inversion, is that SQ links entail scoping (while Q links do not). Put differently, the generalization is not about which merger 'reconstructs' but rather about which merger inherently entails scope: SQ does but Q does not. If so, it automatically follows that the highest SQ-link will appear to be the lowest possible 'reconstruction' site for successive remergers of a given node.

The Round Robin Derived from *Relativised Minimality* on Merge

Recall the round robin constraint:

(368) $* \alpha_j \; \alpha_i \; \ldots \; \alpha_i \; \alpha_j$

The puzzling aspect of this constraint was that it referred to chains, or chain-links, in an unclear way. Now that we understand the underlying structural nature of chains, the constraint involves constituents rather than chain links, and becomes manageable.

The corresponding legal phrase structure is:

(369)

B

A

B

A
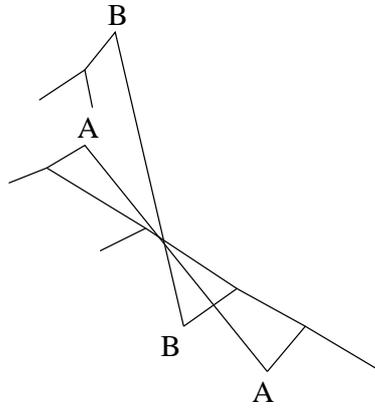
and the robin-violating configuration is:

(370)

A

B

B

A
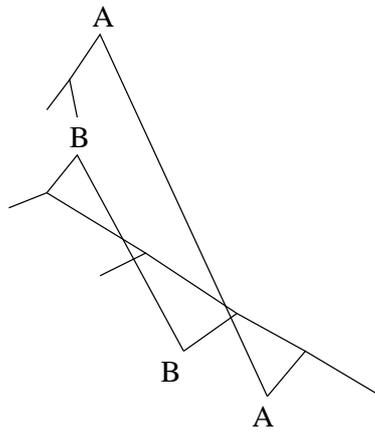
And the traditional minimality violation is:

(371)

A

B

A
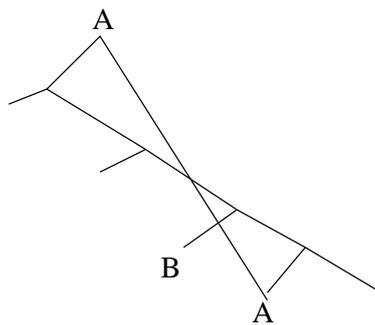
which is a proper subset of the felicitous (369).

150

Given these configurations it is obvious that the illegal remergers cross over the intervener, while the legal remerger does not. This can be stated either in terms of domination or in terms of c-command. Using domination, in (370-371) A contains B in that the remerger of A dominates all mergers of B, while in (369) the remerger of A does not contain (dominate) all mergers of B. In terms of c-command, the lower B is not an intervener in (369) because it c-commands the topmost A.

Let us formalise the notion of 'dominates all mergers' needed for the domination approach:

(372) α X-includes β iff α dominates all X-mergers of β

(Notice the similarity with the notion of 'segments'/'inclusion' traditionally associated to adjunction.)

The two potential formalisations of locality are then (373) in terms of inclusion, and (374) in terms of the traditional notion of c-command; with the relevant difference highlighted:

(373)   X-merge(α, β) iff

α, $\beta \in$ X, and

$\neg \exists \gamma$, $\gamma \in$ X, α **X-includes** γ and γ c-commands β.

(374)   X-merge(α, β) iff

α, $\beta \in$ X, α dominates β, and

$\neg \exists \gamma$, $\gamma \in$ X, α dominates γ, γ c-commands β, and γ **does not c-command** α.

The definition of the intervener in (374) is somewhat cumbersome, although intuitively clear: γ is not an intervener if it c-commands both elements between which it is supposed to intervene. To bring this out more clearly, let us define a notion of 'proper domination':

(375)   proper-dominates(α, β) *iff* dominates(α, β) & $\neg$c-command(β, α)

The formalisation (374) can now be streamlined to (377). The inclusion-based formalisation (373) can also be streamlined, to (376), which simply states that the closest X included in α is attracted. The two real contenders are thus:

(376)   X-merge(α, β) iff

α, $\beta \in$ X, and

$\neg \exists \gamma$, $\gamma \in$ X, α X-includes γ and γ c-commands β.

(377) X-merge(α, β) iff

α, β $\in$ X, and

$\neg \exists$ γ, γ $\in$ X, α properly dominates γ, γ c-commands β.

Which alternative is correct? There is one crucial difference which makes (376) more adequate: inclusion is a relative notion but c-command is not.

Take the case of (369) again, but suppose that the A remerger is not the same kind of merger as downstairs:

(378)   *
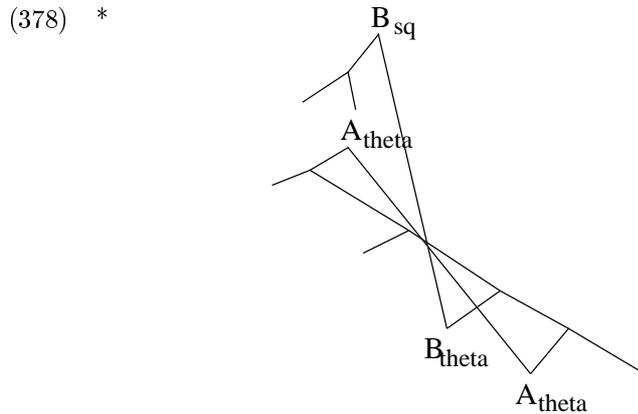
Here B has 'forked' to SQ-style mergers. The round robin now rules out the structure because A has prolongated θ-movement without B. (378) forms a minimal pair with (369): they are identical geometrically, but the nature of the highest B node triggers differing grammaticality.

In the c-command based definition, (377), the distinction is not made: both (369) and (378) are predicted to be good since the lower B always c-commands the higher A and thus never qualifies as an intervener. The definition based on inclusion does however make the correct distinction: once the higher B is of a different type, A does include all θ-occurrences of B, and locality is violated in (378) but not in (369).

The underlying difference between c-command and inclusion is that c-command is a purely geometrical concept which cannot be naturally upgraded to be sensitive to content, while inclusion is natural to state in relativised terms. What the round-robin constraint tells us is thus that (376) is a more correct formalisation of locality (as mentioned above, the asymmetry of the operation remains to be stated).

The remerger B in (369), repeated below as (379a), is thus legal because a does not count as an intervener, not being properly included. What about the remerger A? Why is it not blocked by the boldfaced merger B? At first sight it should be blocked since b is properly included in the remerger A.

(379)   a.   b.

Contrast this with the structure violating locality, (379b). Again *b* is included in A, but now it does count as an intervener. What is the difference?

The difference follows from the locality definitions above: the intervener must c-command the lower term of the merger. No B node qualifies as an intervener since the only relation they can bear to *a* is domination, not c-command. The only b-related node which can in principle c-command *a* is *b*. What is relevant is thus the relative status of *a* and *b*. As expected, *b* c-commands *a* in (379b) but not in (379a). Hence the asymmetry.

One case requires special care: successive cyclic remergers, as in (366). Why doesn't the intermediate Q remerger count as an intervener for the second re-merger? One possibility is to complicate the definition of intervener with a clause stating that the intervener must be distinct from α, β being merged. But a simpler solution is to define domination non-reflexively: if a node does not dominate itself, it follows from the definition of inclusion that no node will ever include its previous mergers (of the same kind). The highest Q projection in (366) therefore does not include the intermediate Q remerger and the latter thus does not qualify as an intervener.

The full round-robin effect thus follows from a simple restatement of the anti-identity principle in terms of domination/inclusion:

(380)   X-merge(α, β) iff

        α, β $\in$ X, and

        $\neg\exists$ γ, γ $\in$ X, α includes γ and γ c-commands β.

This formulation in turn follows from the fact that the higher term of the merger is a node dominating the lower term (which of course follows from the unification of merge and move).

Representations and Derivations Again

An important consequence of this result is that it requires a representational approach: in (379a), it is only after A has remerged that the 'earlier' remerger of B becomes legal. To the extent that this explanation of the round-robin

153

generalization is correct, we thus have some reason to think that the constraints on merge are representational (i.e. apply to the structure resulting from the last step of the derivation) rather than derivational. (In Chomsky's terminology, this suggests that the locality component of merge is itself a 'bare output constraint', leaving syntax to be a trivial and unrestricted concatenation engine.)[29]

Although the strong conceptual argument distinguishing a representational from a derivational approach to linguistic constraints was lost, two empirical arguments surfaced: the contention of Brody (2000a) that language never uses the ordering possibilities afforded by derivations, and the locality-based explanation of the round-robin constraint.

---

[29]Notice that this conclusion is not necessarily true in a top-down derivational system, an intriguing aspect to which I hope to return.

9 Purifying Merge of its 'Specifier' Heritage

For all the appearances of sharp turns, current Principles & Parameters theories contain layers of (tacit) assumptions gradually accumulated over the last 30 years. In some cases these strata play well together, building on top of each other. But in many cases older sediments survive only by historical inertia and their interaction with newer layers obfuscates the resulting theories. This final section aims to clean up one such case: the arrival of assumptions about 'functional projections' over the last 15 years has made older principles about the organisation of phrase structure obsolete, but these obsolete assumptions - traditionally called 'xBar theory' - still survive in various guises, leading to unnecessary complications, loss of insight, and redundancy. In short, xBar theory (in any contemporary variant) is a relic that holds us back. Cleaning up the theory of projection will in turn allow us to express the core 'merge' operation cleanly.

Any theory with 'functional' projections contains these two kernel assumptions (possibly as theorems):

(381)   there exists a concatenation operation (e.g. 'merge').

(382)   there exists an 'fseq' - a sequence of functional projections - such that the output of (381) must respect fseq.

Assumption (381) merely states the core of syntax, regardless of the theory adopted. Assumption (382) spells out the tacit agreement that - say - C is above T but not vice versa.

On the basis of the kernel assumptions we would expect structures of the type:

(383)   CP
         /\
           TP
          /\
            vP
           /\
             VP
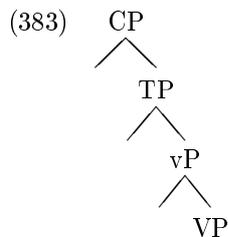
(modulo n-arity of the concatenation operation, which is irrelevant to the point to be made here)

But current theories postulate different - and more complex - structures:

(384)   CP
       /\
       CP
      /\
       TP
      /\
       TP
      /\
       vP
      /\
       vP
      /\
       VP
      /\
       VP
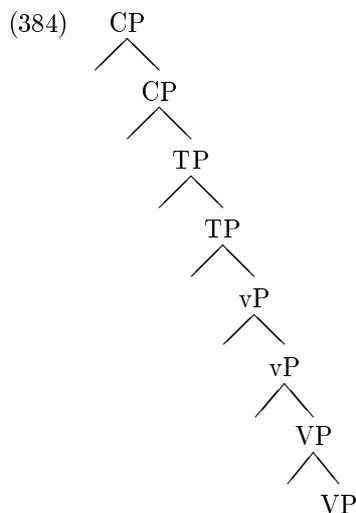
i.e. $[_{f3}$ f3 $[_{f2}$ f2 $[_{f1}$ f1]]] becomes $[_{f3}$ f3 $[_{f3}$ f3 $[_{f2}$ f2 $[_{f2}$ f2 $[_{f1}$ f1$[_{f1}$ f1]]]]]] as each item within fseq is allowed a second shot at concatenation ('2nd-merge', 'specifier') with a set of idiosyncratic properties associated to this second shot (requirement on bar level, agreement/checking relationships, linearity correlates, etc.).

What principle is added on top of (381-382) so as to require the more complex structure (384)? Answers vary widely, but amount to:

(385)   every XP within fseq (i.e. specifiers or adjuncts) must be locally associ-
        ated with a corresponding head.

which is in turn taken to be a reflex of a deeper (tacit) assumption:

(386)   Asymmetric Projection [AP]: a feature f in an XP node cannot legiti-
        mate its mother (it cannot 'project'), but the same f in an X° node can le-
        gitimate a maximal projection (it can 'project')

The surface symptom of Asymmetric Projection is the appearance of the specifier-head-complement triad: only theories with AP have the triad (theories without AP only have a diad: projectingElement-complement, as in (383)). All modern (Principles & Parameters) versions of phrase structure theories have the triad, and thus tacitly add AP to their stock of assumptions.

The natural question to ask is thus: why is this additional apparatus postulated? I.e. what does this additional apparatus do, that the kernel (381-382) does not do? And since the presence of specifiers and heads *is* the hallmark of 'xBar' theory (a misleading term given that the 'bar' node is absent in most current renderings), the question boils down to: why do we need to add xBar theory on top of the kernel assumptions?

156

Given that the immediate consequence of (385-386) is the need to postulate a null head for each attested specifier, an equivalent question is: What do we gain by postulating a null head accompanying each syntagm (specifier) in fseq? Here is a list of the grievances (developed below) suggesting that we gain nothing, and lose some:

- in current theories, Asymmetric Projection has only one use: to serve as a building block for the theory of triggers of merge/move (cf. 'checking', which builds on the availability of both the spec and the head). But since the kernel assumptions already entail a theory of displacement (and an arguably more insightful one), Asymmetric Projection is redundant, it has no use.

- Asymmetric Projection induces the unnecessary cost of having 90%+ heads be null

- Asymmetric Projection induces the unnecessary cost of duplicating every feature (once in the specifier, once in the head)

- Asymmetric Projection induces the unnecessary cost of having a 'double-articulation' of phrase structure: nodes are first assembled into xBar units, and these units are then assembled into an fseq, with no clear relationship between these two layers.

- Asymmetric Projection wreaks havoc with locality: from the point of view of locality, the spec-head relation is identical to a chain relation (i.e. identity), while - in a tree-with-trace syntactic representation - it has very different locality properties. Within such approaches, Asymmetric Projection forces us to complicate our approaches to locality (this point does not hold in the model developed in the preceding section, which made 'movement' an instance of a spec-head relation).

- extending the spec-head relation to all projections is based on the hope that the following (massive) generalization is true: every feature occurs in two nodes, these two nodes are in a spec-head configuration and undergo agreement (i.e. every feature will behave like a ϕ-feature). Embarrassingly, a decade of research has failed to turn up a single convincing case of such an agreement (outside of the starting point, ϕ-features), suggesting that the whole edifice might rest on a wrong generalization. (Even more embarrassing are the recent doubts that the paradigmatic case, ϕ-features, is a canonical case of spec-head agreement at all).

9.1 Doing Away with Specifiers

Insertion without the specifier-head relationship

Currently, the only job of (385-386) is to trigger insertion (merge/move). This is based on the fact that if (385-386) holds, every insertion of an XP into phrase-

structure creates an invariant local configuration between that XP and a corresponding head (the spec-head configuration). It is then trivial to equate the cause of the insertion with the need to create that configuration, which is what all current approaches do (cf. 'checking', 'criteria', etc.).

Should there be some other cause for insertion however, (385-386) would *ipso facto* become jobless and merrily walk into retirement. But there *is* another trigger for insertion: (382), the fseq requirement. In the absence of (385-386), each feature is represented only once (as depicted in (383)) and therefore every insertion (or failure thereof) will alter the fseq. Since (382) regulates possible fseqs, (382) will dictate what can (and must) be inserted, triggering merge/move.

To illustrate, take (383), where fseq = <C, T, v, V>. Now suppose the subject had not moved from its v-position (and no expletive was inserted): the result would be an illegal fseq, <C, v, V>, violating (382). (382) - i.e. fseq - thus forces the subject to move (or an expletive to be inserted), so as to project T (or Agr, or I) and obtain a legal fseq.

To repeat the logic: under the raw kernel assumptions any application of merge alters fseq, and is thus regulated by (382). Given that the kernel itself regulates insertion, (385-386) becomes superfluous. Since (385-386) *is* xBar theory's contribution to phrase structure, this means that xBar theory is unnecessary.

In simple terms, the arrival of functional projections has made the distinction between heads and specifiers obsolete: 'second merge' is unnecessary.

The only structural configuration needed is the 'head-complement' configuration. This entails that the 'head' of a projection can be either a terminal or a non-terminal:

(387)   a. αP
                    $\wedge$
              α βP

        b. αP
                    $\wedge$
            αP βP
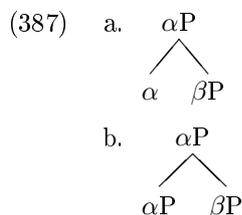
These two configurations are now identical, they both turn a β-terminated fseq, <β, ... >, into an α-terminated fseq, <α, β, ... >.

A simple illustration of this is the case of an interrogative CP: it is an instance of (387a) when headed by *if*, and an instance of (387b) when headed by a wh-phrase:

(388)   a.   VP
                  $\wedge$
             V   QP
                      $\wedge$
                if$_{[+Q]}$ TP

b.

```
          VP
         /  \
        V    QP
            /  \
        DP[+Q]  TP
         /  \   /\
        D   NP etc
```
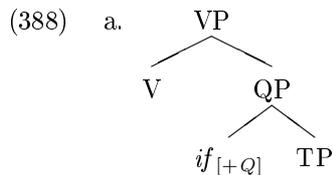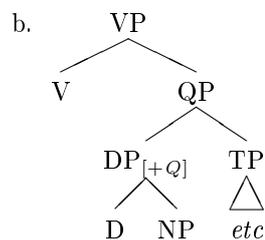
The point to notice is that (388) does not involve any 'Q°' terminal. The $DP_{[+Q]}$ directly projects QP since we threw away the stipulation that XP cannot project.

The Downsides of SpecHead

Retiring xBar turns out to be a positive result on many counts other than simplification of the theory by dropping unneeded assumptions: there is a substantial number of downsides to xBar that we don't have to live with anymore. Let us briefly review these before coming back to investigate the fseq-based alternative to the spec-head relation.

Specifier-head agreement does not extend beyond ϕ-features

The idea that 'specHead is special' (in providing the basis of agreement-like relationships) is a wild empirical bet: it takes the configurational treatment of ϕ-feature agreement as its starting point (e.g. Kayne (1989)), and bets that all features will behave similarly to ϕ-features. If this turned out to be true, we should be finding cases of local duplication of features in droves. To be more precise, what we expect to find are cases where two elements are each independently characterized as +f, and they (underlyingly) occur in the order *...XP[+f] X°[+f]...*, as in the prototypical ϕ-feature case. But such cases are exceedingly difficult to find, if existent at all. To illustrate, consider a number of representative cases, in fact some of the *a priori* most promising phenomena in favor of agreement: CP of the English type, NegP of the French type and yes/no questions in Russian.

Under a simple analysis, the English CP is the locus of two alternations: *that/if* in the head and *wh/⊘* in the spec (setting aside the null variant of *that*). Both alternations seem to be about the same feature, [±question], as expected, and spec-head agreement can proceed. Two sets of facts indicate however that this simple picture is inaccurate. First, languages which realize both the would-be head and the would-be specifier (such as Canadian French) get it wrong: they realize the [+question] specifier with the [-question] head, i.e. *... what that...* .. instead of the expected *... what if...* , thereby casting serious doubt on the presence of wh-agreement in the postulated null counterparts (there seems to be only one case instantiating the expected pattern, Dutch, which however has a

159

complicated and unclear pattern involving multiple complementiser particles.).
Second, more careful analyses of these heads suggests that the *that*/*if* contrast is
illusory: the two heads occupy different positions and thus do not form a featural
opposition at all (Rizzi (1999)). Both of these developments lead to approaches
with no spec-head agreement wrt [±question] in this construction. What we
need to explain instead is the complementary distribution of the *that*-head and
the *if*-head, and the complementary distribution of *if* and a wh-phrases. No
available theory addresses these, to which we'll come back below.

Or consider Russian yes/no questions. A particle (*li*) appears in the CP area;
and it is preceded either by the verb in 'neutral' questions (corresponding to
pseudo-examples such as *saw-li he John*, 'did he see John?'), or by an XP which
becomes the target of the questioning (i.e. *John-li he saw*, 'is it John that he
saw?'). A simple analysis of the latter case has the target XP in a spec-head
configuration with *li*, whereby it is interpreted as the questioned element. No-
tice however that the XP is not independently interrogative, it does not contain
any question marker, so that the construction is compatible with a spec-head
idea but does not provide any evidence for it. Furthermore, there is an indepen-
dent motivation for the anteposition of the XP independent from the would-be
[+question] agreement: *li* is a clitic of the C2 type and as such requires either a
verb or an XP to be anteposed. Finally, the focalised semantics of the anteposed
XP cannot come from a spec-head agreement wrt [+question], and suggests that
the XP rather occupies a different CP-related position: FocusP. This case, while
straightforward, is illustrative of a large class of contexts analysed as specifier-
head agreement; cases in which one of the two elements is not independently
marked with the relevant feature. All such cases are neutral with respect to the
present topic: they are compatible with a specifier-head approach, but provide
no evidence for it.

French negation looks like a promising candidate: it involves two syntactic en-
tities, both of which are independently characterised as negative. At first sight
the underlying order is the expected *XP X°* order, the X° surfacing to the right
of the XP as an artifact of verb movement: *ne-V pas t_v*. Everything seems set
for a NegP with *pas* in a spec and *ne* in a head. As widely observed though,
this won't do, since the above description only holds for finite verbs (and im-
peratives). Infinitives have a *ne pas V* order (*ne pas lutter contre l'injustice est
... * 'ne not fight against injustice is ...'), an order which directly contradicts the
idea that the rightwards position of *ne* is an artifact of verb movement, thereby
canceling a generalized underlying-NegP analysis. What we seem to be looking
at here are two independent occurrences of negation within the French clause
(perhaps along the lines of Zanuttini (1991)). Again, a promising case turns out
to be orthogonal to the issue.

Cases such as the three above being among the best hopes for extending the
'agreement' idea, it is striking that none turns out to provide any evidence, let
alone solid evidence for agreement, and this despite more than a decade of work
along these lines. Given this lack of success, it is a welcome result that the kernel

assumptions make xBar theory obsolete: this allows us to keep the positive property of triggering insertion, while abandoning the unlucky 'agreement' bet.

Unnecessary costs tied to R(spec, head)

Another welcome consequence of retiring xBar assumptions is a series of simplifications in the model. First, and maybe most obviously, there is no more need to postulate a null head accompanying each XP ('spec'). This result is particularly significative once the logic of the 'functional projections' approach is followed through, as in Cinque (1999) and related work: it then becomes quickly clear that more than 90% of heads are null, and more importantly, it becomes plausible that every head postulated on the basis of Asymmetric Projection (xBar) is null. Of course, there is nothing wrong in principle with null heads, regardless of their quantity. But just as obviously, a model achieving the same results while dispensing with them is preferable. Which is just what an fseq-based model promises to do.

Another side effect of the same move is to stop duplicating every feature: under Asymmetric Projection every feature present in an XP must be present twice (since a corresponding (local) head must bear it too). Again, once the logic of functional projections is followed through, the scale of the issue becomes non-trivial. And again, an fseq-based approach does away with this unwelcome property of classical phrase structure.

Yet another dubious side-effect of xBar theory is to split the theory of phrase-structure into two layers: in a first layer, syntactic nodes are assembled into 'mini-structures' (xBar units) and in a second layer, theses mini-structures are assembled into an fseq. Contemporary investigations have been able to make very little sense of this, mostly leaving the issue aside for future scrutiny. Again, retiring xBar relieves us of this unwanted side-effect (by collapsing the two layers into one).

Finally, consider locality. The standard interpretation of locality is that it regulates identity: it defines the domain in which two nodes are allowed to be interpreted as having the same identity. To such a condition, the specifier-head relation will appear to be just another case of two nodes between which identity holds. Yet, the locality condition of the classical specifier-head relation is radically more stringent than any other attested locality (save head-movement perhaps, the status of which is currently unclear). To make a minimal example, consider a classical analysis of a wh-moved element: XP_{wh} $\text{X}°_{wh}$ ... t_{wh}. Given that the relevant locality principle cares about identity with respect to the [wh] feature, the relation between XP_{wh} and t_{wh} looks exactly the same as the relation between XP_{wh} and $\text{X}°_{wh}$. It then comes as a surprise that the two differ with respect to locality, forcing some complication onto the locality conditions in order to prevent a 'distance' specifier-head checking. Again, the whole issue disappears once xBar is retired (and no 'checking' configuration needs be enforced).

9.2 Formalising Merge and Locality

This means that we can dispense with one central component of contemporary syntax: the 'checking' relation, i.e. the spec-head relation. More generally, we can dispense with specifiers.

Recall the logic of the fseq-based triggering of insertion of nodes into phrase-structure: given fseq = $<f_1 \ldots f_n>$, and a phrase-marker whose top node is F_kP (1<k<n), the next step/layer *must* consist of an insertion of f_{k+1}. The sheer need to comply with fseq drives merging operations.

Remerger - 'movement' - follows as a theorem: if a syntagm τ has two features relevant to fseq, f_j and f_k (j<k), and those features are not available from any other source, τ is first merged in the position adequate to f_j, but must be subsequently remerged at the position appropriate for f_k in fseq. In traditional terms, fseq forces τ to move.

A derivation of a simple declarative clause such as 'John always snores' is now as follows. Suppose fseq = <Agr, Asp, v, V>, or substituting 'v' with the more explicit AgentP, <Agr, Asp, Agent, V>, and suppose the DP has already been constructed. The first step must be a grouping of V and DP with DP projecting its [+agent] property, in order to obtain the sub-fseq = <Agent, V>:

(389)   vP = AgentP

```
        vP = AgentP
          /\
         /  \
   DP_agent   V
      /\
     /  \
    John
```
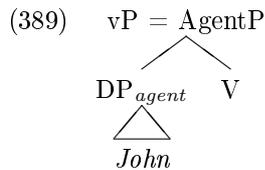
The next step necessarily consists in the merger of the aspectual adverb since this is the only way to comply with fseq:

(390)

```
              AspP
            /      \
           /        \
   AdvP_{+asp}       vP
                    /  \
                   /    \
           DP_{+agent}    V
               /\
              /  \
             John
```
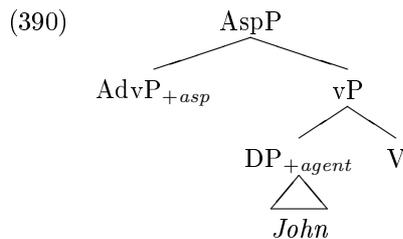
Finally, Agr needs to be provided (again, the exact nature of the label is irrelevant to the logic, substitute T, I, K, EPP, etc.), and this can only be done by remerging the subject and projecting the agr property of the subject:

162

(391)

AgrP
AspP
always
+asp
vP
SU
+agent
+agr
V
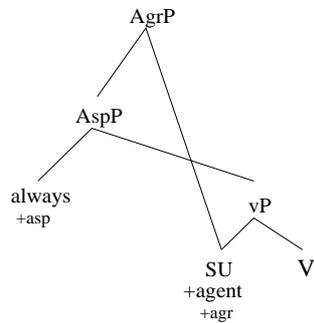
This reformulation is straightforward for such cases, but has a major consequence for locality: if a given features does not occur multiple times, locality cannot be a condition regulating identity. To see this consider wh-movement again. Instead of the earlier (358), repeated here as (392), we obtain (393).

(392)

Q
Q
Q T
V
DPwh

(393)

Q
T
V
DPwh
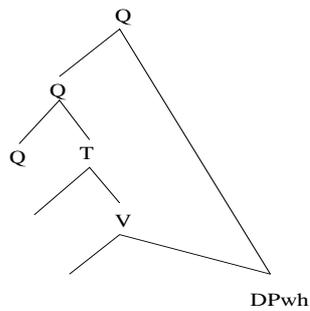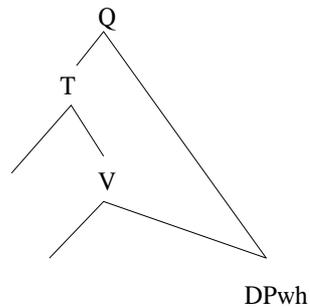
Instead of the earlier distance-grouping of the form *merge(Q, Q)*, (393) involves a distance-grouping of the form *merge(Q, T)*.

More generally, merge never applies under identity, since features are not duplicated. How is locality to be stated in this new situation?

Empirically, the transposed generalization is obvious: upon merging T with Q, only the nearest Q qualifies. A simple formalisation is:

163

(394) X-merge(α, β) iff

$\beta \in$ X, *and*

$\neg\exists$ γ, $\gamma \in$ X, α includes γ and γ c-commands β.

Where 'X-merge(α, β)' stands for 'merge α, β and label the result X, X a class in the feature hierarchy.

Notice that α-merge is now asymmetric: choosing $\{\alpha=Q, \beta=T\}$ versus $\{\alpha=T, \beta=Q\}$ will trigger radically different results. The correct choice is of course predetermined by the fseq constraint: given the nature of fseq, either α is a legal continuation of β, or β is a legal continuation of α, but not both (this also holds at the juncture point between two fseqs, as is easily verified).

In other words, locality does not regulate identity between the landing site and the departure point. It only regulate one of the two nodes, determining who is allowed to travel (i.e. remerge).

This raises the interesting question of the direction of the asymmetry: as stated in (394) only β-like elements qualify as interveners. That is, only elements that are of the same class as the item about to project. This expresses the intuition that remerging Q with T (moving Q above T) is blocked by Q interveners, not by T interveners. But logically nothing precludes the opposite definition:

(395) α-merge(α, β) iff

$\alpha \in$ X *and*

$\neg\exists$ γ, $\gamma \in$ X, α includes γ and γ c-commands β.

This issue of whether it is the moved element or the landing site that defines the class of relevant interveners was always present, but could not be addressed directly since both sites were seen as identical. Given an asymmetric view of *RM*, let us probe the issue more directly.

Successive cyclic remergers such as (366), now represented as (396), are for instance relevant cases:

(396)

164
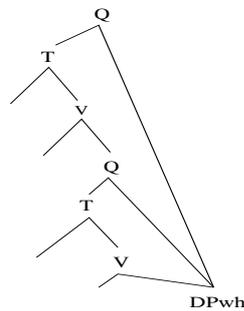
In such cases, the definition (394) yields the correct result: in the merger between the highest T and DP$_{wh}$ the only potential intervener is the intermediate Q, but the intermediate Q is not properly included in the highest T and thus does not qualify.

According to (395) on the other hand, any intervening T will block the merger. (396) does contain an intervening T, the T in the intermediate clause. This T is included in the T of the root clause and thus qualifies as an intervener. If (395) was correct, successive cyclic merger should thus be excluded, contrary to fact.

It therefore follows that it is the projecting element which determines the relevant type of intervener, as in (394). Accordingly, I adopt (394) as the final formulation of locality and merger, possibly the sole remaining syntactic principle.

9.3 The Nature of *fseq*

Given that the need to comply with fseq drives insertion into phrase-structure and displacement within phrase-structure, what is it exactly? The identity of the features within fseq is not paramount to the current line of reasoning (although it is of course of central importance to any current P&P approach, given the kernel assumptions). What rather needs to be known is how much variation fseq allows, if any. Let us distinguish two (traditional) questions:

- is there more than one distinct fseq? (where two fseqs are distinct if they are not in a subset relation)

- given an fseq, must all its features always be present, or are subsets legal instances of fseq?

Consensus seems to have developed around the most restrictive answer to the first question: there is one and only one legal fseq, universally (Starke (1995), Rizzi (1997), Cinque (1999), and numerous others). The second question is more controversial, with three types of answer repeatedly proposed: the 'rigid' approach - everything must always be present (e.g. Starke (1995), Cinque (1999)); the 'peeling' approach - projections can be missing, but only by peeling off from the top (e.g. Radford (1990), Rizzi (1994), Platzack (1996), Cardinaletti and Starke (1999)); or finally the *'laissez-faire'* approach - any projection can be missing (e.g. Wexler (1994)). No argument has been put forward one way or another (pace Cinque (1999), p.133).

Here are however two argument for the third, least normative, line. Consider what happens to NegP in positive clauses. According to the first two approaches, there must be a null 'assertion' operator paralleling the negative operators (i.e. NegP is a PolarityP). The [-positive] operators however have some distinct properties, among which the quality of inducing weak islands ('negative islands'). All

165

things being equal we would expect the [+positive] assertion operators to have the same property. Any value of PolarityP would thus induce weak islands, and hence all clauses should always be weak islands, a *reductio ad absurdum* of the idea of 'positive operators' in a PolarityP of every clause. Is there any evidence that this abstract reasoning is correct? Factives provide a direct confirmation: it is a traditional observation that verbal predicates whose semantics involves asserting that their complement denotes a fact (*regret, know*) induce island effects (the 'factive' island) while their non-fact-asserting counterparts do not (*believe, think*); an observation traditionally described with a factive (assertion) operator inducing weak islands. The approaches to fseq that force the presence of an assertion operator thus seem stuck with the absurd conclusion that every clause is a weak island. (In fact, things get worse as the same logic applies to at least [±question] and [±focus]. Each clause is thus predicted to contain at least three weak islands.).

The same logic is massively required by Rizzi's Modifier-movement. Recall that Rizzi (2001) postulates a third class relevant to *Relativised Minimality* to account for RM effects of one adverb crossing another. Rizzi's implementation rests on a Cinquean structure with multiple adverbial positions. Anteposition of any adverb - except the structurally topmost adverb - will always cross adverbial positions and will thus always be ruled out. A phrase-structure theory which entails the presence of all projections thus entails the absurd conclusion that only the topmost adverb can ever be anteposed.

We thus need a *laissez-faire* approach allowing a projection to be absent if no element in the enumeration requires it, both in order to allow for the absence of NegP/PolP in positive clauses, and to allow for the absence of unused adverbial positions. A radical implementation would replace (382) with:

(397)   there exists an 'fseq' - a unique sequence of functional projections - such that the output of (381) must be a subset of fseq.

But this won't do. (397) is both too permissive, allowing many unattested combinations, and expresses the wrong generalization - the generalization that 'anything can drop'. But that's not what the facts are telling us. The facts are telling us that there is some regularity behind 'who can drop'; and it is this regularity that the fseq-requirement needs to express.

As an illustration, take negation, questions and focus again: we never have any reason to think that [+neg], [+wh] or [+foc] are absent from the syntactic representation but interpreted semantically. Such cases - e.g. *plum tastes good* interpreted as *plum does not taste good* - indeed strike us as aberrant. (Cases where elements are phonologically null but syntactically present, as in the classical analysis of English null complementisers, are of course irrelevant.) The only corresponding cases that ever do suggest radical absence from the syntactic representation (with concomitant semantic interpretation) are [-neg], [-wh] and [-foc].

Building on the traditional observation that speakers have systematic and strong intuitions about 'markedness' of values whereby a declaration is 'unmarked' and an interrogation 'marked', asserting is 'unmarked' and negating 'marked', a flat statement is 'unmarked' and focusing 'marked'; the correct generalization seems to be:

(398)   (only) unmarked values can drop

How is this new generalization to be captured? We could simply stipulate it as a mild version of the *laissez-faire* approach. But that would miss a deeper step: the fact that only unmarked values can 'drop', entails that dropped items are semantically recoverable (i.e. absent nodes are necessarily interpreted as receiving their unmarked value). (398) thus amounts to a 'recoverability condition' in the vein of Vergnaud (1974): an element of fseq can be dropped only if it is recoverable at the LF interface.

This in turn entails that at the level of interpretation, *fseq* is always complete (having recovered all pieces missing in the syntactic representation via the markedness route). There is thus no need to hard-code the droppability of unmarked values within the fseq principle; it follows from the interaction of two simpler assumptions:

(399)   semantic representations must respect fseq
        i.e. there exists an 'fseq' - a unique sequence of functional projections - such that the output of (381) must be LF-interpretable as an instance of fseq

(400)   the interpretive component's reading of syntactic representations is based on recoverability: absence from syntactic structures corresponds to unmarked values

It thus turns out that both the rigid and the *laissez-faire* approaches are right: the rigid approach is correct as far as the interpretative component goes (i.e. everything must always be present); the *laissez-faire* approach is correct for syntax but as an artifact of how interpretation takes place.

As a second argument for the *laissez-faire* nature of the syntactic fseq requirement (and for the markedness-based formulation of that requirement), consider a massive and curious fact: only some values of a given feature trigger remerger (movement). In many languages a [+wh] feature triggers A'-movement to the CP zone, but in no language does a [-wh] feature trigger a similar movement; similarly in many languages a [+foc] feature triggers A'-movement to the CP zone, but in no language does a [-foc] feature trigger a similar movement; or again, some constructions involve A' movement of [+neg] phrases (such as negative inversion in English), but no constructions involve a similar movement triggered by a [-neg] feature. Why?

Again, the generalization seems to be about markedness: only 'marked' values ever move. But why? Why should the same feature be subject to movement under one of its values but not under the other? and why should the moved value systematically be perceived as the 'marked' value by speakers? A rigid view of fseq makes this mystery thicker: it postulates a functional projection corresponding to [-neg], [-foc], [-wh], ready to host movement, leading to the expectation that this projection will be used.

The *laissez-faire* approach on the other hand captures the contrast. Given our two assumptions that (i) features with marked values must be present in the syntactic structure, but features with unmarked values can be absent and (ii) the incentive for remerger (movement) is to extend fseq so as to comply with (399); it follows that unmarked features will never trigger movement. This is because fseq does *not* need to be extended in the case of unmarked features, and thus movement/remerger is superfluous. Marked features on the other hand correspond to projections that cannot be dropped, and thus fseq needs to be extended, through movement or other means.[30] [31]

(An interesting consequence of (399) is that the fseq-requirement is not a constraint of 'narrow' syntax. I will come back to this below).

Absorption

Given this view of fseq, consider the stipulation (251b) above: within the same robin, θ and θ_L do not mix. More generally, what a robin configuration consists of is a recursion of the same feature, and within such a recursion, the feature must have the same value.

In terms of fseq, this maps onto the fact that an fseq of the type $<\ldots \ \alpha \ \beta \ \gamma \ldots>$ can be legally instantiated by $<\ldots \ \alpha \ \beta \ \beta \ \beta \ \gamma \ldots>$, i.e. $<\ldots \ \alpha \ \beta^* \ \gamma \ldots>$, with all occurences of β interpreted as a single instance. If this is what underlies robins, it follows that each occurence of β must have the same value, in order to be interpretable as a single interpretive entity. Different values such as θ and θ_L therefore cannot mix.

When overt, these configurations correspond to the various cases of recursion of a feature (topic phrases in Romance, negative phrases in Bavarian, DP-scrambling in West Flemish, wh-movement in some Slavic languages, *tout*-movement in

---

[30] The same reasoning does not hold for the first merger under the common assumption that every element of the numeration needs to be integrated into the structure, regardless of the value of its feature(s).

[31] An attractive answer would be to adapt the approach taken by Kaye et al. (1985) for broadly similar facts in phonology: the unmarked value simply does not exist - ie features are unary (and presence corresponds to the marked value). Transposing: the reason why the unmarked value never moves is that it doesn't exist. Unfortunately this approach does not seem to be workable in syntax, since realised instances of unmarked values of syntactic nodes abound. Notice however that such an answer would amount to a *laissez-faire* approach to fseq: only marked (ie. existing) features can project.

French). When resulting from covert movement, such configuration correspond to what is traditionally referred to as 'absorption'.

(Notice that these configurations also correspond to the presence of a specifier. I.e. under the current approach, nothing precludes the presence of a specifer in principle; its absence simply follows from the fact that it is unnecessary for the well-formedness of the structure).

9.4   Apparent Problems

According to the fseq logic, the only syntactically relevant geometrical relationship is the complementation relationship linked to sisterhood, now understood as including the 'distant sisterhood' of remergers. This is because syntactic structures are nothing but raw layers of head-complement relationships, with each layer independent of the other.

There are however two types of unexpected relations which seem to occur. First, the agreement relation - though exceedingly rare - does occur with ϕ-features. A second type of case is illustrated by subject-verb inversion in root wh-questions: the insertion of one node (the inverted verb) depends on the insertion of another node (the wh-phrase).

Both of these unexpected dependencies are straightforward to state in terms of spec-head agreement but seem unstatable with fseq, as they do not involve head-complement relations.

Apparent spec-head agreement: ϕ-features

How do we express the fact that the morphology of the verb covaries with the morphology of the subject? Let us set our goals higher: since ϕ-feature agreement is the only credible instance of agreement, we not only want to explain how the agreement between the subject and the verb comes about, but we also want to explain why overt morphological agreement only ever happens with ϕ-features.

An obvious observation in this context is that there is one other phenomenon in grammar which again involves covariation in ϕ-features but in no other features: binding. Simple contrasts such as 'she washes herself' vs. 'he washes himself' (cf. *'she washes himself') are formally identical to subject-verb agreement: the morphology of two nodes in the structure covaries with respect to ϕ-features (but no other features).

Why then do we have twice the same phenomenon? A natural answer is to reduce one to the other: verbal morphology enters into a binding relation with the argument of the verb. This captures the similarity between the two cases and solves the fseq problem: the verb and its covarying argument are in different projections, with no direct geometrical relationship involved: the relationship

is established via binding which does not require xBar-type relations between the binder and the bindee. (This line of thought is reminiscent of approaches which hold that verbal agreement is 'pronominal' - e.g. Rizzi (1982). In that terminology, the present argument is that the 'pronominal nature of agr' allows it to enter binding relationships, and renders the idea of spec-head agreement redundant.)[32]

Apparent spec-head agreement: filling-in

What about inversion-type cases? Why does insertion of a verb above the subject 'depend' upon the insertion a wh-phrase - or a focus-phrase - above the verb, if the two occupy unrelated projections? To make things worse, the same problem arises independently of inversion: in some constructions, insertion of an XP requires prior insertion of a particle. Again two unrelated nodes depend on each other. To take an example, insertion of a focused XP in Gbe requires prior insertion of the particle *wè* (Aboh (1997)), and other such examples abound.

Notice however that there is a clear pattern to the issue: the insertion of a higher node is dependent on the prior insertion of a node immediately below it. Let us call this 'dependent insertion'.

As it turns out, dependent insertion is already expressed in fseq: the fseq requirement dictates that insertion of a given feature depends on prior insertion of all features preceding it in fseq. Temporarily ignoring the markedness issue, the logic is as follows: given fseq=$<\ldots \alpha \, \gamma \, \beta \ldots>$, attempts to insert α directly on top of β will create an illegal fseq and will thus fail. As an example, with fseq=$<$C, T, V$>$ attempting to merge C directly with VP will fail, since T is missing from the sequence. The same logic applies to subject-verb inversion: if α is the projection of the wh-phrase, and β the complement of the inverted verb, not inverting the verb creates $<\ldots \alpha \, \beta \ldots>$, an illegal fseq. Dependent-insertion thus amounts to "fseq-filling".

But given the markedness-based reasoning above $<\ldots \alpha \, \beta \ldots>$ *is* a legal instance of $<\ldots \alpha \, \gamma \, \beta \ldots>$, provided that γ instantiates an unmarked value. An "fseq-filling" explanation thus predicts that dependent insertion will only occur when α requires the *marked* value of γ. As an example, suppose that in subject-verb inversion the verb occurs in Q° (a plausible hypothesis since inversion is found with negation, focus and wh). In that case, the generalization is that projecting whP requires prior projection of a QP-like projection. The filling-in phenomenon is now a consequence of the fact that +wh requires +Q rather than -Q. This type of feature-value dependency is natural given a Cinquean refined

---

[32] There is an interesting question as to why some languages allow binding (agreement) only in some positions - e.g. Arabic is noted for having the verb covary with the subject only if the subject precedes the verb, and similar facts hold of French participles wrt. an object. This is the (independent) general question as to why coreference is possible in some structural configurations but not others, ie. the question as to the nature of Binding Theory.

fseq: if tense is split among several projections for instance, a value +future is 'dependent' on a value +tensed, etc.

Finally, notice that this explanation of dependent-insertion makes different predictions with respect to the standard spec-head approaches: the spec-head approach entails that the only case of dependent-insertion ever found is the case of a YP dependent upon the prior insertion of exactly one X°. The fseq-approach on the other hand has no such entailment: insertion of a YP can require prior insertion of any number of elements, depending on how many fseq items are missing between the insertion site and the top node of the tree to be extended. The elements inserted to fill-in fseq can furthermore be either heads or phrases, while it can only be a head in a spec-head approach. The fseq prediction seems correct. Cases with more than one element filling-in are provided by the Scandinavian wh-constructions where *two* particles ('complementisers') appear between the wh-phrase and the rest of the clause. A case in which a maximal projection is the filler is provided by Jamaican Creole focus structures, in which a particle *precedes* the focused phrase in the CP area (Durrleman (2000)), i.e. a head triggers an XP-filler.

The EPP

Another type of problematic case is the EPP: the EPP states that one particular projection in fseq is an exception in that it can never be dropped. Why?

Contrary to the cases above, this problem is not particular to the fseq approach. Traditional theories are confronted to the same issue and need to appeal to stipulations such as asserting that a particular feature has the property of being 'strong', or simply stating the facts through an 'EPP' feature.

The analysis of ϕ-agreement in terms of binding however opens up a path for explanation: if verbal agreement enters a binding relationship with the argument, the relationship is either pronominal or anaphoric. Suppose it is anaphoric. It then follows that a local antecedent must be present. Supposing that it is a general property of language to have anaphoric agreement on the verb (often null), the EPP follows.

(This opens up intriguing possibilities for the well known correlation between null subjects and rich agreement. We know that anaphors are morphologically poor, accordingly languages with poor verbal inflection qualify as anaphoric-inflection languages and thus need a subject to bind the inflection. Languages with rich inflection on the other hand qualify as pronominal and may thus escape the EPP requirement through the very morphology of the verbal agreement.)

Guilty of Duplication?

One potentially serious objection to the fseq line of reasoning is that it runs the risk of doubling the number of features in fseq. Take any projection XP, and suppose we genuinely find a situations where - in traditional terms - both 'specXP'

and 'X°' are filled; in such a case, fseq would postulate that the 'spec' and the 'head' in fact project two different features, thereby replacing the traditional feature sequence '... α α ...' by '... α β ...' (as in the so called doubly-filled-comp cases which upon inspection turn out to be two different projections, Rizzi (1999)). Now suppose that for every XP, there exists some construction in some language which has both the 'spec' and the 'head' filled. The net result is that the fseq approach would double the size of fseq. Redundancy is cut, but at a cost.

I'll argue below that we do not need to pay this price, but for the sake of the argument suppose that we do need to pay it: just how steep would it be?

While it is certainly unpleasant to double the size of fseq, the loss incurred pales in comparison to the simplification of the theory afforded by the elimination of the specifier-head machinery.

The core asset of the present proposal is to replace a more complex and more heterogeneous theory by a simpler and more homogeneous counterpart through the elimination of the distinction between heads and specifiers: The older - traditional - theory takes 'merge', adds Asymmetric Projection (xBar theory) on top of it, grafts an agreement/checking mechanism on top of this, and then proceeds to leave fseq lying around in a dusty corner, barely used. In doing so, it ends up with merge+fseq+AsymmetricProjection+checking, a curious bag of heterogeneous tools.

In contrast, the new contender only has merge+fseq, which form a homogeneous pair: one creates an object, the other decides the type of the object just created. The newer theory is simpler in that it is a proper subset of the old theory, but maybe more importantly it leaves us with a more homogeneous state of affairs. This is the main gain, supplemented by a number of side-gains: no duplication of features, no gratuitous null heads, no bizarre locality, no two intermixed theories of phrase-structure, no embarrassing mispredictions as to agreement.

The objection we are considering - if it went through - would essentially tell us that the "cutting duplication" side-gain is non-existent because it is offset by the cost of doubling the number of syntactically relevant features. Maybe, but this would shoot down a footnote rather than the main text. The appeal of the core proposal remains intact. [33]

But arguably the objection doesn't hit the footnote. The underlying issue is: is it true that the 'doubly-filled comp filter' extends to all projections (generalizing

---

[33]One might be tempted to claim that another loss would be incurred if fseq turns out to double the size of fseq: we don't cut the amount of empty heads which only serve the purpose of satisfying Asymmetric Projection anymore, since those positions are created again to account for the putative doubly filled projections. This is however not accurate: as elements of fseq can be - and are - dropped via recoverablity, not all of fseq is expressed in any given syntactic structure. The question is thus: does an enriched fseq lead to syntactic structures with those null heads again? No: it is a traditional observation that 'null heads' overwhelmingly express unmarked values of features. As discussed above, features with unmarked values need not appear in syntactic representations, so that no null head will be generated for these cases by a putative increase of fseq.

the spirit of Cheng (1997)) or is it the case that the 'filter' is generally violated - an assumption which the objection rests on?

To address this, let us start by distinguishing two types of terminals: particles and verbs. These two types of terminals appear to have very different properties.

First, particles do not move, while verbs do (or appear to). We never find particle-inversion in contexts such as '... τ XP PRT ...' (where τ is a 'head position' in need of a filler), *'... PRT XP t ...'. What we do find is insertion of another particle or verb, '... PRT2 XP PRT ...'. With verbs on the other hand, we routinely find inversions in '... τ XP V ...' contexts, resulting in '... V XP t ...'.

Second, the (fixed) insertion position of particles has a clear semantics correlate wherever the meaning of the particle can be determined independently, whereas the position of verbs has lost any correlation to meaning or form as inquiry about their syntax has progressed - as detailed below. Let us call these two types of heads *rigid* heads (no movement, semantically motivated) and *soft* heads (movement, no semantic correlate).

Rigid heads appear to behave as the current approach predicts (again limiting ourselves to those particles whose meaning can be independently determined): they do not allow any corresponding specifiers around them. Consider for instance the Gun 'aspect markers' (Aboh (1997)), a set of particles indicating properties such as imperfective, prospective, habitual. They occur preverbally in a strict 'head-initial' context, with a 'habitual > progressive > prospective' order. The straightforward expectation of a spec-head (agreement) theory is '... Adv_{hab} hab-prt $\text{Adv}_{progress}$ progress-prt Adv_{prosp} prosp-prt ...'. But such structures are impossible, eg:

(401)   * Assiba hwehwe no ton
         Assiba often   hab go-out
         Assiba often goes out

This situation seems representative of particles in general, once one takes into account cases which were thought to involve spec-head relationships but were found to be in distinct projections on independent grounds (cf. the Russian *-li* and the English *if* mentioned above). [34]

The potential duplication issue then only arises with soft-heads, i.e. lexical heads. As Cinque (1999) documents, for every adverbial position, there is a (head-initial) language that can position a verbal head to the immediate right

---

[34] There are cases of apparent spec-head agreement between an XP and a particle immediately following it, such as the Gun case of 'topic'-particles which occur directly after a topicalised syntagm. But all such cases are precisely cases where we do not have any independent means of determining the meaning of the particle (eg. the Gun topic particle never occurs by itself, or outside of topic contexts, so that it is impossible to tease apart the meaning of the topicalisation itself from the meaning of the particle).

of that adverbial, thus producing a configuration where both the specifier and the head are filled. Since projections with adverbs in their 'specifiers' span the majority of projections, apparently an fseq approach must cut almost all features into two independent features.

But given this reformulation, the doubling-fseq issue is a side-effect of an independent problem: why is it that verb movement has such widespread optionality as it has according to Cinque's system, with no correlation with either semantics or morphology? why are the boundaries of the 'optionality zone' of verb-movement arbitrary (again, no semantic or morphological correlate)? Why is even the correlation to verb-type lost: infinitives 'move more' than tensed verbs in some constructions, but less in others? In short, the descriptive situation is that verbs can be inserted anywhere, and the language-particular choice of the potential landing sites is entirely arbitrary. It is only because verb-movement turns out to have such bizarre properties that the issue of potential fseq duplication arises.

There are however two regularities that make the situation less desperate than seems at first sight.

First, not all 'verb-movement' is bizarre: subject-aux inversion in questions, V2 antepositions, etc. do not display the above bizarre properties. Such anteposition do not result in optionality, and are correlated to interpretation. In short, such antepositions behave like rigid heads. It is thus only the first, 'default', step of verb-movement which behaves oddly.

Second, verbs are not the only category with these bizarre symptoms. Agreement and negation have the same syndrome: "the evidence points to the possibility of generating a NegP on top of every adverb-related functional projection" Cinque (1999):126, and a similar conclusion is reached by Cinque for "DP-related" projections, i.e. AgrP.

The second observation provides a basis for a solution to the fseq-doubling issue given the fact that verbal forms canonically include an agreement morpheme. Since the agreement morpheme is typically the outermost morpheme of the verb, the mirror generalization and its standard syntactic consequences (Muyskens (1981), Baker (1988), Pollock (1992), Brody (2000b)) entail that the verb will be displaced into the position of Agr. The end-position of normal verb-movement will thus be the arbitrary position of Agr-insertion.

It then follows that the random-insertion of Agr will appear to be a random-insertion of V. In other words, the bizarre character of 'verb-movement' reduces to the bizarre character of Agr-insertion.

Reducing verb-movement to the Agr-insertion issue also provides an explanation of the first generalization above: there is a difference between the first step ('IP')of verb-placement and the higher ('CP') steps because only the first lower part of the trip is conditioned by Agr-insertion. Any further step such as subject-verb inversion is triggered by non-Agr features, with no random-insertion effect.

In terms of fseq, what this all reduces to, is that Agr° has a liberal positioning within fseq - each language apparently making an arbitrary choice - and the verb ends up in that apparently arbitrary position. No fseq doubling required, in fact no fseq growth at all.[35] [36]

---

[35]Notice that the technology of verb 'movement' is used here only for ease of use. I am in fact assuming the more elegant technology proposed by Brody (2000b) whereby no actual movement of the verb occurs; but no difference arises with respect to the above reasoning.

It also goes without saying that allowing language-dependent choice of Agr-positions is a descriptive stopgap at best, which needs to be understood and reduced to more plausible constraints.

[36]A radical solution along the same general line is to allow insertion of the verb in any position in fseq, discarding the traditional idea that verbs are generated in the lowest projection. This would also make verb-optionality parallel to neg/agr-optionality. I will not pursue this line, but I do wish to note that there is surprisingly little evidence for the common assumption that verbs must be generated very low.

10   Overview

What we have been doing, in its essentials, is to examine two core redundancies in current theories, eliminating them and examining the properties of the resulting model.

The elimination of the two core redundancies in turn builds on the observation that current syntactic models include two hidden assumptions (whose removal makes the redundancies disappear). The two hidden assumptions are:

(402)   a.   maximal projections cannot project (or 're-project'), although the potentially projecting feature within them can 'agree'. Only terminals can project.

        b.   the anti-identity locality principle (*Relativised Minimality, Attract Nearest*, etc.) applies 'horizontally', to sister classes of syntactic features, and not 'vertically' to subsets/supersets (subclasses/superclasses) of syntactic features even though a (simple) vertical organisation of classes is presupposed.

The heart of the first redundancy is the fact that all modern theories of syntax (implicitly) assume an ordering principle, stating facts such as C is above T, T is above V, etc. This principle is not given much attention, and because of this, another redundant constraint is postulated: the 'checking' or 'spec-head' agreement constraint. The only goal in life of checking (or spec-head agreement) is to trigger insertion of material into phrase-structure, either via merger or via movement. But the triggering of insertion already follows from the ordering principle: any insertion not compliant with the ordering will yield gibberish.

Allowing each inserted node to project however presupposes that merged maximal projections may project further (not necessarily wrt. the same feature), and thus abandoning the stipulation (402a). By the same token, the notion of 'specifier' disappears from the theory, as would-be 'specifiers' are in fact heads of their own projections. What remains is an approach in which syntactic structures are pure head-complement relations, with the 'head' either a terminal or a non-terminal.

This is a significant achievement by itself, not only because it allows us to preserve all major results of contemporary *Principles & Parameters* theories while dispensing with a central notion in those theories, but also because it shifts the burden of explanation exactly where we want it: away from a theoretically suspicious entity and onto the basic syntactic relation, *merge*. The theoretically suspicious entity is 'second' merge - or more generally the 'spec-x' relation - which is essentially a compositionally vacuous instance of merger: two elements merge under identity (wrt the projecting feature), thus not doing any useful work. In the trimmed down approach on the other hand, all explanation is shifted onto the concatenation of distinct elements into a new compositional superset; i.e. the head-complement relationship.

The second redundancy lies in the fact that the 'move' operation is essentially similar to the 'merge' operation, a similarity which is not captured in standard approaches. In most general terms, the similarity lies in the fact that both operations are 'grouping' operations: uniting two input entities into one output superordinate entity. The two operations are also similar technically: the locality of merge is a special case of the more general locality on move, and the 'labelling' of merge is identical to the labeling of 'move'.

To see this we however needed to go through an extensive trip into the locality of move: a potential identity between the locality of merge and the locality of move only arises if there is such a thing as a unified locality of move. But unifying locality has been a long-standing high-priority task eluding our grasp.

The decisive notion in the solution of this problem is what might be called 'vertical' *Relativised Minimality*: *Relativised Minimality* applied not only to distinct (sister) classes of features, but also to features ordered into a superset/subset (superclass/subclass) relationship. The standard formulation of *Relativised Minimality* already entails that syntactic features are organised into a feature-tree (albeit a very simple tree), but does not take advantage of this property, thereby in effect amounting to the stipulation (402b).

Eliminating (402b), and refining the feature-tree of syntactic features, allows us to unify a large array of otherwise disparate locality effects, thanks to the fact that the standard version of *Relativised Minimality* makes clear - and correct - predictions for the superset/subset cases. This unification of locality then reveals that the locality of merge is in fact the same as the locality of move, leading to a unification of the two notions. As a result, move is nothing else than an instance of 'merge' operating on two non-adjacent nodes.

Again, the result of this simplification is a shift of the explanatory burden away from a suspicious entity and squarely onto the shoulders of the core syntactic principle. The suspicious entity is 'move' - generally recognised as suspicious - and the core syntactic principle is of course merge.

These two explorations taken together have then led us to a strong conclusion: we can express the entirety of current syntactic theory by attributing to syntax only one core operation: 'merge'. No checking, no specifiers, no movement, no chains, etc. Merge, of course, is what syntax is all about; i.e. grouping unrelated tokens into one (compositional) unit. In other words, it seems possible to express all the results of current theories by simply iterating head-complement mergers, without indulging into the rich apparatus standardly thought to be necessary to achieve those results.

References

Aboh, Enoch. 1997. From the syntax of Gungbe to the grammar of Gbe. Doctoral Dissertation, University of Geneva.

Adger, David. 1994. Functional heads and interpretation. Doctoral Dissertation, University of Edinburgh, Centre for Cognitive Science.

Baker, Mark. 1988. *Incorporation. A theory of grammatical function changing.* Chicago: University of Chicago Press.

Baltin, Mark. 1992. On the characterisation and effects of d-linking: Comments on Cinque. In *Current issues in comparative grammar*, ed. Robert Freidin, 249–256. Kluwer Academic Publishers.

Bayer, Josef, Markus Bader, and Michael Meng. 1998. Morphological underspecification meets oblique case: Syntactic and processing effects in German.

Beghelli, Filippo, and Tim Stowell. 1997. Distributivity and negation: The syntax of each and every. In *Ways of scope taking*, ed. Anna Szabolcsi, 59–107. Dordrecht: Kluwer.

Boeckx, Cedric. 2000. Interrogative strategies in French: A reply to cheng and rooryck. Ms., University of Connecticut.

Bošković, Željko. 1997. On certain violations of the superiority condition, AgrO, and economy of derivation. *Journal of Linguistics* 33:227–254.

Brody, Michael. 1992. A note on the organization of the grammar. *UCL Working Papers in Linguistics* 4.

Brody, Michael. 1995. *Lexico logical form.* MIT Press.

Brody, Michael. 1998. The minimalist program and a perfect syntax. *Mind and Language* 13:205–214.

Brody, Michael. 2000a. Developments in elegant syntax. Blagoevgrad Summerschool.

Brody, Michael. 2000b. Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry* 31:29–56.

Brody, Michael. 2000c. On the status of representations and derivations. *UCLA Working Papers* 12:343–367.

Cardinaletti, Anna. 1997. Subjects and clause structure. In *The new comparative syntax*, ed. Liliane Haegeman, 33–63. London and New York: Longman.

Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: On the three grammatical classes. In *Clitics in the languages of europe*, ed. Henk Van Riemsdijk, 145–290. Mouton de Gruyter.

Cattell, Ray. 1978. On the source of interrogative adverbs. *Language* 54:61–77.

Chang, Lisa. 1997. Wh-in-situ phenomena in french. MA Thesis, University of British Columbia.

Cheng, Lisa. 1997. *On the typology of wh-questions*. New York: Garland Publishing.

Chomsky, Noam. 1986. *Barriers*. Linguistic Inquiry Monograph 13. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Venneman, 506–569. Berlin: Walter de Gruyter.

Cinque, Guglielmo. 1977. Towards a unified treatment of island constraints. In *Proceedings of the Twelfth International Congress of Linguists*, ed. Wolfgang Dressler and Wolfgang Meid, 344–348. Innsbruck, Austria: Innbrucker Beitrage zur Sprachwissenschaft.

Cinque, Guglielmo. 1990. *Types of $\overline{A}$ dependencies*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

Clements, George N. 1985. The geometry of phonological features. *Phonology Yearbook* 2:225–252.

Cole, Peter, and Gabriella Hermon. 1998. The typology of wh-movement. *Syntax* 1:221–258.

Comorovski, Ileana. 1996. *Interrogative phrases and the syntax-semantics interface*. MIT Press.

Cresti, Diana. 1995. Extraction and reconstruction. *Natural Language Semantics* 3:79–122.

Dobrovie-Sorin, Carmen. 1990. Clitic doubling, wh movement, and quantification in Romanian. *Linguistic Inquiry* 21:351–398.

Doetjes, Jenny, and Martin Honcoop. 1997. The semantics of event-related readings: A case for pair-quantification. In *Ways of scope taking*, ed. Anna Szabolcsi, 263–310. Dordrecht: Kluwer.

Durrleman, Stéphanie. 2000. The left periphery in jamaican creole. Ms., University of Geneva.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22:1–25.

Fiengo, Robert, and Robert May. 1994. *Indices and indentity*. Cambridge, Mass.: MIT Press.

Fox, Danny. 1995. Economy and scope. *Natural Language Semantics* 3:283–341.

Frampton, John. 1991. Relativized minimality: A review. *The Linguistic Review* 8:1–46.

Fujita, Tanaka, Ito, and Cheng. 1992. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* 360:343.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, Mass.: MIT Press.

Grosu, A. 1972. The strategic content of island constraints. *Ohio Working Papers in Linguistics* 13.

Haeberli, Eric. 1999. Features, categories and the syntax of A-positions. Doctoral Dissertation, University of Geneva.

Haegeman, Liliane. 1993. Object clitics in West Flemish. *GenGenP* 1:1–30.

Heycock, Caroline. 1995. Asymmetries in reconstruction. *Linguistic Inquiry* 26:547–570.

Hiramatsu, Kazuko. 2000. Accessing linguistic competence: Evidence from children's and adults' acceptability judgments. Doctoral Dissertation, University of Connecticut.

Hirschbühler, Paul, and Marie Labelle. 2000. Verb positions in the history of French negative infinitival clauses. Conference presentation, Lund.

Huang, C.-T. James. 1982. Move WH in a language without WH movement. *The Linguistic Review* 1:369–416.

Kaye, Jonatha, Jean Lowenstamm, and Jean-Roger Vergnaud. 1985. The internal structure of phonological representations: A theory of charm and government. *Phonology Yearbook* 2:305–328.

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

Kayne, Richard. 1989. Facets of romance past participle agreement. In *Dialect variation in the theory of grammar*, ed. Paola Benincà, 85–104. Dordrecht: Foris.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kiss, K. E. 1993. Wh-movement and specificity. *Natural Language and Linguistic Theory* 11:85–120.

Kobatake, and Tanaka. 1994. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology* 71:856.

Koopman, Hilda, and Dominique Sportiche. 1990. Subjects. Ms., University of California, Los Angeles.

Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Foris.

Kroch, Anthony. 1989. Amount quantification, referentiality, and long wh-movement. Ms., University of Pennsylvania.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15.2:235–289.

Mahajan, Anoop. 1990. The a/a-bar distinction and movement theory. Doctoral Dissertation, MIT.

Manzini, Rita. 1992. *Locality*. MIT Press.

Marti, Luisa. 1999. Extraction from subjects in Spanish. Ms., University of Connecticut.

McCawley, James. 1988. *The syntactic phenomena of English*, volume 2. Chicago, IL: University of Chicago Press.

McCloskey, James. 2000. Quantifier float and wh-movement in an Irish English. *Linguistic Inquiry* 31:57–84.

McDaniel, Dana. 1989. Partial and multiple wh movement. *Natural Language and Linguistic Theory* 7:565–604.

Müller, Gereon. 1995. *A-bar syntax*. The Hague: Mouton de Gruyter.

Muyskens, P. 1981. Quechua causatives and logical form: A case study in markedness. In *Theory of markedness in generative grammar*, ed. Adriana Belletti, Luciana Brandi, and Luigi Rizzi. Pisa: Scuola Normale Superiore.

Nicolis, Marco. 1999. L-tous, restructuring and quantifier climbing. Ms., University of Siena.

Obenauer, Hans. 1994. Aspects de la syntaxe A-barre. Doctoral Dissertation, Université de Paris VIII.

Ordonez, Francisco. 1998. Postverbal asymmetries in Spanish. *Natural Language and Linguistic Theory* 16:313–346.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In *The representation of (in)definiteness*, ed. E. Reuland and A. ter Meulen, 98–129. Cambridge, Mass.: MIT Press.

Pesetsky, David. 1995. *Zero syntax*. Cambridge, Mass.: MIT Press.

Platzack, Christer. 1996. The initial hypothesis of syntax. a minimalist perspective on language acquisition and attrition. In *Generative perspectives on language acquisition*, ed. Harald Clahsen. Amsterdam: John Benjamins.

Pollard, Carl, and Ivan Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23.2:261–303.

Pollock, Jean-Yves. 1992. Belletti: Generalised verb movement: Aspects of verb syntax. *Language* 68.

Postal, Paul. 1998. *Three investigations of extraction*. Cambridge, Mass.: MIT Press.

Radford, Andrew. 1990. *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English*. Oxford: Blackwell.

Richards, Norvin. 1997. What moves where when in which language? Doctoral Dissertation, MIT.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Rizzi, Luigi. 1988. On the status of referential indices. Ms., University of Geneva.

Rizzi, Luigi. 1990. *Relativized minimality*. Linguistic Inquiry Monograph 16. Cambridge, Mass.: MIT Press.

Rizzi, Luigi. 1992. Argument/adjunct (a)symmetries. In *Proceedings of NELS 22*, 365–381. Amherst, Mass.: GLSA.

Rizzi, Luigi. 1994. Some notes on linguistic theory and language development: the case of root infinitives. *Language Acquisition* 3:371–393.

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

Rizzi, Luigi. 1999. On the projection of int°. Ms., University of Geneva.

Rizzi, Luigi. 2000. Reconstruction, weak island sensitivity, and agreement. Ms., University of Geneva.

Rizzi, Luigi. 2001. Relativised minimality effects. In *A handbook of syntactic theory*, ed. Mark Baltin and Chris Collins, 89–110. Oxford: Blackwell.

Romero, Maribel. 1998. Problems for a semantic account of scope reconstruction. In *Reconstruction: Proceedings of the 1997 tübingen workshop*, ed. G. Katz, S. Kim, and H. Winhart, 127–153.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Bloomington: Indiana University Linguistics Club.

Ross, John Robert. 1984. Inner islands. In *Proceedings of the Tenth Berkeley Linguistic Society*, ed. Berkeley Linguistic Society. California.

Rullman, Hotze. 1995. Maximality in the semantics of wh-constructions. Doctoral Dissertation, UMass, Amherst.

Sperber, Dan, and Deirdre Wilson. 1986. *Relevance: Communication and cognition*. Cambridge, Mass.: Harvard U. Press.

Starke, Michal. 1993. En deuxième position en Europe Centrale. Mémoire, University of Geneva.

Starke, Michal. 1995. On the format for small clauses. In *Small clauses*, ed. Anna Cardinaletti and Teresa Guasti, volume 28 of *Syntax and Semantics*, 237–269. New York: Academic Press.

Starke, Michal. 1997. Loc: A sketch with special reference to anaphora. Ms., University of Geneva.

Starke, Michal. 1999. Unifying locality. Presented at the Pontignano Workshop.

Szabolcsi, Anna. 1999. Strong and weak islands. Ms. NYU.

Szabolcsi, Anna, and Frans Zwarts. 1997. Weak islands and an algebraic semantics for taking scope. In *Ways of scope taking*, ed. Anna Szabolcsi, 217–262. Kluwer Academic Publishers.

Tanaka. 1996. Inferotemporal cortex and object vision. *Annual Review of Neuroscience* 19:109.

Tanaka, Saito, Fukada, and Moriya. 1991. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology* 66:170.

Thomason, R. H., ed. 1974. *Formal philosophy: Selected papers of Richard Montague*. New York: Yale University Press.

Torrego, Esther. 1986. Empty categories in nominals. Ms. University of Massachusetts, Boston.

Vergnaud, Jean-Roger. 1974. French relative clauses. Doctoral Dissertation, MIT.

Watanabe, Akira. 1992. Subjacency and S-Structure movement of *wh*-in-situ. *Journal of East Asian Linguistics* 1.3:255–291.

Wexler, Ken. 1994. Optional infinitives, head movement, and economy of derivation. In *Verb movement*, ed. Norber Hornstein and David Lightfoot, 305–350. Cambridge: Cambridge University Press.

Wilder, Chris. 1994. Coordination, ATB and ellipsis. *Groninger Arbeiten zur Germanistischen Linguistik* 37:291–331.

Wilder, Chris. 1996. Rightward movement as leftward deletion. Ms., Max Planck, Berlin.

Wilder, Chris. 1997. Some properties of ellipsis in coordination. In *Studies on universal grammar and typological variation*, ed. Artemis Alexiadou and Alan Hall, 59–107. Philadelphia/Amsterdam: John Benjamins.

Zanuttini, Raffaella. 1991. Syntactic properties of sentential negation: A comparative study of Romance languages. Doctoral Dissertation, University of Pennsylvania.

Zwart, Jan-Wouter. 2000. Asymmetric verb movement: The 'phonology' of head movement. Talk at the University of Geneva.