Multiattachment syntax, “Movement” effects, and Spell-Out*

Steven Franks
Indiana University — Bloomington

This paper addresses a set of puzzles associated with Spell-Out. Of primary concern is the pronunciation and interpretation of hypothetical intermediate copies of moved constituents. I show that LF $wh$-movement never exhibits any intermediate effects and argue that intermediate effects are best accommodated by rejecting successive cyclic movement in favor of a “one fell swoop” feature-driven approach. I regard “movement” as feature-driven multiattachment, rather than re-merge of actual copies. This relation is unbounded in the syntax; $wh$-movement intermediate effects arise through the attempt to form a chain for Spell-Out purposes. Since these can only involve C (there being no intermediate SpecCPs), all such effects are head effects.

1. The project

This paper weaves together several independent strands of thinking about syntactic representations and derivations to articulate a new grammatical architecture. The specific leading ideas are minimalist in nature, serve to define a larger research program, and can be described as follows:

1. There are no autonomous copies. “Movement” is instead a metaphor for multiple occurrences, i.e., distinct nodes in the tree are linked to the same item (alternatively, call up/point to the same address). This can be represented in terms of multiattachment,¹ as in Gärtner (2002), Frampton (2004), Johnson (2010), Citko (2011), or De Vries (2012), among others.

2. In the spirit of Takahashi (1994) and Boeckx (2003), there is no EPP, “edge,” successive cyclic, or COMP-to-COMP movement. Instead, multiattachment/movement is invoked only when the triggering feature is introduced, creating long-distance associations/“one fell swoop” movements.

3. Since words and phrases are feature sets, with their particular feature geometries, features can be multiattached (i.e., every piece of structure has an address). One
consequence is that *wh*-movement — established as a *syntactic* relation — involves only *wh* features.2

4. Multiattachment can be resolved in various ways (consistent with feature geometry). For example, links can be deleted, resulting in PF ellipsis, Right Node Raising, or the “delayed” pronunciation of clitics, feature bundles can be split off, resulting in resumption, and phrases forced into head positions can lead to *wh*-copying.3

5. “Movement” effects, including intermediate movement and island phenomena, arise as a consequence of the mapping to PF (contra Takahashi and Boeckx), rather than in the syntax per se, and derive from the nature of Spell-Out.4 Since LF movement requires no access to PF information, it does not show any such effects.

It is argued that these mutually supportive ideas combine to produce a restrictive model of grammar that makes strong predictions about how diverse structures should be treated. The attempt to adapt this model to accommodate different phenomena leads to interesting challenges and particular analyses, some of which are sketched out below. The present project of pushing the empirical coverage of multiattachment feature-geometry syntax is programmatic and should be taken in this spirit.

2. On cyclic domains

This section surveys some of the reasons for positing successive cyclic *wh*-movement and some of the approaches one might entertain for inducing cyclicity. At least since Chomsky (1977), the claim that movement is not unbounded has become widely accepted. Thus in (1), *which book* moves from its external merge position as object of *borrowed*, through the lower SpecCP (indicated in outline font), to its ultimate [+Q] SpecCP landing site:5

(1) I wonder [CP *which book* [+Q]-C0 [TP Elisabeth thought [CP *which book* that [TP they borrowed *which book*]]]]?

*Which book* is pronounced in the highest position and deleted in the lowest. As for the middle position (in outline font), I will argue that it does not exist.

2.1 Motivation for intermediate movement

There are many arguments in the literature for intermediate *wh*-movement; cf. e.g. Boeckx (2003, 2008), Felser (2004), or Franks (2005). Types of evidence include Henry’s (1995) claim that T-to-C movement (subject-auxiliary inversion) in embedded clauses in Belfast English serves as a diagnostic that the fronted *wh*-phrase must have passed through the intermediate SpecCP:
(2) \[\text{CP} \text{What did } [\text{TP} \text{Mary claim } [\text{CP} \text{what } [\text{C} \text{did} ] [\text{TP} \text{they steal what}]]]]\]?

Consider also Torreg o's (1984) demonstration that “V-Preposing”/inversion in Spanish is triggered not only by a wh-phrase ultimately landing in SpecCP, but also by its passing through a SpecCP in the course of the derivation:

(3) \[\text{CP} \text{Qué pensaba Juan } [\text{CP} \text{qué que le había dicho Pedro } [\text{CP} \text{qué que había publicado qué la revista}]]]?

‘What did Juan think that Pedro had told him that the journal had published?’

Another putative intermediate effect involves agreement, either on C, as in (4) from KiLega, cited by Carstens (2010: 52), or on T, as in (5) from Ojibwe, due to Lochbiler and Mathieu (2011):6

(4) \[\text{CP} \text{Biki bi- } [\text{TP pro b- á- tend- ilé } [\text{CP} \text{bi- } \text{what 8CA- } \text{2SA- ASP- say- perf } \text{8CA- } \text{TP pro b- á- gül- ilé biki}]]]

‘What did they say they had bought?’

(5) \[\text{CP} \text{Aniish1 Bill gaa- eneendang } [\text{t}_1 \text{John gaa- keda- t what Bill wh.pst- think John wh.pst- say- 3 conj } \text{[t}_1 \text{Mary gaa- giishnedot } \text{t}_1 ]]?

‘What does Bill think John said Mary bought?’

Colloquial German wh-copying, as in (6), also involves C; cf. e.g. Felser (2004) or Rett (2006):

(6) \[\text{CP} \text{Wen glaubt Hans } [\text{CP} \text{wen } [\text{TP Jakob wen gesehen hat}]]]?

‘Who does Hans believe that Jakob saw?’

A more subtle argument can be constructed on the basis of parasitic gaps; see Nissenbaum (2000) for general discussion. In (7), we find credible if subtle cases of intermediate licensing of parasitic gaps, where the clauses containing the parasitic gap are adjuncts of the intermediate VPs, headed by agreed and claimed, respectively:

(7) a. \[\text{Which proposal did the NSF say } [\text{CP which proposal that you agreed [CP that Julia should reject which proposal] [without PRO (=you) considering pg]]]?\]

b. \[\text{Which painting did Karen say } [\text{CP which painting that you claimed [CP that people love which painting] [in order PRO (=you) to get me to look at pg]]]?\]

Perhaps the most famous demonstration of a need for intermediate wh-movement is the phenomenon of intermediate binding:8
(8) \[ \text{CP Which picture of herself did [TP Mary say [CP which picture of herself [TP Bill bought which picture of herself]]]} \]

The connectivity effect in (8), in which the reflexive is bound by Mary, is standardly taken as evidence that the wh-phrase containing herself must pass through the intermediate SpecCP, a position in which it is locally c-commanded by Mary.

2.2 Reasons to constrain movement: wh-island and (other) relativized minimality effects

Island phenomena have long provided a textbook argument for successive cyclic wh-movement. If movement were unbounded, the reasoning goes, (9b) should be possible alongside (9a):

(9) a. What does David believe [CP (that) we bought what]?
   b. *What does David wonder [CP where we bought what]

Under GB, the combination of Subjacency and the Strict Cycle Condition forced what to move through the lower SpecCP, but in (9b) that position is occupied by where, hence is not an available landing site. With no “escape hatch,” the embedded clause becomes an “island,” in the classic terminology of Ross (1967).

The problem presented by (9b) can also be seen as one of skipping over an intervening wh-phrase in the attempt to relate two wh-phrase sites. Rizzi (1990) later relativized this intervention effect to accommodate illicit Long Head Movement and Superraising, as in (10):

(10) a. *Have John could have eaten the cake?
   b. *John seems that it is likely John to win.

While the focus in this paper is on A′-movement, the same issues arise here as well, and a solution must ultimately be sought in terms of leading ideas 2 and 5 — movement is unbounded and its effects derive from the exigencies of pronunciation.

2.3. Ways of forcing recourse to an escape hatch

The question of what drives movement has been a recurring theme of generative grammar. The GB tradition, stemming from Ross (1967), characterized movement as free but subject to constraints. Under minimalism, however, all operations require motivation, and the standard answer has been to posit some feature to induce the movement. It is nonetheless unclear, under the standard minimalist assumptions in (11), why wh-movement should ever target intermediate landing sites:

(11) a. Movement is driven by a feature of the target. [Attract]
b. Syntax is built from the bottom up:
i. Merge can introduce items only at the root of the tree. [The Extension Condition]

ii. Move also respects the Extension Condition.

   c. The feature driving wh-movement is on the interrogative C0. [The Wh Criterion]

Since intermediate landing sites are not [+Q], this particular combination of assumptions gives rise to the “look-ahead” puzzle: at the point in the derivation where intermediate movement applies, the motivating information is not yet available (i.e., the attracting feature only pertains to the ultimate landing site).

One might envision the following ways of tinkering with the system in (11) to obtain intermediate movement:

1. With respect to (11a), one could argue that moving items themselves have some feature driving the movement, i.e., one could reinstate Move.
2. With respect to (11bii), one could allow movement without extending the tree.
3. With respect to (11c), one could endow intermediate sites with an attracting feature.

These have all been proposed, in one form or another. Bošković (2007) optionally endows wh-phrases with a feature that is involved in wh-movement, but argues that there is no feature checking at intermediate movement sites. Boeckx (2003) makes use of chain formation at the point when the attracting head is merged; intermediate sites are simultaneously created in order to minimize chain links. Chomsky (2000) optionally endows intermediate C0 with the EPP property, to attract wh-phrases that need to move through the relevant landing site; in Chomsky (2008) this is recast as the “edge” feature of a phase head.

Boškovič’s insight that the moving element has a formal inadequacy (namely, an uninterpretable feature) which drives the movement is conceptually akin to the proposal of “agnostic” movement in Franks (2006) and Franks and Lavine (2006). The basic idea of agnostic movement was that whenever a phase is completed and it contains an element with an unchecked feature, that element moves to the phase edge in order to remain visible to a potential higher checker. Failure to move would ensure inevitable crash. Thus, movement takes place in the absence of knowledge of an eventual checker, i.e.,agnostically, just in case an appropriate element may be merged and save the derivation.

3. When intermediate effects might be expected but are missing

This section examines several problems for the successive cyclic movement hypothesis. First, there is the “look-ahead” problem noted above. Second, actual intermediate effects turn out to be limited to heads, suggesting that SpecCP is never implicated. Finally, it is shown that LF/covert movement invariably behaves like no movement.
3.1 Some things that do not happen with overt wh-movement

There are good reasons to suspect that a long-distance fronted wh-phrase never actually occupies any intermediate SpecCP position along the way.

3.1.1 The “look-ahead” problem

From the perspectives of bottom-up syntax and cyclicity, it is striking that when wh-movement fails it is the lowest copy (in an A-position) rather than an intermediate one (in an A’-position) which ends up being pronounced:

(12) a. Who thinks that John bought what?
    b. *Who thinks what (that) John bought?

That is, if in the course of the derivation what moves to the embedded SpecCP in (13), then we would expect it to remain and be pronounced there once further wh-movement into the matrix clause is blocked by who:

(13) *Who [TP who thinks [CP what (that) [John bought what]]]?

Neither Chomsky (2000) nor Bošković (2007) have anything insightful to say about the ungrammaticality of (12b); what cannot “know” that it should not move to the local SpecCP without looking outside its own clause. Their only solution is mechanical, with what having the option of moving or not.

In fact, “failed” wh-movement — movement which hypothetically starts but cannot be completed — should not be allowed to start in the first place. Otherwise, we might expect it to license parasitic gaps, but this never happens. (14) patterns like (15), and not like the examples in (7):

(14) Who thinks [CP that John bought what (*after trying on pg)]?
(15) Which agency rejected which proposal (*after considering pg)?

(cf. Which proposal did you reject e after considering pg?)

Intermediate movement of what in (14) can no more license a parasitic gap than can LF-movement of which proposal in (15).

The obvious minimalist solution to the look-ahead problem raised by failed wh-movement is thus to take seriously the idea that it never actually starts, i.e., movement only takes place when motivated. Since in (12) [+Q] what needs a [+Q]-C0, it should in fact wait to move until the matrix [+Q]-C0 is merged, in which case it (re)merges at the top of the tree. Hence there should be no intermediate what in either (12a) or (12b). Under the multiattachment approach presented later, what will be attached as the object of bought and who as the subject of thinks. Its [+Q] feature will be linked to the matrix C0, as will that of who. Spell-Out will then attempt to access the relevant phonological features, but reaches who first, realizing it as SpecCP. This leaves what associated in LF with the matrix C, allowing it to scope (in the pair-list question) but
only permitting pronunciation in its A-position. Crucially, no attachments to any intermediate sites are created.

3.1.2 Intermediate copies are not pronounced

In general, movement that is not motivated never surfaces. This was just demonstrated for syntactically failed wh-movement, but it also pertains to wh-movement which fails for PF reasons. A particularly telling illustration of this fact concerns the status of intermediate copies of wh-phrases which encounter clearly PF problems. This can be seen in the well-known case of Billings and Rudin’s (1996) constraint against sequences of homophonous wh-phrases.\footnote{As discussed in Bošković (2002: 364–376) and Franks (2009: 114–118), inter alia, in multiple wh-fronting languages such as Bosnian/Croatian/Serbian (BCS), Bulgarian, or Romanian the second of two adjacent identical wh-phrases cannot be pronounced, as in BCS (16):

\[(16) \ Šta_1 šta_2 [šta_1 uslovljava šta_2]?\]

‘What conditions what?’

The PF nature of this effect is confirmed in several ways. First, as Bošković (2002) observes, it disappears if the wh-words are not adjacent:

\[(17) \ Šta_1 neprestano šta_2 [šta_1 uslovljava šta_2]?\]

‘What constantly conditions what?’

(cf. \(\tilde{\text{š}}\)ta_1 neprestano šta_2 [šta_1 uslovljava šta_2]?)

Second, consider the pattern in Bulgarian (18):

\[(18) a. \ Koj na kogo kogo e pokazal?\]

‘Who pointed out who to whom?’

b. Koj na kogo e pokazal kogo?

c. Koj na koj kogo e pokazal?

The order *na kogo kogo ‘to whom whom’ in (18a) is not felicitous, so the second kogo is pronounced in its lower occurrence. Interestingly, as noted in Billings and Rudin (1996), replacing dative na kogo ‘to whom’ by its colloquial variant na koj ‘to who’, as in (18c), removes the homophony and obviates the need for lower pronunciation. A third piece of evidence confirming the PF nature of the constraint, discussed by Bošković (2002) but now such a compelling discovery as to have entered the canon of textbooks such as Hornstein, Nunes, and Grohmann (2005), is that parasitic gaps are licensed as if the homophonous phrase had indeed moved in the syntax. Romanian (19a) is just like BCS (16); (19b, c) show that (unlike its English translation) the unpronounced direct object is syntactically active:

\[(19) a. \ Ce_1 ce_2 [ce_1 precede ce_2]?\]

‘What precedes what?’
b. $\text{Ce}_1 \text{ ee}_2 [\text{ee}_1 \text{ precede } \text{ce}_2]$, [fără să influenceze pg]?
   ‘What precedes what without influencing it/*pg?’

c. $\text{Ce}_1 \text{ ee}_2 [\text{ee}_4 \text{ demonstratează } \text{ce}_2 \text{ că Maria a cumpărat } \text{ce}_2]$
   [fără să citească pg]]?
   ‘What demonstrates that Maria bought what without reading it/*pg?’

With this background, consider what happens in long-distance multiple $wh$-questions. 
Bošković (2002: 372–374) examines the issue and concludes that, although the data are somewhat murky and there are possible interfering factors, the traditional successive-cyclic analysis is not supported.13 Strikingly, it turns out that whenever a $wh$-phrase cannot be pronounced in its head position — even for superficial, PF reasons such as the constraint against homophonous sequences — it is the tail (or a scrambled/focused copy in the lowest clause) rather than any copy in any putative intermediate SpecCP position that gets pronounced:

(20) $\text{Ce} (*\text{ce}) \text{ crede Ion [CP (*\text{ce}) că [(%\text{ce}) a determinat } \text{ce}]])$\textsuperscript{14}
   ‘What does Ion think that determined what?’

(21) Ivan $\text{kakvo} (*\text{kakvo}) \text{ misli [CP (*\text{kakvo}) če [(%\text{kakvo}) obuslavja } \text{kakvo]}]]$\textsuperscript{15}
   ‘What does Ivan think conditions what?’

The correct observation about Romanian (20) and Bulgarian (21) is that $wh$-phrases are pronounced in the highest position to which they have any motivation to move. Typically this is the matrix scope position (i.e., where the [+Q] feature is checked), except when this comes up against the homophony constraint. In sum, intermediate copies are not available to be pronounced.

3.1.3 An exception and a restriction

There is one exception to this generalization. This is the $wh$-copying construction introduced in (5) and repeated in (22a). A salient and much discussed aspect of this phenomenon is that unequivocal phrases do not exhibit it, as in (22b), from Fanselow and Mahajan (2000).

(22) a. $[\text{CP Wen} \text{ glaubt Hans } [\text{CP wen } \text{TP Jakob wen gesehen hat}]]$?
   ‘Who does Hans believe (who) Jakob saw?’

b. $^* [\text{CP Welchen Mann} \text{ glaubst du } [\text{CP welchen Mann } \text{TP sie welchen Mann liebt}]]$?
   ‘Which man do you believe (which man) she loves?’
   (cf. Welchen Mann glaubst du daß sie liebt?)

The same pattern arises in other languages, such as the following pair from Yugoslav Romani, adapted from Nunes (2004: 18–19) citing McDaniel (1986):

(23) a. $[\text{CP Kas} \text{ misline } [\text{CP kas o Demir dikhlâ kas}]]$
   ‘Who do you think (who) Demir saw?’
b. *[{CP Save chave} mislinea [{CP save chave o Demiri dikhà save chave}]]

‘Which boy do you think (which boy) Demir saw?’

Although there is some variation with regard to the availability of wh-copying for PPs, I will concur with Nunes that this is a head restriction and apparent PP wh-copying invariably involves a complex (or morphologically reanalyzed) head. For example, although Fanselow and Mahajan (2000) report instability for examples like (24a), (24b) with a fused head is perfect:

(24) a. (*) An wen glaubst du, an wen sie denkt?
   ‘About who do you believe she is thinking?’

b. Wovon glaubst du, wovon sie träumt?
   ‘What about do you believe she is dreaming?’

Felser (2004: 550) marks (24a) with ‘%’, noting the following Afrikaans example from Du Plessis (1977: 725):16

(25) Met wie het jy nou weer gesê [{CP met wie} met wie gaan Jan trou]?
   ‘Who did you say again Sarie thought Jan is going to marry?’

Afrikaans (25), as well as McDaniels’s (1986) Mit wem glaubst du, mit wem Hans spricht? ‘With whom do you believe (with whom) Hans is speaking?’, are extremely unusual. Nunes (2004: 169, fn. 35) suggests that these ‘involve fusion between the preposition and the wh-word, followed by fusion with the intermediate C0.’ However derived, it thus seems that wh-copying is only available for heads (be they complex or simple). In Section 4.2, this restriction will be shown to fall out naturally from the unbounded multiattachment system advocated below.

A second important observation about wh-copying is that it is only putative copies in C0 (or SpecCP) that can be pronounced. In a sense, this is the flip side of what we saw in the previous subsection: the copy in the wh-phrase’s original tail position is not accessible for PF purposes; (26) is not a possible variant of (22):

(26) *[{CP Wen glaubt Hans [{CP wen} [TP Jakob wen gesehen hat]]}]

I will argue that this restriction is once again a matter of the wh-copying effect being limited to heads, in that it can only arise when syntactic phrases are forced into head positions.

3.2 The non-existence of intermediate LF movement

LF wh-movement is yet another type of failed movement. Here, overt movement does not take place, either because the SpecCP landing site is otherwise filled (as in the
English look-ahead examples of Section 3.1.1) or because the language lacks overt wh-movement in general (as in Chinese). Interestingly, to my knowledge none of the diagnostics for intermediate wh-movement obtain for LF movement. This fact requires explanation and, as proposed in Section 4, the obvious way to tackle it is to create movement/multiattachment chains only on the PF side, when lexical material needs to be accessed for pronunciation.

Section 2.1 described some phenomena associated with the assumption of hypothetical intermediate copies for overtly moved constituents. These facts are complemented by their curious absence under LF wh-movement. The pattern that only overtly moved wh- phrases ever show LF reconstruction effects is underappreciated but robust. Thus, when the wh-phrase cannot occupy its matrix SpecCP scope position (because it is otherwise occupied), not only is the tail pronounced, as in the Bulgarian and Romanian examples which encountered the constraint against adjacent homophonous items, but all intermediate effects also disappear. We can see this in the impossibility of intermediate binding of herself by Mary in (27) and in the absence of intermediate V-Preposing of dijo and había publicado in Spanish (28):

(27) *When did Mary say [CP which picture of herself(that) [Bill bought which picture of herself]]?

(28) [CP Quién pensaba [CP qué que Pedro dijo [CP qué que Juan había publicado qué]]? ‘Who thinks that Pedro said that Juan published what?’

Evidence for putative copies (in outline font) is clearly absent for failed (covert) wh-movement. These judgments contrast markedly with those for successful (overt) movement in (7) and (3), respectively.

Turning to parasitic gaps, compare the impossible sentences in (30) with those in (6), repeated as (29):

(29) a. *Which proposal did the NSF say [CP which proposal that you agreed [CP that Julia should reject which proposal] [without PRO (=you) considering pg]]?

b. Which painting did Karen say [CP which painting that you claimed [CP that people love which painting] [in order PRO (=you) to get me to look at pg]]?

(30) a. *Which agency said [CP which proposal that you agreed [CP which proposal that Julia should reject which proposal] [without PRO considering pg]]?

b. *Which woman said [CP which painting that you claimed [CP which painting that people love which painting] [in order PRO to get me to look at pg]]?
Once again, LF movement has no intermediate effects, supporting my contention that there is no intermediate movement in such failed cases.

Consider, finally, LF \textit{wh}-movement in languages such as Chinese that lack overt \textit{wh}-movement. In (31a) the reflexive \textit{ziji} can be bound by either a local or more distant 3rd person antecedent. In (31b), on the other hand, since the more distant potential antecedent differs from the intervening \textit{wo/ni ‘I/you’} in person features, \textit{ziji} can only be locally bound and \textit{Zhangsan} is not accessible:

\begin{enumerate}
\item a. \textit{Zhangsan$_1$ xiangxin Li-Laoshi$_2$ renwei ziji$_{1/2}$-de erzi zui congming.}
\textit{Zhangsan believe Li-teacher think self-de son most smart}
‘Zhangsan believes that Teacher Li thinks that his son is the smartest.’
\item b. \textit{Zhangsan$_1$ xiangxin \textit{wo/ni$_2$ renwei ziji$_{1/2}$-de erzi zui congming.}}
\textit{Zhangsan believe I/you think self-de son most smart}
‘Zhangsan believes that I/you think that my/your/*his son is the smartest.’
\end{enumerate}

Interestingly, such person blocking effects in Chinese are not obviated by LF \textit{wh}-movement. This is shown by (32b), hence the contrast with its English counterpart:

\begin{enumerate}
\item a. \textit{Zhangsan$_1$ xiangxin Lisi$_2$ xihuan ziji$_{1/2}$-de na yi zhang zhaopian ne?}
\textit{Zhangsan believe Lisi like self-de that one CL photo q}
‘Which picture of himself does Zhangsan believe that Lisi likes?’
\item b. \textit{Zhangsan$_1$ xiangxin \textit{ni$_2$ xihuan ziji$_{1/2}$-de na yi zhang zhaopian ne?}}
\textit{Zhangsan believe you like self-de that one CL photo q}
‘Which picture of yourself does Zhangsan believe that you like?’
\end{enumerate}

\textit{(not: ‘Which picture of himself does Zhangsan believe that you like?’)}

The fact that LF \textit{wh}-movement in Chinese does not escape person blocking effects shows that it, too, does not proceed in a stepwise fashion.

In sum, covert movement neither feeds binding nor triggers V-Preposing nor licenses parasitic gaps. This is puzzling if one assumes a bottom up syntax with successive cyclic movement steps, since the fact that the movement is ultimately not viable can only be determined at the end of the derivation. There is no local reason not to take the intermediate steps, steps which would under standard models (i.e., with successive cyclic, barrier, or phase-based opacity) be necessitated in order to reach the \textit{wh}-phrase’s ultimate landing site.

4. A multiattachment solution and its consequences

In this section the general form of a solution to these puzzles is laid out. The basic ideas are that movement only occurs when truly motivated (rather than stipulated or freely), that it should be construed in multiattachment/occurrence (rather than copy) terms,
that the job of the syntax is to establish a multiattachment relation between relevant sets of features (rather than between words or phrases per se), and that this relation is unbounded. All intermediate movement effects are the consequence of Spell-Out — a term I reserve explicitly for the mapping to PF — searching for the material associated with the featural link established in the syntax. The observed effects are therefore morphological or phonological in nature.

My reasoning, assumptions, and expectations are as follows:

1. Lexical items, including grammatical formatives, are structured bundles of features.
2. Syntactic “movement” is the pairing of features. Thus, a [+Q]-C looks for an unvalued [uQ] feature (finding it on a wh word). I represent this as linking of features (perhaps a notational variant of Chomsky’s Move F or Agree). This happens as soon as the [+Q]-C enters the derivation.
3. Structures are built, from the bottom up, in separate work-spaces and attached to each other as needed (or required for linearization; cf. Nunes 2004).
4. LF requires access to the Lexico-Conceptual Features of the item whose [Q] feature is multiattached, but this is unrestricted (and is conceivably established by syntactic multiattachment).
5. It is only Spell-Out (to PF) which requires access to Morpho-Phonological information, in the effort to pronounce constituents containing multiattached features. This process attaches such constituents as close as possible to the triggering feature (perhaps a notational variant of “pied-piping” in the Move F system). This essentially recreates EPP-type movement, but only as a consequence of Spell-Out.
6. All intermediate effects derive from the exigencies of the mapping to PF. Since LF does not involve pronunciation, there can be no LF movement intermediate effects. Moreover, since there are only intermediate head (rather than Spec) sites to be attached to and the Extension Condition prohibits new internal structure from being created, all intermediate effects have to be morphologically compatible with heads.

In this light, we now return to the phenomena described in Sections 2 and 3 above.

4.1 Overt wh-movement

Consider simple local wh-movement, in which a [+Q]-C points to a [Q] feature on some category lower in the tree.
The diagram in (33) is meant to convey a number of ideas: (i) feature sets point to other feature sets (the solid line), which are technically addresses; (ii) the [+Q] in C actually only points to the [uQ] feature of which (which consequently becomes valued as [+Q], here shown as “u>+Q”); (iii) multiple nodes can point to the same address (i.e. multiattachment); (iv) to resolve multiattachments and thereby initiate Spell-Out, the [+Q] feature must probe down the tree (the dotted line), attempting to access Morpho-Phonological information; (v) if accessed, that PF information is then attached, resulting in effect in wh-movement (the dashed line). As presented, movement thus reduces to three distinct steps; I return to the question of their independence in Section 5.2 below.

The proposal in (i) is taken from Franks and Herring (2011), in which we restyle lexical items and phrases as “pointers.” This term is borrowed from computer science and refers to the memory location of a data object (as opposed to the information it contains). Trees are thus recursive data structures wherein nodes are pointers either to other nodes (the recursive case) or lexical items (the base case). Herring’s idea is that rethinking syntactic structures as assemblies of pointers allows operations at later stages to affect — through use of shared memory addresses — syntactic objects formed in earlier stages without compromising any commitment to bottom-up assembly of the tree.

There are two essential reasons why the phrase which book in (33) ends up pronounced in SpecCP. Firstly, and most importantly, I take this to be a morphological
fact: as a phrase, *which book* cannot be pronounced in $C^0$, even though the [+Q] in $C$ targets it.\(^{25}\) I will argue throughout this section that this PF fact — the morphological impossibility of realizing a phrase in a head position (without special manipulation) — is what explains many ostensibly syntactic phenomena. Second, *which book* attaches as close as possible to the probe, extended the tree by creating a new SpecCP. I will argue below that the impossibility of creating new, clause internal phrasal structure (i.e. the Extension Condition) is what explains why ostensibly intermediate movement behaves as it does: it is not addition of structure, but rather new association of features with existing structure.

Consider next a straightforward long-distance *wh*-movement case, such as (8a) *What does David believe that we bought?*, diagrammed in (34). Crucially, the probing [+Q], in looking for Morpho-Phonological information associated with the [Q] feature of *what*, forms what amounts to a chain (the dotted line). In [+Q]’s search for the initial merge site, every node/data structure all the way down is accessed. Because intermediate $C$ heads are featurally-compatible, they become “entangled” with the probe, represented by the large dot. All this means is that these intermediate nodes must enter into the multiattachment network. Entanglement is in essence a top-down implementation of island effects without successive cyclic movement. One can think of this metaphorically: if a probe encounters a categorially similar higher node, in order to keep the search “alive” it must stop to examine that node’s lower/dependent features and make sure they are not already valued; cf. (36) below for a search which dies.\(^{26}\)
Ultimately, then, island effects, relativized minimality, and reconstruction are going to be side-effects of the need to adapt syntactic representations for the purpose of initiating Spell-Out. I briefly describe some of these below.

4.2 A general approach to islands and reconstruction

Consider, first, a simple multiple *wh*-question, such as *Who bought what?*

In step I, the [+Q]-C is linked to the [uQ] features of both *who* and *what*. This provides the required LF interpretation. However only the higher *wh*-phrase — which is accessed first — can be pronounced in SpecCP because English does not tolerate multiply filled SpecCP. In terms of the mechanics of the model, this means once a matching phrase is accessed under step II, the search for additional phrases ceases. In the third Spell-Out step, that phrase then attaches — in effect, “moves” — as close as possible to the [+Q]-C that probed for it, i.e., extending the tree to create SpecCP. The lower *wh*-phrase remains *in situ*, but the sentence is perfectly fine. Note that this is different in multiple *wh*-fronting languages, in which the search for additional *wh*-phrases continues.

A typical island, such as (8b) *What does David wonder where we bought?,* cannot however be derived:
The intermediate [+Q]-C, which should cause *where* to move to its SpecCP, is encountered by the matrix [+Q] probing down the tree looking for *what*. Failure to access *what* prevents access to the Morpho-Phonological information needed for Spell-Out, hence *what* cannot “move” out of the island.

Islands, in this model, are interpreted as the blocking effect of an intervening feature. One important challenge which remains is to understand why certain features matter more, such as the [+Q] as in (8b)/(36), and others less, such as the Yes/No and definiteness features of (37):

(37) a. *What does David wonder whether we bought?*
    b. *What does David believe the/Julia’s claim that we bought?*

Ultimately we will need an account for why (8b) *What does David wonder where we bought?* is worse than these, but one intuitively appealing idea is that a [−wh, +Q] C or a [+def] D are similar but not identical to the probing head, hence somewhat less prone to entanglement.

Be that as it may, since islands arise through the mapping to PF, such opacity effects are absent in LF. For example, unlike (38b), the pair-list reading is available in (38a), even though *what* is deeply protected from moving overtly:

(38) a. *What does David believe the/Julia’s claim that we bought?*
    b. *What does David believe the/Julia’s claim that we bought?*
(38)  a. Who left the party because of Karen’s claim that she felt sick after having eaten what?
    b. ** What did you leave the party because of Karen’s claim that she felt sick after having eaten?

In (38a), the [+Q] in matrix C⁰ is linked to both who and what, but only the first is accessed and pronounced. No intermediate links are established.

No intermediate link is ever created under LF movement, hence binding of herself by Mary is impossible in (27) *When did Mary say (that) Bill bought which picture of herself? The failure of LF movement to void person blocking effects in Chinese (32) can be understood similarly, except that in English the search fails when a [+Q] link is encountered, whereas in Chinese, which lacks wh-movement, it is never launched in the first place.²⁸

On the other hand, in (7) Which picture of herself did Mary say (that) Bill bought?, the wh-phrase is successfully accessed (as evidenced by the fact that it is pronounced in the matrix SpecCP), so Mary can bind herself in its intermediate position. Linkages here are exactly as in (34). Since the intermediate C⁰ is implicated, it is attached to the entire phrase which picture of herself. The phrase cannot be pronounced in this position, since it involves a head dominating a phrase, but that is a purely PF matter. For LF purposes, on the other hand, binding is perfectly available. This reveals an important conclusion: the probe down the tree in search of the phrase containing the multiattached [Q] feature, i.e., the dotted line in my diagrams, must take place prior to the mapping to LF.

This conclusion is corroborated by the parasitic gap facts, in which intermediate parasitic gaps can be licensed by overt movement, but not by LF movement. A relevant minimal pair — (29b) versus (30b) — is repeated below:

(39)  a. Which painting did Karen say that you claimed [that people love e] [in order PRO to get me to look at pg]?
    b. *Which woman e said that you claimed [that people love which painting] [in order PRO to get me to look at pg]?

The intermediate C, after claimed, is accessed by the matrix [+Q] in searching for which painting in (39a), but not in (39b).²⁹ Hence the parasitic gap is only licensed in the former, the matrix clause of which has essentially the structure in (34).

4.3 More on multiple wh-questions and [wh] linking

It is important to note that, in the system envisioned, there are in principle no locality constraints on the first step, that of linking [+Q] with [uQ]. That is why this linking is tantamount to LF movement. There are however several things that might go wrong with particular attempted linkings. Consider again the ungrammatical (36) *What does David wonder where we bought? As presented, the crash is a matter of matrix [Q]-C
becoming entangled and conflicting with embedded [Q]-C. Presumably, the problem is one of scope and independent of whether we assume Multiple Spell-Out (so that where moves) or not. But I also must assume the Wh Criterion, which in my system means every [+Q, +wh] must be linked to (at least, and in languages like Italian, exactly) one [uQ] (no vacuous quantification) and that every [uQ, +wh] must be linked to (exactly) one [+Q] (no free variables). For example, if the embedded [+Q] in (36) were attached to the [uQ] of both where and what, then the matrix [+Q]-C would end up vacuously quantifying. Other than that, however, the first step of feature linking is free.

The envisioned system thus consists of the following three steps:

(40) Step I: [+Q] is freely linked with any [uQ]
   a. this happens as soon as [+Q] enters the structure
   b. [+Q] is consequently at the top, so linking can only be down (i.e, not with some [wh] merged later/higher)
   c. [+Q] can link with as many [uQ] as it wants, except:
      i. if it links with none, then derivation crashes in LF
      ii. for non multiple question languages (e.g., Italian), it links with just one

(41) Step II: [+Q]-C traverses the tree in search of the phrasal constituent(s) containing [Q] linked in the first step.
   a. C becomes entangled with similar heads, especially other C
   b. if the intervening head has values (for similar features), then the search cancels

(42) Step III: Wh-movement of accessed lexical material to [+Q]

Technically, all these steps are the same kind of operation — pointing to/calling up an address — which I have referred to as “linking” or simply “attachment.” Step I is the core of syntax, since its linking results in feature valuation. That is, it corresponds to standard minimalism’s probing and valuing a feature (except that in my system it is the functional category probe that comes with a fixed feature which sets the value of the matching feature on the goal). Steps II and III correspond to Spell-Out, since the linking under the former introduces cyclic effects and the latter results in category movement/pied-piping.

Consider in this light some possible multiple wh scenarios and their interpretations (reflected in the answers):

(43) Q: Who wonders [who bought what]?
   A1: John wonders who bought what, Mary wonders who bought what, …
   A2: John wonders who bought an apple,
       Mary wonders who bought a pear, …

The question in (43) has a (simplified) feature structure as in (44):
Let us investigate the linking possibilities when step I applies to (44).

In the embedded clause, [+Q]-C can link with the [uQ] features of both of who and what, as in (45a), or with those of who alone, as in (45b):

(45) a. … [CP [+Q]-C [TP [uQ, wh]→ o bought [uQ, wh]-at ]]]

b. … [CP [+Q]-C [TP [uQ, wh]→ o bought [uQ, wh]-at ]]]

When the matrix [+Q]-C is merged and the embedded structure is as in (45a), it links with the [uQ] feature of main clause who, as show in (45a). On the other hand, when the embedded [+Q]-C is as in (45b), then the main clause [+Q]-C links with the [uQ] features of both main clause who and embedded what. This is shown in (46b):

(46) a. [CP [+Q]-C [TP [uQ, wh]→ o wonders]

[CP [+Q]-C [TP [uQ, wh]→ o bought [uQ, wh]-at ]]]

b. [CP [+Q]-C [TP [uQ, wh]→ o wonders]

[CP [+Q]-C [TP [uQ, wh]→ o bought [uQ, wh]-at ]]]

Notice that the option in which embedded [+Q] skips the [uQ] of who and just links with that of what does not give rise to the sentence in (43). The problem is that Spell-Out would cause what rather than who to front in the embedded clause. And indeed, this is what happens in (47), which under my linking system is able to escape Superiority:

(47) Q: Who wonders [what who bought]?

A: John wonders what Mary bought, Sally wonders what Bob bought, …

This surprising result is not expected under traditional approaches but follows straightforwardly from the way step I works. Consider the derivation. In the embedded clause, [+Q] is able to skip the first [uQ] and link only with the second one:
Then, when the matrix [+Q]-C merges, it links with the [uQ] features of both who subjects, as in (49):

(49) \[CP [+Q]-C [TP [uQ, wh]-o wonders [CP [+Q]-C [TP [uQ, wh]-o bought [uQ, wh]-at ]]]

Keep in mind that all this just involves step I above, in which [+Q] is freely linked with any [uQ], so there can be no entanglement or island effects. These arise only by virtue of step II, which initiates Spell-Out. Spell-Out requires first to find the material dominating the linked [Q] (step II) and then, to front it (step III). This means that, in the embedded clause in (49), what is going to move rather than who, producing (47).

The fact that this sentence is unambiguous — the embedded who must have matrix scope — is crucial. It is the higher [+Q] which saves the derivation. This is an important consequence of my model, since it provides a reason why a simple violation of Superiority, such as (50), is much worse than (47).

(50) * John wonders [what who bought].

4.4 Lower wh occurrences are never pronounced in intermediate SpecCPs

We saw above that failed wh-movement, as in (11) and repeated below, never gives rise to intermediate pronunciation of the second wh-phrase:

(51) a. Who thinks that John bought what?
   b. *Who thinks what (that) John bought?

This makes sense, because once the matrix [+Q] encounters who, the search ceases, so what is never accessed and no intermediate links are ever created, either for PF or LF purposes.

The multiple wh-fronting language facts described in Section 3.1.2 present a different kind of problem. Recall what the issue was: when a wh-word immediately follows an identical item it comes up against a surface constraint prohibiting sequences of adjacent homophonous elements; this leads not to pronunciation of the next “copy” down, but rather to pronunciation in its original clause. Here are two additional Bulgarian paradigms:
The wh-phrase kakvo is best in its initial merge position, after the embedded verb. Variants with kakvo in other positions are acceptable to the extent that reordering/scrambling of the direct object is pragmatically viable in general (for functional sentence perspective reasons, such as focus).

Here, however, and unlike English (51), we had independent evidence that a chain is formed linking intermediate C\(^0\)s ultimately to the merge position of kakvo in the lowest clause: parasitic gaps are licensed just as if the movement were overt. More patiently, the actual surface form of the second wh-word must be available to determine whether the constraint is invoked; cf. example (18). 34 So why isn’t the offending wh-phrase pronounced in the next available position down?

The answer, I contend, hinges on which positions are “available.” Intermediate SpecCPs are not, because, with the rejection of successive cyclic movement, there are none. Intermediate C\(^0\)s are not, because only head material can be realized in them, not phrases. 35 Thus, as indicated by the structure in (54), PF is forced to resort to pronunciation in a phrasal position, and this is the tail of the chain (modulo A-scrambling, which introduces a higher phrasal position in the lowest clause, as in (53e) and (53f)).
Bulgarian ordinarily fronts all wh-phrases; this is indicated by the two dashed lines in (54); the structure is simplified to represent multiple SpecCPs and tucking-in (Richards 2001), but the same derivation would obtain if a system such as in Bošković 2002 were adopted instead. Only in PF is it determined that the second instance of kakvo cannot be pronounced in the matrix SpecCP. The phrase is attached to the embedded C0 head (če), but cannot be pronounced here because it runs afoul of morphological compatibility. Hence it surfaces in its original clause, even though, for all intents and purposes, it has moved in the syntax. The multiple wh-fronting data thus fall out from my system, in which there are no copies, just occurrences, and these are created in the mapping to Spell-Out. If there were literal copies, with chains formed in the course of the syntactic derivation, one would expect to see pronunciation in the next highest position when the highest encounters PF difficulties.

Interestingly, this is what happens when enclitics find themselves in a position without prosodic support in PF. Since Franks (1998/2010) and Bošković (2001), a standard account of “clitic third” in Wackernagel type clitic languages such as BCS is that, when the head of a clitic chain fails to be prosodically supported, the next lower viable copy is pronounced. Thus (55a), simplified from Franks (1998/2010), has a structure as in (55b):

(55) a. pro Stalno mi je kupovao knjige.
    constantly me.dat aux.IIIsg bought books
    ‘(He) was constantly buying me books.’
b. \[pro \text{mi}je [[stalno]_ω \text{mi} \text{je} [\text{kupovao} [\text{mi}je [\text{knjige}]]]]\]

Since it does not project a prosodic word ("ω"), \text{pro} cannot host the clitic cluster \text{mi je}, and the next highest occurrence is attempted. The same thing happens in (56), where the parenthetical \text{tvoja mama} 'your mother' is flanked by intonational phrase boundaries ("#"):

\[(56) \#Ja#, \#tvoja [mama]_ω #, \#sam\-ti [[obećala]_ω sam ti] igračku#.
‘I, your mother, promised you a toy.’

The reason, I contend, that \text{wh}-phrases and clitics behave differently when faced with a PF problem reduces to their differential status as phrases versus heads. Spell-Out chains can only exploit available structure. Crucially, this means that the result of the offending \text{wh}-phrase \text{kakvo} not being able to be pronounced high in (54) is that it surfaces in the next highest \text{phrasal} position. Clitics, on the hand, surface in the next highest \text{head} position.

4.5 All intermediate effects involve heads

One consequence of the present model which does not follow under standard assumptions of successive cyclicity concerns the nature of intermediate effects when they do occur. In general, since the multiattachment chains established for Spell-Out purposes cannot add nodes in the middle of the tree, we expect all intermediate movement effects to involve heads. This prediction is borne out by the \text{wh}-copying phenomenon discussed in Section 3.1.3. Recall the representative minimal pairs from colloquial German in (22) and (23) and Fanselow and Mahajan’s (2000) Berlin dialectal (24), repeated in (57)–(59):

\[(57) \text{[CP} Wen \text{glaubt Hans [CP wen [TP Jakob (*wen) gesehen hat]]]?}
‘Who does Hans believe (who) that Jakob saw (who)?’

\[(58) * \text{[CP Welchen Mann glaubst du [CP welchen Mann [TP sie welchen Mann liebt]]]?}
‘Which man do you believe (which man) she loves?’

\[(59) a. * \text{An wen glaubst du, an wen sie denkt?}
‘Who do you believe she is thinking about?’

\[b. Wovon glaubst du, wovon sie träumt?]
‘What do you believe she is dreaming about?’

Descriptively, the overall pattern for pronunciation of intermediate \text{wh} was as follows: (i) only intermediate \text{wh} elements in \text{COMP} are pronounced, never those inside the source clause, and (ii) only \text{wh}-words are pronounced, not \text{wh}-phrases. This is similar to what happens with the clitics in (55) and (56), except that here all intermediate sites are pronounced, not just the highest one.
How does our model lead to this pattern? The key to understanding what is going on is that the patterns of pronunciation do not have to do with the traditional (and ill-understood) distinction between A- and A’-positions, but rather between head and phrasal positions. Consider the derivation of (57):³⁶

The architecture of my model of Spell-Out basically captures Nunes’s (2004) insight that such cases involve morphological restructuring (“fusion”, in Distributed Morphology terms), although the details are somewhat different:

1. Step I: The [uQ] feature of *wen* is attached to the matrix [+Q]-C and is consequently valued as [+Q]. (This happens upon merge of [+Q] in C and is represented by the solid line; it is analogous to Move F.)

2. Step II: The matrix [+Q] probes down to determine the content of *wen*, becoming entangled with the intermediate C along the way. (This happens for Spell-Out and is represented by the dotted line; it is analogous to Form Chain.)

What does PF do with this? Ordinarily, the wh-phrase *wen* is pronounced in the main clause SpecCP (represented by the dashed line of step III), and nothing else is possible. Any potential intermediate position will be dominated by C⁰, but pronunciation of the wh-phrase there would be morphologically incompatible — a phrase in a head position is ineffable. However, these dialects resort to a strategy which is able to resolve the incompatibility: *wen* is reanalyzed as a head, meaning all of its features
are fused into a single bundle. In and of itself, however, this reanalysis would have no real consequences: treating *wen* as a head would not result in its pronunciation in the intermediate sites — it would just mean that *wen* should be pronounced in the matrix C rather than SpecCP position. Something else must be happening. Indeed, multiple realizations implies that there must be multiple “copies.” However, these are imperfect copies, not created through movement (which is multiattachment after all), but rather through fission (i.e., spinning off of morphosyntactic features) and concomitant disruption of the chain. There is thus a phrasal version of *wen*, pronounced in the matrix SpecCP, and a head version, pronounced in the intermediate C⁰.³⁷

Some evidence for this can be found in the following German dialect example from Fanselow and Ćavar (2001: 18):

> (61) **Welchen Mann** denkst du [CP *wen* [TP er kennt]]?
> ‘Which man do you think he knows?’

This appears to combine elements of (57) and (58). Crucially, what (61) shows is the phrase being pronounced in the matrix SpecCP and a comparable head being pronounced in the intermediate C⁰. The situation is never the other way around. Apparently, the attachment of the phrase to the intermediate head position leads to independent access of the entire set of morphosyntactic features that constitute *wen* (roughly, [+wh, +hum, +acc]). Technically, this could be accomplished in either of two ways. The features could be split off in PF, so that there would be two items, with their own addresses, letting each be pronounced. Alternatively, and I think preferably, the morphological restructuring involves linking the intermediate C⁰ to the feature set that amounts to *wen*, whereas the matrix SpecCP remains linked to *welchen Mann*, schematically as follows:³⁸
This approach raises the obvious question of why other morphologically simplex \textit{wh}-phrases do not undergo similar fission. Nunes’s (2004: 42–3) answer, with which I concur, is that the vagaries of apparent intermediate copy pronunciation are “due not to syntactic computations proper, but to the degree of permissiveness of a given dialect or idiolect to morphological reanalysis,” with the caveat that “the more complex a constituent, the smaller the likelihood that it will undergo” fission.

Also telling is what happens in more deeply embedded contexts. It turns out, according to Rett (2006: 356), that “the copy construction is multiply iterable, but only optionally so”; cf. also Höhle (2000):

\begin{enumerate}
  \item \textit{Wen} glaubt John daß Hans meint daß Mary getroffen hat?
  \vspace{0.5em}
  \hspace{1em} ‘Who does John believe that Hans thinks that Mary saw?’
  \item \textit{Wen} glaubt John \textit{wen} Hans meint daß Mary getroffen hat?
  \item \textit{Wen} glaubt John \textit{wen} Hans meint \textit{wen} Mary getroffen hat?
  \item \textit{Wen} glaubt John daß Hans meint \textit{wen} Mary getroffen hat?
\end{enumerate}

What this means is that, for each attachment site, the morphosyntactic features of \textit{wen} can be independently attached or not.39

Another way of thinking of this phenomenon, suggested by Jason Merchant (p.c.), is that \textit{wen} in C is an instance of some kind of “complementizer agreement.” My
approach “turns this on its head” in that complementizer agreement should itself be seen as akin to \(wh\)-copying. That is, it is another head effect, an instance of \(C^0\) realizing the morphosyntactic features of an associated \(wh\)-phrase. Recall the KiLega example in (4), with relevant aspects reproduced below:

\[
\text{(64) } [\text{CP } \text{Biki } bi- [\text{TP } \text{pro } báàtendilé } [\text{CP } \text{bi- } [\text{TP } \text{pro } báàgúlilé } \text{biki}]]]?
\]

“What did they say they had bought?”

\(Bi\) marks class 8 on all the complementizers, since this is the class to which \(biki\) ‘what’ belongs. Here, every \(C^0\) links (on the dotted step I line) to a single set of valued class features, i.e., those of \(biki\). Assuming \(C\) in this language has a classifier feature set, these will be shared with those of \(biki\) and valued accordingly.

Interestingly, just as in German dialects where \(daß\) and \(wh\)-copying are in complementary distribution, Kinyalolo (1991: 171) points out that complementizer agreement in KiLega is blocked by the presence of an overt complementizer. Compare his ungrammatical (65b), with complementizer \(búno\) ‘that’, and (65a), with no complementizer:

\[
\text{(65) a. } [\text{CP } \text{Kúní } ku- \text{u- } \text{á- búl-ile } \text{17where } \text{17CA-11sg-aspect-perf} \\
[\text{CP } ku- \text{b- } \text{énd-ílélé } \text{Kúní}]]?
\]

‘Where did you say they had gone?'

\[
\text{b. } *[\text{CP } \text{Kúní } ku- \text{u- } \text{á- búl-ile } \text{17where } \text{17CA-11sg-aspect-perf} \\
[\text{CP } \text{búno } ku- \text{b- } \text{énd-ílélé } \text{Kúní}]]?
\]

that \text{17CA-2sa-aspect-perf}
‘Where did you say that they had gone?'

Kinyalolo (p.c.) explains that (65b), with \(búno\) ‘that’, “calls for the absence of \(ku\)- on the embedded verb”:

\[
\text{(66) } [\text{CP } \text{Kúní } ku- [\text{TP } \text{pro } \text{uábúlile } [\text{CP } \text{búno } [\text{TP } \text{pro } \text{bándilé } \text{Kúní}]]]]?
\]

‘Where did you say that they had gone?’

The parallelism with German \(wh\)-copying is striking.\(^{40}\) In both, the initial \(wh\)-phrase becomes entangled with the intermediate \(C^0\), which can either instantiate its morphosyntactic features or be realized as \(daß/búno\) ‘that’.\(^{41}\)

For me, the point that this phenomenon is limited to head marking is crucial.\(^{42}\) If triggered by phrasal movement through SpecCP, as in the standard account, it remains unexplained why the only overt consequence is putative Spec-head agreement. In my system, on the other hand, there are no intermediate Spec positions.\(^{43}\) The realization of features on C in Germanic dialects (and Romani) and KiLega comes instead from the morphological incompatibility of associating a \(wh\)-phrase with a head position.
Other intermediate *wh*-movement effects discussed above, such as Belfast English (2), which involves higher appearance of the auxiliary, and Spanish (3), which involves higher appearance of the verbal complex, are similarly restricted to heads and presumably also obtain from an account in which multilattachment chains are formed for Spell-Out purposes. While coming to grips with the details of these phenomena is beyond the scope of this paper, and I have explicitly put aside head movement, my suspicion is that what is at stake here is that the information attached to T (the auxiliary in Belfast English, the verbal complex in Spanish) is also attached to C (i.e., there is T–to–C “movement”), and that when, under step II, the *wh*-phrase becomes entangled with C, this for some reason causes pronunciation of T at the higher C site. As always, there are additional complexities (e.g., Torrego’s 1984 inversion data suggest entanglement with the lowest C is optional and *why* — which crosslinguistically behaves as if merged in SpecCP — shows no entanglement effects at all). Similar issues of implementation will arise for other well-known asymmetries, such as those incurred by subject versus non-subject or main versus embedded clause *wh*-movement, interrogative versus relative clauses, whether or not “anti-agreement” at the root obtains (cf. fn. 38 above), and so forth. While I hope to tackle such matters in future work, the point remains that, since they can only exploit existing positions, all these effects are limited to head sites.

5. Conclusions and comparisons

The theme of this volume, and the workshop from which it emanates, is to explore new directions in “minimalist” style generative syntax. In doing so, we ask what specific mechanisms are conceptually and/or empirically necessary in order for the syntax to accomplish its minimal task of mediating between form and meaning. In this paper I have subjected to critical scrutiny one of the most persistent of traditional generative syntax constructs — the notion of successive cyclic movement — and proposed a novel architecture to instantiate movement effects. A minimalist model of movement was put forward in which syntax is reduced to relationships between sets of features. Discontinuous dependencies are unbounded and expressed through feature multilattachment. It is only the need to access associated Morpho-Phonological information that leads to the formation of intermediate attachments. Yet, the various phenomena addressed represent but a fraction of what the overall model must ultimately explain. In this final section, I offer a brief summary of my speculations and consider some larger related issues.
5.1 Review of the proposal

Syntax is about establishing discontinuous relationships, and the core syntactic operation is one in which compatible features share their values. As described in Section 4 above, step I entails a [+Q]-C probing and linking with unvalued [uQ] in its scope. This, to my mind, is a truly minimal syntactic operation. Technically, linking in step I means that the “u” [Q] come to share the “+” [Q] information provided to C. And since C merges with the root of the tree, [+Q] can simply search all terminals for [uQ], without regard to structure (because, it necessarily c-commands everything else). Moreover, any wh-phrase in the sentence can thus be valued in this step; it does not matter how “protected” that phrase is, hence [+Q] can point to elements not accessible to the probe in the standard model:

(67) a. Who died [DP the day after filming [TP [DP a movie [PP about [DP a man [CP who discovered what]]]]]]?
   b. Fellini died [PP the day after filming [TP [DP a movie [PP about [DP a man [CP who discovered what]]]]]]??

It is only the subsequent attempt to pronounce what in its scope position that leads to problems:

(68)** What did Fellini die the day after filming a movie about a man who discovered what?

This is the purview of steps II and III, which, taken together, amount to the word order change in the surface string known as “movement.” Step II is the only one which considers structure. In it, the tree is traversed to identify the lexical material containing the [Q] valued under step I, as a prerequisite for step III, which attaches the relevant phrase to the root of the tree, enabling it to be pronounced there.

Pinning down places where different models make different predictions is not easy, since most systems are flexible enough to handle most phenomena. The differences really boil down to which phenomena are taken as a point of departure and what one’s perspectives are on how language “should” work. My model differs from more familiar minimalist systems in a number of ways. It takes as fundamental the feature linking operation (step I) essential for LF interpretation, abstracting away all movement effects as a consequence of Spell-Out to PF (steps II and III). The limitation of overt cyclic effects to heads, as discussed in Section 2, falls out directly from my model, since the syntax only has intermediate heads to work with in the first place. Eschewing successive cyclic movement also offers a simple solution for the “look-ahead” problem created by “failed” wh-movement, as described in Section 3.1.1, and similarly explains the non-existence of successive cyclic LF movement effects. Instead, all intermediate effects are established in the attempt to access lexical material as part of the Spell-Out process, through entanglement with compatible intervening heads. Because my model has both bottom-up and top-down/left-to-right components (cf. also Phillips 2003,
although his “top-down” is purely linear), it allows phenomena to be conceptualized in various ways. For example, for me island effects are top-down rather than bottom up, so that the crash point would be encountered earlier in (69a) than under standard minimalism and the locus of the crash point in (69b) would be the higher rather than lower intermediate wh:

(69)  
  a. *What do you wonder who John told (that) he bought what?
  b. *What do you wonder who John asked where he should buy what?

Another contrast is that, for me, the source of functional information is in the functional categories themselves and not the associated lexical items. That is, a V might receive its past tense feature value from T and an N its accusative case value from Asp, not the other way around. In short, functional categories exist to endow lexical categories with their grammatical properties. And selection of actual lexical material for insertion (“late insertion,” under Distributed Morphology) cannot take place until their grammatical features have been valued. These distinctions are subtle to be sure, and difficult to discriminate.

5.2 Some timing issues

Another major aspect of my model of movement is its need to take the entire structure into consideration. For step I, as (38a) and (67) show, structural details are irrelevant, and step III operates in one fell swoop. It is only step II which examines structure, with entanglement and the possibility of introducing intermediate head effects.

This begs the question of when Spell-Out takes place and, in particular, whether we want to assume some sort of iteration, i.e., “Multiple Spell-Out” (MSO). For MSO to make any sense, it must apply from the inside out, at least as far as step II is concerned. And while I remain agnostic, my best guess is that step III should be able to take place once the feature structures it spells out are complete (i.e., everything is valued). This relates to questions raised by an anonymous reviewer about the dashed line in (36) (the structure for (8b) *What does David wonder [CP where [+Q]-C [TP we bought what where]]?) and whether the wh-island effect is caused by where, which is after all also [+Q], rather than the interfering [+Q]-C, as I argue. But, even assuming MSO, step III cannot apply to where in the lower CP since what is still [uQ], hence cannot be spelled out.45

On the other hand, as Nunes 2004 observes, MSO provides an elegant account of why specifiers and adjuncts are generally strong islands.46 Consider the adjunct clause in (70):

(70)**What did [TP John fall asleep [PP after reviewing what]]?

Here entanglement cannot be the issue. Instead, in order for linearization to proceed, it is proposed that the phrase containing the wh-word must be sent to Spell-Out before
merging. \(47\) This analysis is couched within Nunes’s “sideward movement” model, whereby the Extension Condition is circumvented by employing a separate workspace for non-complement phrases (i.e., phrases that do not merge with heads). This is why the parasitic gap in (71) is acceptable:

\( (71) \quad \text{What did [TP John buy what [PP after reviewing what]]?} \)

Nunes’s system of linearization capitalizes on copies having an independent existence, thus presenting a challenge to show how a multiattachment system can accommodate the facts.

In (71), there is just one instance of what (two would imply a pair-list reading, i.e., What did John buy after reviewing what?), so what is attached as the object of both main clause read and adjunct reviewing. The structure is built up until the matrix [+Q]-C merges. Under step I, this links to the [uQ] of what, which is an unbounded relation. When sent to Spell-Out there are two avenues for accessing the entire phrase — the information needed for PF is attached both inside the main clause and the adjunct clause. Given (70), the attempt that involves the adjunct must for some reason fail, but this ends up being irrelevant, since access to what in the main clause works straightforwardly. The result is that what is pronounced in SpecCP and interpreted in both clauses.

Although the MSO account of why the adjunct in (70) is an island seems sensible to me, I am non-committal about what would force separate Spell-Out of adjuncts and specifiers. If it is the need for unambiguous linearization, then this consideration must override the one that there be no “u” left unvalued. \(48\) Be that as it may, there is however one stipulation required by the copy system which the multiattachment account avoids: copies, once created, must be used immediately, they cannot be “reserved” to be remerged later. Consider a parasitic gap within an island:

\( (72) \quad * \text{What did Bill [TP decide to buy what [CP because John became rich [PP after reviewing what]]?} \)

Imagine that what were first merged with reviewing, then, following Nunes’s sideward movement system, a copy made and the derivation allowed to proceed, merging the adjunct clause after reviewing what with because John became rich and the copy of what reserved to merge eventually with buy. Under this scenario, (72) could be derived; indeed, there would be no islands. Nunes deals with this problem by regarding copying as a Last Resort operation, so that triggers for copying must be locally available. Under the multiattachment approach, on the other hand, there are no copies to be kept in reserve until needed. Because for me the two steps of Nunes’s system — copy and merge — are really one — attach — no stipulation about immediate use is necessary.

A final timing issue concerns the question of whether steps II and III might more perspicuously be collapsed into a single operation. Aren’t these both Spell-Out after all? One reason to keep them separate may be the intermediate binding facts discussed in Section 4.2. In Which picture of herself did Mary say (that) Bill bought?, step I links
the [uQ] of which with the matrix [+Q]-C, but step II also attaches the DP which picture of herself to the intermediate C:\^49

(73) \[_{CP} \, C \, [_{TP} \, Mary \, say \, [_{CP} \, C \, [_{TP} \, Bill \, bought \, D]]]]\]

which picture of herself

Clearly, although the phrase which picture of herself cannot be pronounced under this C, it can be interpreted there. Hence, LF must interpret the results of step II. (On the other hand, although it does not add anything, I can think of no obvious reason why LF should not also be able to see the results of step III.)

The parasitic gap facts just discussed may also shed some light on the issue. Spell-Out must discriminate between (71) and (72). Here is how it might work: even though the after reviewing what clause is spelled out for linearization purposes, the [uQ] of what must remain accessible in (71), but not in (72). This is presumably because what is sent to Spell-Out contains a higher occurrence of what in (71), but not in (72).

Another way to think of this is analogous with feature accessibility in morphology: after reviewing what has a visible [uQ] feature, but once spelled out inside of a larger item (because John became rich) that feature is inaccessible. So if step II is what establishes these successive relations, in a bottom-up fashion, they are disrupted in (72) but not in (71), and this necessarily precedes the final step of attempting to attach the wh-phrase at the [+Q]-C site.

5.3 Conclusion

The general system proposed here should extend, mutatis mutandis, to other kinds of chains. Fleshing out the details is a problem for future work, but the basic idea is this: chains are formed by features probing down the tree, linking to feature sets of similar types, and if these entangled features are already valued, then the search for Morpho-Phonological information ceases. We have seen that an intervening [+Q] on C^9, which I take to mean it is valued, creates island effects in this way. Other types of chains can presumably be formed or disrupted in a similar manner, expressing the insights of Rizzi (1990).

Consider briefly agreement, which I believe similarly illustrates the role played by featural compatibility and the general blocking effect evinced by elements of the same type.\^50 Like a [+Q] probe, agreement features traverse the tree in search of matching features. Taking agreement to be “covaluation” or the sharing of feature values, note that multiple heads can mark the same agreement pattern, as in Bulgarian compound tenses such as the future perfect in (74):
(74) Za malko šteše da izpusneš vlaka.
for little aux.IIsg to miss.IIsg train.DEF
‘You almost missed the train.’

This is because there is no T node intervening between the two targets with [uAgr] features, šteše and izpusneš. On the other hand, long-distance agreement, as described by Bhatt (2005) for Hindi-Urdu or Polinsky and Potsdam (2001) for Tsez, is blocked by an intervening matching head that fails to agree (in Hindi-Urdu, long-distance agreement of a matrix verb with the object of an embedded infinitive is only possible if that infinitive also agrees, and the infinitive can only agree as a side effect of agreement with the embedded object) or an intervening agreeing head (in Tsez, an embedded tense blocks the matrix verb from entering into a long-distance agreement relationship with a lower potential target). I take these facts to mean sets of agreement features can be multiattached and multiply realized, but only when there is no disruption of the chain.51

This is a work in progress. There are many as-of-yet unresolved details about the architecture of the system. Nonetheless, much has been suggested about how “movement” should be understood, how features might work, and the possible mechanics of Spell-Out. It is hoped that these suggestions will lead to new and insightful treatments of much more than the problems examined in these pages.

List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Argument</td>
</tr>
<tr>
<td>A’</td>
<td>Non-argument</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>Agr</td>
<td>Agreement</td>
</tr>
<tr>
<td>ASP</td>
<td>aspect</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary</td>
</tr>
<tr>
<td>BCS</td>
<td>Bosnian/Croatian/Serbian</td>
</tr>
<tr>
<td>C</td>
<td>Complementizer</td>
</tr>
<tr>
<td>CA</td>
<td>complementizer agreement</td>
</tr>
<tr>
<td>CED</td>
<td>Condition on Extraction Domains</td>
</tr>
<tr>
<td>CL</td>
<td>classifier</td>
</tr>
<tr>
<td>Conj</td>
<td>conjunction</td>
</tr>
<tr>
<td>D</td>
<td>Determiner</td>
</tr>
<tr>
<td>DAT</td>
<td>Dative</td>
</tr>
<tr>
<td>ECM</td>
<td>Exceptional Case Marking</td>
</tr>
<tr>
<td>EPP</td>
<td>Extended Projection Principle</td>
</tr>
<tr>
<td>GB</td>
<td>Government and Binding</td>
</tr>
<tr>
<td>hum</td>
<td>human</td>
</tr>
<tr>
<td>LF</td>
<td>Logical Form</td>
</tr>
<tr>
<td>MSO</td>
<td>Multiple Spell-Out</td>
</tr>
<tr>
<td>P</td>
<td>Phrase</td>
</tr>
<tr>
<td>PERF</td>
<td>perfect</td>
</tr>
<tr>
<td>PF</td>
<td>Phonetic Form</td>
</tr>
<tr>
<td>pg</td>
<td>Parasitic gap</td>
</tr>
<tr>
<td>PST</td>
<td>past</td>
</tr>
<tr>
<td>Q</td>
<td>question</td>
</tr>
<tr>
<td>SA</td>
<td>subject agreement</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>Spec</td>
<td>Specifier</td>
</tr>
<tr>
<td>T</td>
<td>Tense</td>
</tr>
<tr>
<td>u</td>
<td>unvalued</td>
</tr>
<tr>
<td>V</td>
<td>Verb</td>
</tr>
<tr>
<td>#</td>
<td>intonational phrase boundary</td>
</tr>
<tr>
<td>ω</td>
<td>prosodic word</td>
</tr>
<tr>
<td>III</td>
<td>3rd person</td>
</tr>
<tr>
<td>II</td>
<td>2nd person</td>
</tr>
</tbody>
</table>
Notes

* This paper expands upon ideas that emerged in discussions with Joshua Herring, who is exploring computational aspects of what he refers to as “call-by-reference” syntax. I am grateful to him for helpful input. The material has also evolved through presentations at Syntaxfest (Indiana University, June 2010), the Slavic Linguistics Society annual meeting (University of Chicago, October 2010), the LSA Winter Meeting (Pittsburgh, January 2011), and the Universities of Edinburgh (May 2012, February 2013), York (February 2103), Venice (March 2013), Nova Gorica (March 2013) and Zagreb (March 2013). I thank those audiences, as well as an anonymous reviewer for this volume.

1. The narrower term “multidominance” is more common, especially as it pertains to how multiattached nodes are linearized. I do not deal with linearization in this paper, although this seems to be the focus of the bulk of the Spell-Out and multidominance literature; cf. e.g. Nunes (2004), Franks (2009), or Citko (2011), and references therein.

2. Another aspect of explicit feature geometry to be explored in work in progress is that agreement — construed as “feature sharing” in the sense of Frampton and Gutmann (2000) — can be partial, multiple, or blocked.

3. Wh-copying is discussed below; the other effects of delinking are treated in work in progress.

4. Hornstein, Lasnik, and Uriagereka (2003) also propose that islands are the consequence of Spell-Out.

5. I employ [+Q] for the feature on C⁰ which marks the domain of interrogative scope, rather than [+wh], which belongs to the wh-word. The latter also has an unvalued [uQ] feature which is ultimately valued as [+Q] by the [+Q]-C (in my system, through linking). I take [+Q] to subsume both [−wh] Yes/No and [+wh] wh questions, which is why, in some languages, these can coordinate (cf. e.g. Zanon 2014).

6. Another famous example is the particle aL in Irish, which McCloskey (2001) shows to mark wh-agreement. These differ from French (i), cited by Takahashi (1994: 61), since here participles only agree with local fronted objects:

(i) Combien de chaises a-t-il cru/*crues que Marie a repeintes?
   ‘How many chairs has he thought that Marie has painted?’

7. Examples such as (6) are widely discussed in the literature for various Germanic dialects (and Romani); for analysis and data see in particular Nunes (2004: 38–43), Felser (2004), and Rett (2006), as well as references therein.

8. The intermediate binding argument dates to Barss (1986).


10. In his survey of “Alternative Views on Successive Cyclicity,” Boeckx (2008: Section 6.4) characterizes all these approaches as “Greed-based.”
11. An anonymous reviewer suggests (12b) could be excluded because *thinks* selects for a declarative rather than interrogative complement. This, however, is an LF rather than PF matter, whereas the decision of which copy to pronounce — generally, the highest/first one — is resolved in the mapping to PF.

12. This constraint is part of a larger family of PF constraints against sequences of homophonous elements.

13. He argues that features rather than phrases move (“Move F” of Chomsky 1995), but also suggests that the second *wh*-phrase might move in one fell swoop, since the first has already respected Subjacency (cf. Richard’s 2001 “Principle of Minimal Compliance”).

14. Note that judgments are identical even if the adjunct parasitic gap clause *fără să influențeze* ‘without influencing’ is added. As expected, *Cine ce crede Ion că a determinat?* ‘Who does Ion think determined what?’, with multiply fronted non-homophonous *wh*-phrases, is perfect.

15. This example is modified from Bošković (2002). First, *Ivan* is topicalized to make it sound more natural, and second, his reported judgments are modified because Bulgarian speakers I have asked find *kakvo* ‘what’ absolutely impossible before *če* ‘that’ and stylistically marked before the verb *obuslavja* ‘conditions’.

16. Du Plessis (1977) and this particular example are repeatedly cited in the literature, but I know of no subsequent study of *wh*-copying in Afrikaans.

17. It is sometimes claimed (e.g. Boeckx 2008, Fox 2008) that LF movement exhibits cyclic scope effects.

18. In (30) control of PRO is irrelevant, since neither putative intermediate SpecCP *wh*-phrase can license the parasitic gap.

19. Thanks to Yuyin Hsu (p.c.) for judgments and glosses.

20. My ideas about Spell-Out are conceptually close to those developed in Ochi’s (1999) articulation of the dual chain formation system of Chomsky (1995). For Ochi, category (as opposed to feature) chain formation was a kind of pied-piping motivated by the need (of the moving element, hence “Greed”) to overcome the “feature-scattering” PF deficiency created by formal feature movement. Of course, in my system there is no literal movement, hence no deficiency, just the need to access the category in order to pronounce it.

21. I do not deal with T-to-C movement in this paper; (33) thus represents the embedded portion of *I wonder which book Julia should read*. Other irrelevant details pertaining to the internal structure of the clause, such as vP and multiattachment/movement of subject and object, are put aside. Pulling the [uQ, +wh] feature out of the D head *which* is a shorthand representation for the assumed hierarchical/nanosyntactic geometric structure; -ich abbreviates the feature subtree left behind, namely, *which* minus its *wh* features. Interestingly, whereas for interrogative *which* we want [+wh] to dominate [Q], since the relative pronoun is [+wh, −Q], for C we want [+Q] to dominate [wh], since Yes/No questions are [+Q, −wh]. Space considerations preclude representation of these details in the (simplified) diagrams below.
22. To make it clear that it corresponds to movement (in that the multiattached phrase will be pronounced in its higher occurrence), the dashed line of step III is represented throughout this paper with an arrow going up the tree.

23. Since, as just described, trees are ultimately data structures, with nodes understood as addresses pointing to (calling up) other addresses, the reader should not ascribe any special formal significance to the various metaphorical “attachment” terms employed descriptively in this paper.

24. An anonymous reviewer wonders if this admits violations of Chomskyan conditions such as the strict cycle or the bans against tampering and internal extension of the tree. I think not, assuming these lower objects still have features awaiting valuation hence have not yet been sent to Spell-Out.

25. Movement is independently blocked to $C^0$ in (33) because should is there.

26. As an anonymous reviewer points out, this system recreates Rizzi’s (1994) Relativized Minimality in that only intervening data sets of similar “type” interfere. The reviewer suggests that, while I cast entanglement in derivational terms, since the entire structure is present at this point, it could also be conceived of as a representational constraint.

27. The diamond which terminates the dotted line from the matrix [+Q] indicates that the search is blocked. Also note that I marked the dashed line as “?” rather than numbering it “3” because, as discussed in Section 5.2 below, embedded where in (36) should actually never have an opportunity to move.

28. The kinds of phenomena discussed in Bošković (1997: ch. 3), whereby ECM/object shift (movement through SpecAgr$_O$P) in wager-class verbs and French propositional infinitivals only takes place if forced by wh-movement, are readily amenable to my system as well. Consider French (i):

(i) Qui [Pierre croit-il [Agr$_O$ [qui avoir acheté des fraises]]?]

‘Who does Pierre believe to have bought some strawberries?’

Curiously, ECM in *Pierre croit Marie avoir acheté des fraises. is ungrammatical: the embedded subject only receives case if it undergoes overt wh-movement. For me, Spell-Out causes qui to be linked to Agr$_O$ in (i), but no chain is formed otherwise.

29. An anonymous reviewer questions the relevance of these examples, in which the purpose clause is attached to the claim clause. But it is the intermediate C (after say) which is crucial to licensing the pg, since (i), in which the intermediate C (after asked) is [+Q], leads to unacceptability:

(i) *Which painting did Karen buy [because David asked [why she liked pg]]?

30. A [+Q, −wh]-C presumably links with some feature of T.

31. How large a constituent “pied-pipes” is a vexed question. My suspicion is that step III targets the maximal unit bearing the relevant feature ([+Q], in this case), which is minimally the phrase projected by the wh-word at hand.
32. Although the consequences are different (feature valuation, cyclicity effects, movement), it should be borne in mind that all three operations are ones of pointing to a node’s address, i.e., introducing a new attachment.

33. Thanks are due to Teodora Radeva-Bork and Lilia Schuerks (p.c.) for corroboration and discussion.

34. Prosody of course also matters. In fact, if the second adjacent wh-word in (52b) and (53b) or examples (18a) and (21) above receives special intonation it becomes more acceptable. This fact only serves to highlight the superficial nature of the restriction, making it even more mysterious why intermediate SpecCPs are unavailable (under the traditional movement account).

35. Crucially, wh-phrases that consist of a single word in PF must still project a phrase, otherwise, as an anonymous reviewer observes, they might end up pronounced under C⁰. (See Section 4.5 for a fission approach to wh-copying.)

36. Here, for ease of representation, I collapse the feature structure. I also abstract away from irrelevant details of German clause structure. My view however is that German is head-initial and that both T⁰ and the verb — hat and geschehen in (60) — end up final due to “roll-up” phrasal movements.

37. An anonymous reviewer asks whether the matrix instance of wen could be analyzed similarly. Indeed it could, but (61), in which matrix Welchen Mann is phrasal and embedded wen is a head consisting exhaustively of morphosyntactic features, shows that this kind of analysis is not always viable.

38. Recall that the up arrow on the dashed line means “pronounced here/movement” although technically, like all the other links, it is a call down the tree to the information located at an address. Note also that the matrix C⁰ does not become entangled: Welchen Mann (*wen) denkst du wen er kennt? This “anti-locality” effect is, I believe, the same as that observed more generally for “anti-agreement”: as noted by Ouhalla (1993: 480) for Tamazight Berber, anti-agreement only “arises when the subject is extracted and moved to the nearest SpecCP position” and not “when the subject is extracted over a long distance, e.g., from an embedded clause.” (On the other hand, as seen below, main clause (i.e. local) wh-agreement obtains in KiLega, and realization of main clause T features in C (i.e., matrix T–to–C movement) is widespread.)

39. Comparisons with so-called “partial” or “scope marking” agreement, which most of the wh-copying literature also addresses, are tempting. I put these aside due to lack of space.

40. As for (65a), without bùno ‘that’, Kinyalolo (p.c.) adds that leaving ku- off the embedded verb “would seem like there is a direct question following.”

41. It is not possible to determine whether the kind of independent variation reported for German (63) is also possible here, since KiLega speakers are not comfortable with additional embedding.

42. Irish aL, which as described in McCloskey (2001) marks wh-agreement, is also exclusively a head effect.

43. Although this follows because the Extension Condition (11b) is crucially retained, it falls out naturally. I believe, from the system. If nodes are understood as addresses pointing to sets
of other addresses, adding structure internal to the tree would entail insertion of a new address in media res.

44. For a recent comparison between multidominance and copy theories of movement, see the discussion in Larson and Hornstein (under review). Ultimately they conclude that one can "translate Copy talk into Occurrence talk without empirical loss" and that "the systems look different though they do the same things."

45. This means that the reason (8b) is worse than (37a) ?*What does David wonder whether we bought? cannot be because the combination of [+Q, +wh] where in SpecCP and the [+Q]-C somehow renders the embedded clause in (8b) "doubly impenetrable." Moreover, this kind of account would leave the likewise degraded status of (37b) ??What does David believe the/Julia’s claim that we bought? unexplained. In Section 4.2 it was instead suggested that entanglement is sensitive to similarity of features, so that a probing [+wh, +Q]-C clashes most with another [+wh, +Q]-C, less with a [−wh, +Q]-C, and least with a [+def]-D.

46. These are the Condition on Extraction Domains (CED) effects of Huang 1982. Note that the CED only restricted overt movement.

47. From the perspective of Kayne’s (1994) asymmetric c-command approach to linearization, a specifier once spelled out not only constitutes an opaque, impenetrable whole, but also looks like a head, hence asymmetrically c-commands and precedes all terminals within its complement:

(i) **Who did [_{TP [_{DP stories about who surprise the teacher}]}]?**

48. If there exist features which never get valued in the syntax and which the morphology eventually fills in "by default,” then it cannot be that Spell-Out never tolerates any unvalued features.

49. One mystery I put aside here is why the intermediate binding effect is absent in so many other languages, including Slavic ones. For me, this would have to be a matter of whether or not attachment to this intermediate C occurs. See Kang (2013) for another idea.

50. The idea that movement and agreement are two sides of the same coin is hardly novel, although agreement is more usually regarded as syntax rather than movement as PF; see for example Miyagawa (2010).

51. The architecture of the system is inspired by phonological feature geometry; cf. also Bonet (1995) for morphology. It is thus unsurprising that these effects seem comparable to the spreading of features in phonology, where geometrical considerations are at play.

References


Larson, Bradley and Hornstein, Norbert. under review. “Copies and occurrences.”


