Wh-Copying, Phases, and Successive Cyclicity[1]

Claudia Felser

1.   Introduction

Within the minimalist framework outlined by Chomsky (1998, 1999, 2001), the derivation proceeds 'by phase', that is, syntactic structures are built in a bottom-up, one-phase-at-a-time fashion. Phases, which according to Chomsky include the 'propositional' categories CP and (transitive) *v*P, can thus be seen as defining local computational domains. Phases are constructed by successive application of the two basic structure-building operations *Merge* and *Move*. Overt movement, which presupposes abstract agreement, is induced only by heads that carry an EPP feature. Agreement - and hence, movement - is triggered by the need to eliminate uninterpretable features of both the attracting head (the *Probe*) and the attractee (the *Goal*).[2] The requirement that movement must result in feature-checking is known as the *Last Resort* condition (Chomsky 1995: 280). The operation *Spellout* applies cyclically in that each phase is 'spelled out' - that is, is passed on to the phonological and semantic systems for evaluation and interpretation - at the point at which the next higher phase is completed (compare Chomsky 1999: 10). The idea that phases constitute relatively independent units of computation is captured by the *Phase Impenetrability Condition* (PIC):

(1)   *Phase Impenetrability Condition*:
        In phase α with Head H, the domain of H is not accessible to operations outside
        α, but only H and its edge.

(Chomsky 1998: 22)

---

[2]   In Chomsky's (1999) terminology, uninterpretable features are actually *unvalued* features that need to be assigned a PF-value through agreement, which then allows for them to be deleted from the representation. As nothing crucial hinges on the use of any specific terminology here, I shall continue to use the terms 'uninterpretable feature' and 'checking' as in Chomsky (1995).

That is, once a phase has undergone spellout, all elements contained within it, with the exception of the head and its specifier(s), become inaccessible to further syntactic operations. A derivation will crash (at the phase level) if one or more elements within a phase undergoing spellout still contain any uninterpretable features.
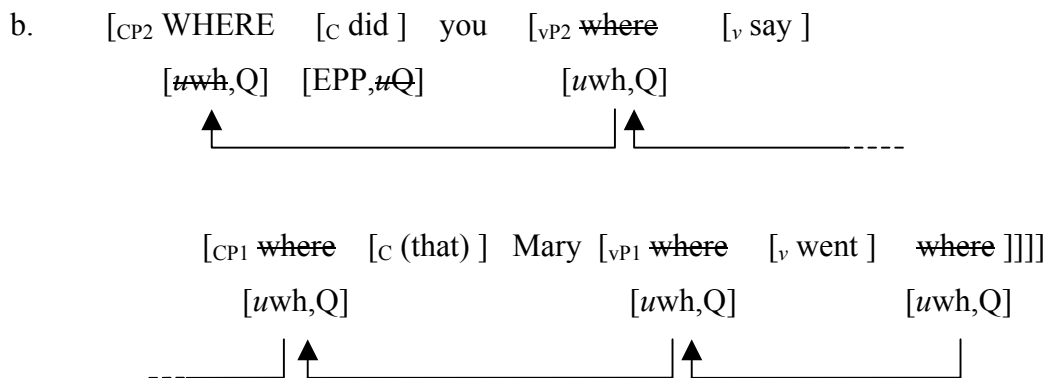
The assumption that derivations proceed by phase serves to reduce 'operational complexity' in that for each derivational cycle, only a subarray of lexical items including a single C or *v* only needs to be drawn from the lexicon and held in active memory (Chomsky 1998: 19f.). Although the minimalist framework does not, of course, claim to be a theory of syntactic processing, Chomsky's appeal to memory load could be seen as reflecting a more general trend to take seriously questions such as the extent to which the grammar may be constrained by requirements of the performance systems or by general cognitive limitations, and how grammatical knowledge and the processing system interact (compare, among others, Chomsky 2001, Hawkins 1994, Jackendoff 1999, Phillips 1996, or Steedman 2000). Working memory (WM) limitations in particular are known to impose severe constraints on our capacity to produce and comprehend sentences (Just & Carpenter 1992, King & Just 1991). From the point of view of language processing, cyclic spellout would seem to make sense insofar as it helps minimise the computational or memory cost incurred by keeping (semantically and phonologically 'complete') partial phrase markers active in working memory.

Several authors have noted, however, that Chomsky's phase-based approach to syntactic derivation, in conjunction with the copy theory of movement, raises new questions about the nature of, and motivation for, certain types of successive-cyclic movement (compare, among others, Atkinson 2000, Heck & Müller 2000, McCloskey 2000a, and Radford 2001). Consider long *wh*-raising, the prototypical case of an unbounded dependency. The idea that long *wh*-movement proceeds in a series of local steps has for a long time been a core assumption of generative-transformational theory, and appears to be supported by a large body of empirical evidence from a variety of sources. Many languages, for example, exhibit morphophonological reflexes of successive-cyclic *wh*-raising such as complementiser agreement in Irish (McCloskey 2000a,b, 2001), deletion of the verbal prefix *men-* in Malay/Bahasa Indonesia (Cole & Hermon 2000, Saddy 1991), or tonal downstep in Kikuyu (Clements et al. 1983, Sabel 2000). Other standard arguments in support of successive-cyclic *wh*-movement include floating quantifiers in Irish English dialects (McCloskey 2000b,c), certain types of reconstruction effect (Barss 2001, Fox 2000), partial *wh*-movement in a variety of languages including Malay/Bahasa Indonesia (Cole & Hermon 2000, Saddy 1991) and Ancash Quechua (Cole 1982), children's use of 'medial' *wh* in long-distance questions such as *Who do you think who's in the box?* (De Villiers et al.

1990, McDaniel et al. 1995, Thornton 1990), and *wh*-copying in languages like German and Romani (Höhle 2000, McDaniel 1986, Reis 2000).

According to Chomsky (1998), *wh*-expressions carry both an uninterpretable [wh] feature that renders them 'active', i.e. available for movement, and an interpretable operator feature [Q]. In *wh*-raising languages such as English, interrogative C - in addition to an uninterpretable [Q] feature that must be eliminated through agreement - contains an EPP feature that triggers overt movement of the closest available *wh*-expression. The uninterpretable [wh] feature carried by the latter will be checked under Q-agreement. For long *wh*-raising structures such as (2a) below, Chomsky assumes that the *wh*-expression moves through the edge of each intervening phase before finally reaching its ultimate landing site, the specifier of the matrix C, where it will also be pronounced (the prefix *u* indicates that a feature is uninterpretable).

(2)　　a.　　Where did you say (that) Mary went?

　　　　b.　　[$_{CP2}$ WHERE　[$_C$ did]　you　[$_{vP2}$ ~~where~~　[$_v$ say]
　　　　　　　[~~*u*wh~~,Q]　[EPP,~~*u*Q~~]　　　　　[*u*wh,Q]

　　　　　　　[$_{CP1}$ ~~where~~　[$_C$ (that)]　Mary [$_{vP1}$ ~~where~~　[$_v$ went]　~~where~~]]]]
　　　　　　　[*u*wh,Q]　　　　　　　　　[*u*wh,Q]　　　　　　[*u*wh,Q]

Given the Phase Impenetrability Condition (1), successive-cyclic movement through the edge of intervening phases is necessary to ensure that the *wh*-expression remains accessible to further syntactic operations once the remainder of the phase containing it has undergone spellout.

A question that immediately arises here though is how movement to intermediate positions should be triggered given that both *v* and declarative C are, presumably, unable to check the uninterpretable [wh] feature of the moved *wh*-expression. Indeed, if they did, the *wh*-expression would no longer be active, thus leaving the uninterpretable [Q] feature of the matrix C unchecked. The *Triggering Problem* is stated informally in (3).

(3)   *The Triggering Problem:*
        On the assumption that agreement (and hence, movement) is triggered by matching but
        uninterpretable features of the probe, what triggers movement of a *wh*-expression to the
        specifier of intermediate non-interrogative heads?

A second problem with the derivation sketched in (2b) arises from the assumption that spellout applies automatically to phases at the next phase level up (compare also Atkinson 2000). Note that at the point at which vP1 undergoes spellout (i.e., upon completion of the next higher phase, CP1), the copy of the moved *wh*-expression in (Spec,vP1) still contains an uninterpretable [wh] feature. This will not be a problem at this stage in the derivation if, as Chomsky suggests, material at the edge of a phase PH_n can escape spellout in the sense that it remains 'visible' to probing heads within the next higher phase PH_{n+1}.[3] However, on the assumption that no look-ahead is possible beyond PH_{n+1}, it is unclear why, at the point at which PH_{n+1} is spelled out (i.e., upon completion of PH_{n+2}) the presence at the edge of PH_n of a copy of a *wh*-item still carrying an uninterpretable [wh] feature does not cause the derivation to crash. Let us refer to this problem - which in fact occurs whenever successive-cyclic *wh*-movement needs to cross more than one phase boundary - as the *Convergence Problem* (4).

(4)   *The Convergence Problem:*
        If phases undergo cyclic spellout at the next higher phase level, then why does the
        spelling-out of phases containing a copy of a *wh*-item that is still active at the point of
        spellout not cause the derivation to crash?

Atkinson (2000) notes that the convergence problem would disappear if phases were defined on the basis of convergence rather than in absolute (i.e., categorical) terms. If only convergent phases can undergo spellout, then it will not be until the matrix CP (= CP2) in (2b) has been

---

[3]   On the assumption that at any given stage in the derivation, all copies created thus far share the same
        feature values, one might wonder to what extent the copy at the *foot* of the chain presents a problem for
        spellout at this stage, though.

completed that anything at all is sent to the interfaces.[4] Observe, however, that allowing for more than two (or in fact, a potentially unlimited number of) phases being kept in working memory at the same time would appear to run counter to the main motivation for introducing the concept of 'phase' in the first place, namely the acknowledgement of the fact that working memory, or computational work space, is limited. I will return to this issue in section 3.3 below.

Elaborating further the idea that certain syntactic operations may apply non-locally, Radford (2001: 37) suggests that both the convergence problem and the triggering problem could be overcome at the same time if it were assumed that rather than creating a trail of phonetically null copies, long *wh*-raising takes place in a single step. Let us refer to this hypothesis as the *Single-Step Hypothesis* (SSH):[5]

(5)   *The Single-Step Hypothesis*:
        There are no intermediate steps in long *wh*-movement.

Under this view, the drivation of (2a) above would involve the following steps (note that the representations below are simplified, and that only features that are directly relevant to the present discussion have been included):

---

[4] Chomsky (1998: 20ff.) also discusses the possibility that phases are defined in terms of convergence, but rejects it on the grounds that it would render lexical selection less economical, and because it occasionally appears to yield the wrong empirical results. By way of illustrating the latter point, Chomsky (1998: 21) observes that if phasehood was dependent upon convergence, then the constituent labelled α in (i) below would not constitute a phase. Since phases are constructed on the basis of lexical (sub-)arrays, the initial lexical array must be larger than the set of items required for constructing α. If the extended lexical array included expletive *there*, however, then nothing would block 'premature' insertion of the expletive in (Spec,TP) of the lower clause, which would have the undesirable consequence of preventing the subject DP *John* from raising to this position.

(i)   Which article is there some hope [$_\alpha$ that John will read t_{wh}]

Note, however, that this problem does not arise under the assumption that *there*-type expletives are quasi-argumental (as has been proposed by Felser & Rupp 2001), which implies, among other things, that the possibility of merging an expletive is constrained by a predicate's selectional properties. An alternative solution to the above problem has been offered by Atkinson (2000: 118ff.), who suggests that we might need to distinguish between two different notions of 'phase', a lexical selection phase and a derivational phase, with the latter but not the former being defined in terms of convergence.

[5] See Postal (1972) for an earlier suggestion to this effect.

(6) a. *Build vP1:* [$_{vP1}$ Mary [$_{v'}$ went *where*]]
$$[u\text{wh,Q}]$$

b. *Build CP1:* [$_{CP1}$ [$_{C}$ (that)] Mary [$_{vP1}$ ... [$_{v'}$ went *where*]]]
$$[u\text{wh,Q}]$$

c. *Build vP2:* [$_{vP2}$ you say [$_{CP1}$ [$_{C}$ (that)] Mary ...

... [$_{vP1}$ ... [$_{v'}$ went *where*]]]]
$$[u\text{wh,Q}]$$

d. *Build CP2:*

[$_{CP2}$ WHERE [$_{C}$ did] you [$_{vP2}$ ... say [$_{CP1}$ [$_{C}$ (that)] Mary ...

[$\cancel{u\text{wh}}$,Q] [EPP,$\cancel{u Q}$]

... [$_{vP1}$ ... [$_{v'}$ went ~~where~~]]]]]

[$\cancel{u\text{wh}}$,Q]

Under the assumption that only convergent phases can undergo spellout, its application will be delayed until CP2 has been completed, and the uninterpretable [wh] feature of *where* has been checked by interrogative C.

If the SSH is correct, however, then how do we account for the various types of evidence suggesting that operator movement takes place in successive-cyclic fashion? Radford (2001) argues that many of the phenomena traditionally claimed to support successive-cyclic *wh*-raising are in fact reanalysable without assuming local *wh*-movement to intermediate positions. He notes, for example, that what appear to be morphophonological reflexes of intermediate movement steps might simply be reflexes of abstract agreement in the sense of Chomsky (1998). Examples of quantifier 'stranding' in intermediate (Spec,CP) positions are also open to alternative analyses, perhaps in terms of base adjunction of Q-elements to projections of intermediate C heads, along the lines suggested by Radford (2001: 48). Upon critical re-examination, other arguments in

favour of intermediate steps in *wh*-movement also turn out to be far from compelling. Note, for instance, that the force of any arguments derived from reconstruction effects depends crucially on our assumptions about how reconstruction works - specifically, whether it is contingent upon movement, abstract agreement, or neither of these two - and at what linguistic level it applies. As regards partial movement in languages like Malay/Bahasa Indonesia, the only evidence that a *wh*-expression in the specifier of a local C ever undergoes any further raising is, again, semantic in nature. What is more, once we abandon the (strong) continuity view on language development, the significance of the evidence from child language is also called into question. Specifically, one could argue that children's use of medial *wh* results from misanalysis of sentence structure, or represents a particular type of surface strategy of asking long *wh*-questions, rather than revealing covert properties of the adult grammar.[6]

The presence of identical PF-copies of a moved *wh*-expression at intermediate positions in the adult language would seem to constitute a more serious challenge for the SSH though. The *wh*-copying strategy of asking long *wh*-questions is attested in numerous languages including German, Frisian, Afrikaans, and Romani.

(7)   a.   *Wen* glaubst Du, *wen* sie getroffen hat?   GERMAN

             who think   you who she met   has

             'Who do you think she has met?'

   b.   *Wêr* tinke jo *wêr't*   Jan wennet?   FRISIAN

             where think you where that-CL J.   resides

             'Where do you think that John lives?'

                                            (Hiemstra 1986: 99)

---

[6]   Based on the observation that children would frequently answer the embedded 'who' question rather than the matrix question in sentences such as *How did Big Bird say who to paint?*, de Villiers et al. (1990), for example, have claimed that children's initial grammars lack successive-cyclic *wh*-movement altogether. According to Roeper (1990), children's use of medial *wh* reflects an intermediate stage at which they have acquired local movement but have not yet acquired adult-like long-distance movement.

c.   *Waarvoor* dink julle *waarvoor* werk ons?   Aᴀғʀɪᴋᴀᴀɴs

         wherefore think you   wherefore work we

         'What do you think we are working for?'

                                                  (du Plessis 1977: 725)

d.   *Kas*   o Demìri mislenola *kas*   i Arìfa dikhla?   Rᴏᴍᴀɴɪ

         whom Demir think   whom   A.   saw

         'Who does Demir think Arifa saw?'

                                     (adapted from McDaniel 1989: 569n.5)

If the intermediate copy proves to be 'real' in the sense that it represents a phonetically realised trace of a moved *wh*-expression higher up in the sentence, thus reflecting an earlier stage in the derivation, then the SSH cannot be correct, at least not for languages that exhibit the *wh*-copying phenomenon.

      The main goals of this paper are to assess the validity of the SSH in the light of what appear to be overt reflexes of intermediate movement steps, and to sketch a possible way of overcoming the more general difficulties that successive-cyclic movement poses for a phase-based theory of syntax that were noted above. In the next section, I will examine the German copy construction in some detail, showing that *wh*-copying does indeed present one of the most compelling pieces of evidence in favour of successive-cyclic *wh*-movement. The finding that *wh*-copies are best analysed as spelled-out intermediate traces of successive-cyclic *wh*-movement is argued to support a convergence-based view of 'phase', but leaves open the answer to the Triggering Problem (3). The latter will be addressed in section 3, where I argue that in the absence of any plausible formal trigger for local *wh*-raising to non-interrogative heads, processing considerations might provide the key to answering the question of what should motivate intermediate steps in long *wh*-movement. The main points and conclusions are summarised again in section 4.

2.   *Wh-copying in adult German*

2.1   *Preliminary observations*

Many varieties of present-day colloquial German permit the use of the *wh*-copying strategy, which is exemplified by (8a-c) below.

(8)   a.   *Wie* glaubst Du, *wie* sie das gelöst hat?

how believe you how she that solved has

'How do you believe that she has solved that?'

b.   *Warum* glaubst Du, *warum* sie das getan hat?

why   believe you why   she that done has

'Why do you believe she has done this?'

c.   *Wovon* glaubst Du, *wovon* sie träumt?

of.what believe you of.what she dreams

'What do you believe that she dreams of?'

(Fanselow & Mahajan 2000: 220)

Wh-copying is possible only with verbs that permit long *wh*-extraction from a finite complement clause (so-called 'bridge' verbs), and which select a non-interrogative complement. Verbs selecting interrogative complements such as *fragen* 'ask', on the other hand, do not license the copy construction (cf. [9c]).[7]

---

[7]   The copy construction has often been treated on a par with the so-called 'scope marking' or *was...w* construction, in which the 'expletive' *wh*-pronoun *was* 'what', rather than a contentful *wh*-phrase, occcupies the (Spec,CP) position of the matrix clause.

(i)   *Was* glaubst du, *wovon*   sie träumt?
      what believe you of.what she dreams

The observation that the copy construction shares several of the properties of the *was...w* construction has led many researchers to conclude that the two are mere variants of each other, both of which represent special instances of long-distance *wh*-movement (Bayer 1996, Höhle 2000, among others). This assumption is called into question, however, by the fact that the copy construction but not the *was...w* construction patterns with long-distance *dass* questions in a number of respects. These include the possibility of conjoined questions appearing in the embedded clause (Dayal 2000), the availability of a wide scope reading of the embedded *wh*-expression (Pafel 2000, von Stechow 2000), the acceptability of *scheinen* 'seem' as a matrix predicate (McDaniel 1986, cited in Höhle 2000), and the possibility of intervening *dass* clauses (Reis 2000). Observe further that in contrast to the *was...w* construction, the copy construction does not normally allow full *wh*-phrases to appear at intermediate positions. In Felser (2001), I argue in favour of an 'indirect dependency' analysis of the *was...w* construction in the spirit of Dayal (1994) according to which *was* is a CP-proform and argumental, and

(9)   a.   Du fragst/*glaubst, *wovon* sie träumt.

                   you ask / believe of.what she dreams

        b.   *Wovon* glaubst /*fragst du, dass sie träumt?

                   of.what believe / ask   you that she dreams

        c.   *Wovon* glaubst /*fragst du, *wovon*   sie träumt?

                   of.what believe / ask   you of.what she dreams

Not all speakers of German accept the copy construction, but as Höhle (2000: 257n.7) observes, its use is not obviously linked to any particular dialect areas, either. In standard varieties of German, the *wh*-copying strategy appears to be restricted to pronominal *wh*-expressions. That is, most of the speakers who accept sentences like (9a-c) above do not accept *wh*-copying structures that involve non-pronominal *wh*-expressions. The examples below, for instance, are judged to be ill-formed by Fanselow & Mahajan (2000) - although they note that speakers' judgements tend to be less consistent for examples like (10b) that involve a prepositional *wh*-phrase.[8]

(10)   a.   **Welchen Mann* glaubst Du, *welchen Mann* sie liebt?

                  which   man   believe you which   man   she loves

             'Which man do you believe that she loves?'

---

the embedded clause a secondary predicate. An analysis along these lines appears to be supported by the following contrast noted by Hinrichs & Nakazawa (2001) (my translations):

(ii) Was Hans sagt, wen er verdächtigt, das/*den habe ich überprüft.
      what H.   says who he suspects   that/*him have I   evaluated
      'I checked what Hans says as to whom he suspects.'

(iii) Wen Hans sagt, wen er verdächtigt, *das/den habe ich überprüft.
      who H.   says who he suspects   *that/him have I   evaluated
      'I checked the person who Hans says he suspects.'

If *was* 'what' in (ii) were a mere placeholder for *wen* 'who', then the resumptive pronoun should be *den* 'him' rather than *das* 'that', as in (iii).

b.   **An wen*   glaubst Du, *an wen*   sie denkt?

of whom believe you of whom she thinks

'Who do you believe that she thinks of?'

In short, it appears that only *wh*-phrases that are 'single morphophonological words' (Fanselow & Mahajan 2000) can normally be spelled out at an intermediate position.

Observe further that *wh*-copies are licensed only in derived positions but not in situ (cf. [11a]), and that in the absence of an overt complementiser, a copy *must* appear at the left periphery of the lower clause (cf. [11b]).

(11)   a.   **Wovon* glaubst du   dass sie *wovon*   träumt?

of.what believe you that   she of.what dreams

b.   *Wovon* glaubst du *(*wovon*)   sie träumt?

of.what believe you (of.what) she dreams

In *wh*-copying structures that involve multiple embedding, each intervening clause must be introduced by a separate *wh*-copy (compare Reis 2000: 395):

(12)   *Wen* glaubst du, *wen* / *dass Peter meint, *wen* Susi heiratet?

who believe you who / *that P.   thinks who S.   marries

'Who do you believe Peter thinks that Susi is marrying?'

It would seem, then, that the copy construction necessarily involves partial movement, and that each intermediate copy must be locally A'-bound by an identical copy in the next higher clause.

---

[8]   Höhle (2000: 258), for instance, only assigns a question mark to an example similar to (10b) (his [19e]); cf. also Höhle's note 8.

2.2   Are intermediate copies real?

If intermediate copies are indeed overt realisations of traces left behind by successive-cyclic movement, then their existence seriously calls into question the Single-Step Hypothesis (5). As has been pointed out by Nunes (2000), the presence of multiple PF occurrences of a syntactic item is also potentially problematic from the point of view of Kayne's (1994) *Linear Correspondence Axiom* (LCA), which maps hierarchical into linear order. Note that the presence of multiple copies would not pose a problem for either the SSH or the LCA if it could be shown that intermediate *wh*-items are not what they appear to be (i.e., spelled-out traces at intermediate [Spec,CP] positions). Let us therefore examine the possibility that *wh*-copying does not involve multiple PF-occurrences of moved *wh*-pronouns at all, but instead represents a special case of complementiser agreement, analogous to *wh*-agreement in Irish. In Irish, the complementiser *go* that normally introduces finite complement clauses (cf. [13a] below) is replaced by *aL* in clauses out of which operator movement has taken place as in (13b).

(13)   a.   Creidim gu-r   inis sé bréag.
                   I-believe **go**-PAST tell he lie
                   'I believe that he told a lie.'

          b.   an t-ainm a   hinnseadh dúinn a   bhí ar an áit
                   the name   **aL** was-told   to-us **aL** was on the place
                   'the name that we were told was on the place'

                                                               (McCloskey 2000a: 4f.)

An analysis of *wh*-copying in terms of complementiser agreement has been proposed by Thornton & Crain (1995) for children's medial *wh*-sentences. Under this view, the intermediate 'copy' is actually a complementiser that agrees with a long-distance moved *wh*-expression to the point of phonetic identity.

(14)   [CP2 *Wovon*$_i$ [C' glaubst du [CP1 (*t*$_i$) [C' [C wovon] sie   *t*$_i$ träumt]]]]
                    of.what   believe you   COMP   she   dreams

To the extent that complementiser agreement can be triggered by application of *Agree* alone, *wh*-copying provides evidence only for local agreement but not for successive-cyclic movement.[9]

There are several arguments against an analysis of the copy construction in terms of complementiser agreement, though. For one thing, on the assumption that intermediate copies are *wh*-agreeing variants of the declarative complementiser *dass*, the *wh*-agreement analysis incorrectly predicts that the copy construction and the corresponding *dass* constructions should pattern alike syntactically. Among other things, the copy construction but not ordinary long-distance questions exhibits the following 'island effects' (the (b) examples are borrowed from Reis [2000: 395] and Fanselow & Mahajan [2000: 219f.]):[10,11]

(15) a. *Wen* glaubst du nicht, dass sie liebt?

          whom believe you not   that she loves

          'Who don't you think that she loves?'

    b. *Wen glaubst du nicht, *wen* sie liebt?   (Negative Island)

          whom believe you not   whom she loves

(16) a. *Wen* möchtest du, dass sie liebt?

          who want   you that she loves

          'Who do you want her to love?'

---

[9] Note, however, that *wh*-copying differs from complementiser agreement in Irish (and other forms of *wh*-agreement in other languages) in that the use of the copy strategy is optional in German.

[10] Again, individual speakers' sensitivity to such island effects appears to vary. Simpson (2000: 163n.12), for example, cites the following example involving a volitional predicate as fully grammatical:

  (i) *Wen* willst du, *wen* Jakob besticht?
      whom want you whom J.   bribes
      'Who do you want Jakob to bribe?'

[11] Following Reis (2000: 395), *wh*-in situ phrases are also excluded from the matrix clause of the copy construction:

  (i) *Wen hat Peter *wann* gesagt, *wen* er besuchen will?
      whom has P.   when said   whom he visit   will

b. **Wen* möchtest du, *wen* sie liebt? (Volitional Island)

who want you who she loves

(17) a. *Wen* bewies sie, dass Fritz liebt?

who proved she that F. loves

'Who did she prove that Fritz loves?'

b. **Wen* bewies sie, *wen* Fritz liebt? (Factive Island)

who proved she who F. loves

What is more, the *wh*-agreement analysis is unable to account for the fact that intermediate copies and overt complementisers can co-occur in dialects of German in which the 'doubly-filled COMP filter' does not hold (compare also Bayer 1996: 248n.63):

(18) *Wer* glaubst du, *wer* dass du bist?

who think you who that you are

'Who do you think you are?'

(Fanselow & Mahajan 2000: 222)

Last but not least, given that complementisers are heads, the *wh*-agreement analysis cannot explain why some speakers also permit copies of full *wh*-phrases to appear at intermediate positions (compare Höhle 2000: 258n.8). Taken together, these observations indicate that an analysis of the copy construction in terms of *wh*-agreement is not in fact tenable.

If intermediate copies are not simply complementisers in disguise, then it would appear that they must be either specifiers or adjuncts of non-interrogative C. Nunes (2000) argues for the latter view, in accordance with his more general claim that PF-copies of moved items are adjoined to functional heads rather than substituted in specifier positions. The adjoined *wh*-copy and its host will then be subject to a process of morphological reanalysis, which effectively turns the two into a single word. On the assumption that the LCA is blind to the internal structure of words (compare Chomsky 1995: 337), the LCA will then not be able to 'see' the copy, thus eliminating the indeterminacy problem noted earlier. Adjunction of an intermediate *wh*-item to embedded C is illustrated in the (simplified) representation in (19).

(19)   [$_{CP2}$ *Wovon$_i$* [$_{C'}$ glaubst du [$_{CP1}$ [$_C$ *wovon$_i$* [$_C$ Ø]] sie *t$_i$* träumt]]]

         of.what   believe you   of.what   she   dreams

Conceivably, head adjunction is motivated by the presence of some 'affixal' feature on a null (or phonetically reduced, as in the Frisian example [7b] above) C head. Notice that besides accounting for the fact that *wh*-copying is normally restricted to single morphophonological words, the head-adjunction analysis also allows for the possibility that a copy and an overt complementiser co-occur, as is the case in (18) above. A potential problem with this proposal though for the construction under consideration is that on the assumption that excorporation is barred by a constraint against affix stranding (Lasnik 1981), it should be impossible for a head-adjoined *wh*-copy to undergo any further movement. Thus a safer assumption might be that both intermediate and ultimate steps of *wh*-movement target CP-specifier positions, and that intermediate copies may subsequently undergo PF-cliticisation to embedded C (compare also Fanselow & Mahajan 2000: 221).

(20)   [$_{CP2}$ *Wovon$_i$* [$_{C'}$ glaubst du [$_{CP1}$ *wovon$_i$* [$_C$ Ø] sie *t$_i$* träumt]]]

         of.what   believe you   of.what   she   dreams

While intermediate *wh*-copies may or may not be visible to the LCA, there is some evidence from scopal data that they 'count' for semantic interpretation. Pafel (2000: 340), for instance, observes that *wh*-copying sentences such as (21a) below definitely have a pair-list reading, with the *wh*-pronoun *wo* 'where' having narrow scope with respect to the universal quantifier *jeder* 'everyone'. He is less certain, however, as to the availability of a wide scope reading of *wo* (an intuition that I share - but cf. von Stechow 2000: 467 for a different view). The ambiguity of the corresponding long-distance *wh*-question (21b), on the other hand, seems uncontroversial.

(21) a.   *Wo*   glaubt   jeder,   *wo*   die besten Weine wachsen?
             where believes everyone where the best   wines grow

      b.   *Wo*   glaubt   jeder,   dass die besten Weine wachsen?
             where believes everyone that   the best   wines grow
             Both: 'Where does everyone think that the best wines grow?'

By way of accounting for the above contrast, Pafel (2000: 348) proposes a constraint on relative scope determination to the effect that for a *wh*-expression to be able to outscope another quantifier, both the head of the *wh*-chain and the intermediate copy must be able to take scope over that quantifier.

Another interpretive difference between long-distance questions and the copy construction has been pointed out by Reis (2000). She observes that whereas long-distance *dass* questions such as (22a) below that contain an inconsistent proposition are ambiguous between an inconsistent and a consistent reading, the corresponding *wh*-copying structure (22b) patterns with simple *wh*-interrogatives such as (22c) in that it allows for an inconsistent reading only (the symbol # marks inconsistence).[12]

(22) a.   *Wo*   glaubt   sie, dass Fox populärer   ist als er ist?
           where believes she   that F. more.popular is than he is?
           'Where does she believe that Fox is more popular than he is?'

      b.   *#Wo*   glaubt   sie, *wo*   Fox populärer   ist als   er ist?
            where believes she   where F. more.popular is than he is?

      c.   *#Wo*   is Fox populärer   als   er ist?
            where is F.   more.popular than he is?

The contrast between (22a) and (22b) would thus seem to support Pafel's intuition that the lower rather than the higher *wh*-copy matters for certain interpretive purposes.

To the extent that the interpretive differences between long-distance *wh*-interrogatives and the copy construction noted above are real, they suggest that *wh*-copies are more than just PF-realisations of intermediate traces. If, however, intermediate *wh*-copies are also included in LF

---

[12] The consistent reading requires that the inconsistent object of attitude be attributable to two different sources capable of believing, and can be paraphrased roughly as follows: "For which place X, in her belief worlds is Fox more popular at X than Fox is popular at X in the real world" (compare von Stechow 2000: 468).

representations, then how come the use of the *wh*-copying strategy does not lead the derivation to crash at this level, or at least give rise to an interpretive dilemma? The apparently paradoxical situation that in complex *wh*-interrogatives, the *wh*-pronoun can be semantically interpreted at a non-operator position can plausibly be accounted for by certain special properties of *wh*-pronouns in languages like German. Note that from a semantic point of view, *wh*-expressions consist of two parts, a *wh*- or operator part, and an indefinite or 'core' part (Katz & Postal 1964, Bayer 1996, among others).

(23)   *wer, wo, warum....*   --> $[Q_{wh}]$ + 'something'

Based on the observation that in German, *wh*-pronouns can also be used as non-interrogative indefinites, some authors have suggested that under certain conditions, these two parts can be spelled out independently (compare Brandner 2000, Cheng 2000). If this is correct, then the copy strategy might provide a means for selectively spelling out (or rather, selectively deleting) the two parts of a *wh*-expression, along the lines illustrated in (24) below.[13]

(24) $[_{CP2}$ [Q_{wh} ~~something~~] ... $[_{CP1}$ [~~Q_{wh}~~ SOMETHING] ...]]

Individual languages vary as to whether or not they allow for *wh*-expressions to be split up into their constitutive parts. As for German, the possibility of *was-für* split and other types of separation construction (see e.g. Pesetsky 2000: 67-69) provides independent evidence that the language does indeed permit this option.[14]

---

[13] See also Hiemstra (1986) for a pre-minimalist analysis of *wh*-copying in Frisian in terms of selective feature movement.

[14] As Cheng (2000: 96) notes, however, relating the availability of partial *wh*-movement to morphological properties of *wh*-pronouns in a given language (i.e., whether or not they can also be used as indefinites) seems problematic in view of the observation that Frisian, for example, appears to lack the indefinite use of *wh*-pronouns. I therefore suggest that the availability of *wh*-separation constructions in a language might provide a more reliable clue as to whether or not that language also permits *wh*-copying. Clearly, the validity of this generalisation needs to be tested by further cross-linguistic investigation, though.

(25)   a.   *Was für Bücher* hast du gekauft?

                what for books   have you bought

        b.   *Was* hast du *für Bücher* gekauft?

                what have you for books   bought

                Both: 'What kind of books did you buy?'

Notice that if the above suggestion is along the right lines, it provides a natural account for Pafel's (2000) observation that the presence of a copy in the lower clause appears to prevent a moved *wh*-expression from outscoping a quantifier in the matrix clause. As the two *wh*-copies are part of a single but discontinuous *wh*-expression, for the *wh*-expression to take scope over another quantifier, both of its parts must do so.[15]

Interestingly, assuming that *wh*-copying involves the separation of the *wh*-part from its (indefinite) restriction might also help account for the island effects notes earlier. Specifically, the

---

[15] Additional evidence for selective spellout comes from long *wh*-movement in Afrikaans, where a preposition can optionally be stranded at an intermediate (Spec,CP) position, as in (iii) below (du Plessis 1977: 724).

(i)   *Waarvoor* dink julle werk ons?
       wherefore think you work we

(ii)   *Waar* dink julle werk ons *voor*?
       where think you work we for

(iii) *Waar* dink   julle *voor* werk ons?
       where think you   for   work we

Note that the use of *voor* instead of *vir* 'for' in (iii) indicates that the stranded preposition is part of the compound *wh*-pronoun *waarvoor* rather than a free preposition as in PPs like *vir wat* 'for what'. Contrary to bound *voor*, the morphologically independent preposition *vir* cannot in fact be stranded (cf. ibid.).

(iv)   *Waar/wat* werk ons nou eintlik *voor*?
       where/what work we now actually for

(v)   **Waar/wat* werk ons nou eintlik *vir*?
       where/what work we now actually for

Given that the stranded prepositional element in (iii) is a bound morpheme, an analysis in terms of base adjunction of *voor* to some projection of embedded C (along the lines suggested by Radford 2001 for floating quantifiers in Irish) does not seem a feasible alternative.

possibility of quantifiers or other operators intervening between two copies may be ruled out by the following general condition barring operator 'intervention effects' that has been proposed by Pesetsky (2000) (compare also Höhle 2000: 262f., and references cited there):

(26)   *Intervention Effect:*
        A semantic restriction on a quantifier (including *wh*) may not be separated from
        that quantifier by a scope-bearing element.

<div align="right">Pesetsky (2000: 67)</div>

Under this view, the negative island effect illustrated by (15b) above, for example, would result from a violation of condition (26) - the same condition that renders bad the (b) examples in (27)-(29) below.

(27)   a.   *Wen*   hat Hans   *alles* gesehen?
                whom has H.   all   seen
                'Who all did Hans see?'

        b.   ??*Wen*   hat niemand *alles* gesehen?
                  whom has no-one   all   seen

<div align="right">(Pesetsky 2000: 68f.)</div>

(28)   a.   *Was* hat Gretchen heute *Schönes* gemacht?
                what has G.   today nice   done
                'What nice thing did Gretchen do today?'

        b.   ??*Was* hat niemand heute *Schönes* gemacht?
                  what has no-one   today nice   done

<div align="right">(ibid.)</div>

(29)   a.   *Was* hat Karla *für Bücher* gelesen?
                what has K.   for books   read

b.   **Was*   hat niemand *für Bücher* gelesen?

         what has nobody   for books   read

<div align="right">(Höhle 2000: 263)</div>

In sum, the above observations indicate that intermediate copies of a long-distance extracted *wh*-expression do indeed represent spelled-out traces of successive-cyclic movement, and as such are most likely located in intermediate CP-specifier positions. What is more, there is some indication that the intermediate rather than the highest copy of a long-distance extracted *wh*-phrase matters for relative scope determination.

2.3   *What permits wh-copying?*

The fact that *wh*-copying is restricted to (a subset of) declarative-taking verbs indicates that for the copy strategy to be available, the complement clause *must* be headed by non-interrogative C. Setting aside, for the moment, the more general question of how *wh*-movement to the specifiers of intermediate non-interrogative heads is formally triggered (an issue that is discussed in section 3 below), another question that has not yet been addressed is what grammatical property or properties of *wh*-copying languages should permit the *spelling-out* of multiple copies of a moved *wh*-item.

McDaniel et al. (1995) propose a unified analysis of medial *wh* in English child language and *wh*-copying in languages like German and Romani in terms of underspecification. Specifically, they argue that grammars that permit *wh*-copying lack the [±pred(*icate*)] feature which according to Rizzi's (1990) feature system distinguishes relative clauses from declaratives and questions in the adult language. The spelling-out of intermediate copies is permitted in this case because the specifier of a C-head that is not specified for the feature [pred] at all is capable, in principle, of hosting a *wh*-expression:

> Grammars that do not contain the [pred] feature do not make a structural distinction between a relative clause and the lower clause of a long-distance question; whatever may appear in the Spec or C of one may appear in the Spec or C of the other. (McDaniel et al. 1995: 736)

Note that McDaniel et al.'s account diverges from Rizzi's (1990) original feature system in that it assumes that both relative clauses and embedded declaratives are headed by a [-wh] C , and that a *wh*-phrase may appear in the specifier of non-interrogative C if the *wh*-phrase is A'-bound (McDaniel et al. 1995: 734).[16] Among the objections that can be raised against McDaniel et al.'s underspecification analysis are the following, though. First, note that it seems intuitively rather implausible to claim that in languages like German or Romani, no distinction is made between relative clauses (which are semantically construed as predicates) and declarative complement clauses (which are argumental). Secondly, it is unclear to what extent a CP that is not clearly specified as declarative is capable of meeting the selectional requirements of the (subset of) declarative-taking verbs that permit *wh*-copying. If anything, the underspecification analysis would seem to predict that *wh*-copying should be *less* restricted than ordinary long-distance extraction - whereas in fact the opposite seems to be true (cf. the various types of 'island effect' noted earlier). The underspecification hypothesis moreover makes the incorrect prediction that German relative clauses can optionally be headed by the declarative complementiser *dass* 'that':

(30) a. Ich kenne den Mann, der   hier wohnt.

       I   know the man   who here lives

       'I know the man who lives here.'

    b. *Ich kenne den Mann, (der)   dass hier wohnt.

       I   know   the man   (who) that   here lives

Third, as the authors point out themselves, their assumption that children's initial grammars lack the [pred] feature is potentially problematic from a learnability perspective: If grammars that allow *wh*-copying are 'smaller' than grammars that do not, then what property or properties of the input tells (e.g.) English-speaking children that their target grammar *disallows wh*-copying? Finally, note that McDaniel et al.'s underspecification account is difficult to reconcile with the minimalist assumption that overt movement applies only as a last resort, to save a derivation that would otherwise be doomed to crash.

---

[16] Note, however, that in only free relatives are normally introduced by a *wh*-pronoun in German (whereas headed ones are introduced by a *d*-word).

What property of *wh*-raising languages like German might require that intermediate copies be phonetically realised under certain conditions? Recall that intermediate copies *must* be used in the absence of an overt complementiser, and that in standard varieties of the language, intermediate copies and complementisers are mutually exclusive. This indicates that German is subject to a requirement to the effect that *either* the head *or* the specifier of an embedded CP must contain phonetic material (compare also Fanselow & Mahajan 2000: 221). In contrast, no such visibility requirement appears to hold for embedded CP in English (cf. the grammaticality of [31b] below):

(31)   a.   *Wen glaubst du, sie liebt?

           who believe you she loves

       b.   Who do you think she loves?

In German, then, there are (at least) two alternative ways of rendering the embedded CP in long-distance questions PF-visible: by merging an overt complementiser in C (cf. [32a]), or, if the complementiser is phonetically null, by pronouncing the intermediate *wh*-copy (cf. [32b]).[17]

(32)   a.   Wen$_i$ glaubst du   [$_{CP}$ t$_i$'   [$_C$ dass]   [$_{TP}$ sie t$_i$ liebt]] ?
       b.   Wen$_i$ glaubst du   [$_{CP}$ wen$_i$ [$_C$ Ø]   [$_{TP}$ sie t$_i$ liebt]] ?

If Fanselow & Mahajan (2000: 221) are correct in that intermediate copies undergo obligatory PF-cliticisation onto null C (in standard varieties of German), then the relevant generalisation

---

[17] Moving the finite verb to C might provide a third way of satisfying the requirement that embedded declarative CP must be overtly marked:

(i)   Wovon glaubst du   *träumt$_i$* sie   *t$_i$* ?
      of.what believe you dreams   she

As it is not entirely clear though whether what appears to be extraction from V2 clauses involves genuine embedding, or parenthesis (as has been argued e.g. by Reis 2000), this possibility will not be discussed any further here.

might be that null declarative C in German - but not, for example, in English - is affixal in that it must be associated with phonetic material.[18]

2.4   Summary

Having shown that an analysis of *wh*-copying in terms of complementiser agreement is empirically inadequate, I concluded that intermediate *wh*-copies are best analysed as spelled-out traces of successive-cyclic movement. The *wh*-copying phenomenon thus provides strong evidence against the SSH. At the same time, though, the observation that the presence of 'active' *wh*-copies in embedded clauses does not cause the derivation to crash would seem to support a convergence-based view of phases, along the lines proposed by Atkinson (2000) and Radford (2001), and contra Chomsky (1998). That is, if the spelling-out of a [wh] feature is contingent upon feature-checking or 'valuation', then the presence of overt *wh*-copies in embedded clauses indicates that at the point at which these copies underwent spellout, the *wh*-item's uninterpretable [wh] feature had already been checked (or 'valued') through Q-agreement with the matrix C.

I further suggested that rather than creating semantically redundant PF-copies of a long-distance extracted *wh*-expression, *wh*-copying might offer a way of selectively spelling out the indefinite or core part of the *wh*-expression at a lower position, whereas the higher copy primarily serves the purpose of indicating that the main clause is a *wh*-interrogative. Under this view, some otherwise mysterious restrictions on *wh*-copying can be argued to follow from a constraint against separating a quantifier from its restriction which according to Pesetsky (2000) holds universally. The fact that in standard varieties of German, *wh*-copying is restricted to single morphophonological words can be attributed to the affixal character of embedded C in these varieties, which require the copy to be a potential PF-clitic.

---

[18] Alternatively, it may be that full *wh*-phrases are excluded from the copy construction for prosodic reasons, which would be in line with the observation that the specifier position of embedded declarative C in German appears to be unable to host any kind of focused constituent (compare Simpson 2000, and example [33] in the main text). Note that in contrast to full *wh*-phrases, pronominal *wh*-expressions can remain unstressed, thus allowing for the embedded clause in the copy construction to retain the prosodic contour of the corresponding *dass* clause.

3.   What triggers intermediate movement steps?

3.1   Optional P-features

While there is strong empirical evidence in favour of successive-cyclic *wh*-movement, the theoretical motivation for intermediate steps in movement is far less clear. As regards *wh*-copying in German, the assumption that embedded CP must be 'visible' may account for the possibility of spelling out an intermediate *wh*-trace, but does not answer the more fundamental question of what should motivate *wh*-raising to the specifier of a [-Q] head in the first place. Given that according to the Last Resort condition, movement can only apply if it results in feature checking, several authors have proposed that intermediate movement steps are triggered by pseudo-interrogative or other 'peripheral' (*force, focus,* or similar) features in intervening phase heads. Collins (1997), for example, has argued that intermediate movement steps are triggered by uninterpretable, non-interrogative [wh] features (or [Q] features, in Chomsky's 1998 system) in intermediate C-heads. On the assumption that the spelling-out of a *wh*-feature is a reflex of Q-agreement, the hypothesis that intermediate movement steps are triggered in essentially the same way as is the ultimate step provides a natural explanation for the fact that in languages that exhibit overt reflexes of successive-cyclic *wh*-movement such as complementiser agreement or *wh*-copying, these reflexes usually involve morphological *wh*-marking (compare also McCloskey 2000a: 8). It remains unclear, however, how such features should ever come to be associated with *non*-interrogative heads - or, conversely, why their presence in intermediate C heads does not violate the selectional requirements of the matrix verb.[19] What is more, note that we would normally expect a *wh*-expression that has entered into a Q-agreement relation with an intermediate C head to become inactive, i.e. unavailable for further movement. As this is clearly not the case (at least not in *wh*-raising languages), though, it appears that we are dealing with [Q] features that are 'strong' enough to act as probes for agreement, but at the same time, too 'weak' to check the uninterpretable [wh] feature of the goal.

A more feasible alternative to Collins' approach to successive-cyclic *wh*-movement is the idea that intermediate steps are triggered by features other than those involved in checking a *wh*-expression's uninterpretable *wh*-feature. Sabel (2000), for example, has suggested that a [focus] feature may be responsible for triggering local *wh*-movement to the specifier of non-interrogative

---

[19]   McCloskey (2000a: 31) points out, however, that the (temporary) presence of pseudo-*wh* features in intermediate C does not conflict with selectional requirements of the matrix verb because these features will be eliminated *before* the matrix clause is constructed.

heads (a similar suggestion has been made by Cole & Hermon 2000 for Malay). If this feature is 'strong', as he claims is the case in German, then partial movement (i.e., the spelling-out of *wh*-expressions in the specifier of non-interrogative C) is possible. If it is 'weak', as in English, partial movement is unavailable.[20] Note, however, that the assumption that embedded declarative C in German *generally* contains an uninterpretable [focus] feature is problematic because it makes the incorrect prediction that in non-interrogative contexts, other constituents should also undergo focus-movement to local CP-specifier positions. As has been pointed out by Simpson (2000: 176f.), however, nothing except *wh*-expressions ever seems to appear in embedded (Spec,CP) position in German:

(33) a.   *Er glaubt, Susi (dass) das Geld   gestohlen hat.

    he believes S.   (that) the money stolen   has

    Intended: 'He thinks that Susi has stolen the money.'

  b.   *Sie denkt, den Wein (dass) er getrunken hat.

    she thinks the wine (that) he drunk   has

    Intended: 'She thinks that he has drunk the wine.'

In short, even in a language in which the [focus] feature in C is supposed to be strong in the sense of being able to trigger overt substitution in its specifier, there is no independent evidence that embedded (Spec,CP) is a focus position.

Given that in the absence of a potential checker, the presence of uninterpretable peripheral features in intermediate phase heads would cause the derivation to crash, Chomsky (1998) suggests that P-features are only *optionally* added to C or *v* heads.

(34)   The head H of a phase PH may be assigned an EPP- and P-feature.

(Chomsky 1998: 23)

---

[20] In the framework outlined by Chomsky's (1998, and later), the strength metaphor is replaced by the presence versus absence of EPP features in agreement-triggering heads.

Together with a movement-triggering EPP feature, the presence of uninterpretable P-features in C and *v* ensure that a *wh*-item that is active because it (still) contains unchecked features is drawn to the edge of each intermediate phase, and thus remains accessible to further computation. There are, however, reasons for being sceptical about a solution to the Triggering Problem in terms of optional P-features as well. For one thing, peripheral features - which must be part of the featural make-up of both intermediate phase heads and a 'matching' *wh*-expression - do not make any obvious contribution to interpretation, at least not to the interpretation of the clause within which they are checked (compare McCloskey 2000a: 6). If anything, as McCloskey points out, the presence of such features in intermediate phase heads renders the mapping from syntax to semantics unnecessarily complicated. What is more, creating a potentially very large number of ultimately redundant *wh*-copies does not appear to be very much in line with minimalist views on economy, and would also seem to put unnecessary extra strain on working memory (compare Radford 2001: 37).[21] Last but not least, the hypothesis that movement-triggering features are present only when needed (that is, to trigger intermediate movement steps) would appear to describe, rather than derive, successive-cylic movement.[22] To conclude, in view of the fact that in there is no independent motivation for the presence of uninterpretable peripheral features in intermediate phase heads in long-distance questions, the idea that 'pseudo'-interrogative or other P-features can optionally be added to C or *v* before a phase is completed fails to offer a satisfactory solution to the Triggering Problem (3).

3.2   Phase Balance

Given that the presence of pseudo-interrogative or other peripheral features in intermediate phase heads is difficult to justify empirically, let us consider the alternative possibility that intermediate steps in movement are *not* feature-driven. A suggestion to this effect has been made by Heck & Müller (2000) as part of an optimality-theoretic approach to long-distance *wh*-raising. They propose that rather than being triggered by uninterpretable P-features of local C heads, intermediate movement steps serve to satisfy a condition dubbed *Phase Balance* (PB).

---

[21]   Note, however, that the existence of intermediate movement steps or copies poses a problem only if a global economy measure is applied, but not necessarily at a local level.

[22]   On the assumption that P-features are drawn from the lexicon, it is also not clear to me to what extent (34) is compatible with Chomsky's (1998) views on lexical selection.

(35)   *Phase Balance*:
         Phases must be balanced: If P is a phase candidate, then for every feature F in the
         numeration there must be a distinct potentially available checker for F.

<div align="right">(Heck & Müller 2000: 104)</div>

On the assumption that material on the left edge of a phase will remain accessible to further computation after that phase has undergone spellout, PB forces a *wh*-expression capable of checking uninterpretable features of some higher head to move to the edge of each current phase. As in Heck & Müller's model the PB constraint on phases is ranked higher than the Last Resort condition, derivations involving movement steps that are not feature-driven may still converge. The PB account is superior to feature-driven accounts in that it renders unnecessary the awkward step of adding uninterpretable, movement-triggering features to phasal heads and *wh*-expressions on an *ad hoc* basis.

Note, however, that the assumption of two different mechanisms capable of driving movement is not unproblematic, either. As McCloskey (2000a) notes, if the spelling-out of a [wh] feature is the result of feature-checking, then the PB approach has difficulty accounting for the fact that in languages that exhibit morphosyntactic reflexes of successive-cyclic *wh*-movement, the [wh] feature is typically spelled out at intermediate positions as well. What is more, note that the PB presupposes that at any stage during the derivation, the computation has access to the *complete* numeration (Heck & Müller 2000: 104). The assumption of this type of 'look-ahead', however, is incompatible with Chomsky's (1998) claim that each phase is constructed from a lexical subarray containing a single C or *v* only, and difficult to reconcile with recent attempts to eliminate all acyclic devices, including any form of look-ahead, from the theory of grammar (compare, among others, Collins 1997, Frampton & Gutmann 1999).

In short, while the assumption that intermediate movement steps are not feature-driven eliminates most of the problems faced by feature-driven accounts, the PB approach to successive-cyclic movement does, in turn, raise concerns about multiple feature-spellout and look-ahead. In the next section, I suggest that these two problems become irrelevant once successive-cyclic *wh*-movement is viewed from a processing perspective rather than from the point of view of bottom-up, right-to-left structure-building.

3.3   *A processing perspective*

The long-neglected questions of how and to what extent the form of the grammar may be shaped by processibility demands and general cognitive limitations have recently become a focus of interest also among generative linguists (see, in particular, Hawkins 1994, as well as the

introductory sections of Chomsky 2001). Comparing and integrating insights and findings from linguistic theory, language typology, and language processing studies led Hawkins (1994: 409) to conclude that "many fundamental and abstract structural properties of grammars can be explained by simple considerations of processing ease", such as the observation that in head-initial languages 'heavy' constituents tend to be postposed, whereas head-final languages prefer preposing. One property of the cognitive system that is known to constrain our ability to produce and comprehend sentences is our limited working memory capacity (Just & Carpenter 1992, King & Just 1991, Gibson 1998, to name but a few). Sentence processing is generally assumed to proceed incrementally from left to right, with each new incoming word or phrase being integrated into the current partial phrase marker as soon as possible (Frazier 1987a, Phillips 1996, among others). Given that working memory resources are finite, it is widely assumed in the psycholinguistic literature that rather than having continuous access to left context, the parser only has a limited 'viewing window' (see, for instance, Berwick & Weinberg 1984, Frazier & Fodor 1978, or Frazier 1985). Regarding the processing of syntactic dependencies, there is evidence that if the parser encounters a displaced element (or 'filler') earlier on in the sentence, it will postulate an associated 'gap' as early as grammatically possible. This parsing principle is known in the psycholinguistic literature as the *Active Filler Strategy* (Frazier 1987b, Frazier & Clifton 1989), or *Minimal Chain Principle* (De Vincenzi 1991). A filler must be retained in working memory until the gap position at the foot of the chain has been located, where it will then be associated with its subcategoriser or other licenser.[23]

Reconsider now complex *wh*-interrogatives. During left-to-right syntactic processing, a *wh*-item appearing in sentence-initial position will immediately be identified as an operator (as well as signalling to the parser that it is dealing with a *wh*-interrogative), so that *wh*-feature checking can be considered 'done'. At the same time, however, the *wh*-item is uninterpretable in that it cannot, at this stage during the parse, be associated with any specific thematic or participant role.[24] In order to construct a representation for a *wh*-question that is convergent in that in

---

[23] These assumptions are supported by results from brain imaging studies involving the measurement of event-related brain potentials (ERPs) during the processing to syntactic dependencies. Kluender & Kutas (1993a,b), for example, found what appear to be electro-physiological correlates of the memory cost incurred by keeping a filler active in working memory (so-called 'LAN-effects'), and Kaan et al. (2000) have claimed that the relative difficulty of integrating a displaced constituent with its subcategoriser may be reflected in the relative size of an ERP component known as the 'P600'.

[24] To the extent that selection also involves a form of feature-checking (Svenonius 1994), the *wh*-item may carry unchecked selectional features as well.

addition to ensuring that all uninterpretable features are checked and eliminated, it also satifies the θ-criterion and selectional requirements, it is vital that the *wh*-filler is ultimately linked to its subcategoriser (or other licensing head). On the assumption that thematic-role assignment takes place at the point of first merger (Chomsky 1995: 313), the *wh*-filler in long-distance interrogatives must remain available for reconstruction at its base position while the parser continues to construct a potentially fairly large number of new phases. Given that WM resources are limited though, chances are that if movement has taken place across several phase boundaries, the filler may fade from working memory before it can be reinstated at its original position (compare Fazier & Clifton 1989, Gibson 1998). From this perspective, intermediate steps in movement serve the function of 'refreshing' the filler at each processing cycle (that is, at the beginning of the construction of each new phase) so as to ensure that it is kept active, and to facilitate its retrieval from working memory once the foot of the chain is encountered. Or, as Bayer (1992: 25) puts it, an intermediate copy can be seen as "a device that *restarts* the gap-search". As we saw earlier, some languages even permit the spelling-out of intermediate copies under certain conditions.

Psycholinguistic evidence that intermediate *wh*-traces form part of the mental representations of sentences that are constructed during processing comes from reading-time studies by Frazier & Clifton (1989) and Gibson & Warren (2000). In a segment-by-segment self-paced reading experiment, Gibson & Warren, for example, found that *wh*-dependencies involving several short steps (as in [36a] below) were easier to establish during online processing than were 'single-step' dependencies of similar overall length, as in (36b).

(36)　　a.　　The manager *who* the consultant claimed *t*' that the new proposal had

　　　　　　　pleased *t* will hire five workers tomorrow.

　　　　b.　　The manager *who* the consultant's claim about the new proposal had

　　　　　　　pleased *t* will hire five workers tomorrow.

Specifically, reading times would increase at the segment containing the subcategoriser for the moved *wh*-expression (e.g., at the segment containing the verb *pleased*) in the complex noun

phrase condition (36b) relative to the embedded-CP condition (36a).[25] On the assumption that higher reading times reflect an increase in processing difficulty, the authors interpret their findings as evidence that intermediate linguistic structure is utilised during the online processing of *wh*-dependencies.

Recall that given the problems with feature-driven accounts of successive-cyclic movement noted earlier, and our conclusion that spellout is based on convergence, there appears to be no plausible formal motivation for including intermediate traces in the phrase marker under construction. What is more, allowing for spellout to be delayed potentially indefinitely seems hardly compatible with Chomsky's central claim that the derivation proceeds 'by phase'. Observe that from a processing perspective, however, the idea that phases should constitute local computational units is intuitively very plausible, on the assumption that at any point during processing, the human parser has a limited viewing window only (Berwick & Weinberg 1984, Frazier & Fodor 1978, among others). From the point of view of language processing, postulating intermediate copies of moved items offers a way to overcome working memory limitations in that it allows for an uninterpretable filler to be kept active over distances exceeding the local computational work space, until it can be associated with its (thematic, or other) licenser. What I would like to suggest, therefore, is that the *Phase Balance* condition, which allows for movement to take place in the absence of local triggering features, is best understood as a constraint imposed by processing considerations, and should thus be reformulated along the following lines:

(37)   *Phase Balance (revised)*:
       For each locally uninterpretable element X, a copy of X must be postulated at the
       edge of each newly processed phase to ensure that X remains accessible to the parser
       for as long as necessary to ensure convergence.

By causing an active element to be 'refreshed' at the edge of the phase currently being constructed, condition (37) ensures that this element remains well within the viewing window of the syntactic processor until the minimal constituent containing the agreeing probe (from a bottom-up structure-building perspective) or the subcategorising head (from a left-to-right processing perspective) has been assembled.

---

[25]   The effect of the difference in subject-verb distance between the two conditions in (36) was controlled
       for independently.

Let us now reexamine the problems of multiple spell-out and look-ahead that arose from the *Phase Balance* condition in its original formulation as a constraint on structure-building in (35) above. In view of the fact that syntactic processing, including the postulation of intermediate copies, must be performed on the basis of PF-representations, indicating the presence of such copies through *wh*-agreement marking or *wh*-copying does not in fact seem such a bad idea.[26] Note further that from a processing perspective, children's use of medial *wh* in non-*wh* copying languages such as English can be accounted for by the fact that children's WM capacity is generally more limited than that of adults. Intermediate traces, then - whether spelled out or not - quite literally serve as memory aids, ultimately ensuring the interpretability of the overall sentence. From the point of view of the grammar, multiple spellout of the [wh] feature will not cause the derivation to crash as long as the highest member of the chain has entered (or will ultimately enter) into the required checking relation.

The second potential problem with Heck & Müller's (2000) version of PB noted above arose from the necessity for assuming look-ahead, in the shape of access to the complete numeration. Again, this problem simply disappers once we redefine PB as a constraint on language processing rather than on right-to-left phase construction. The revised Phase Balance condition (37) ensures that a syntactic item that is not fully interpretable within the phase currently being processed is carried forward into the next processing cycle, until it can be associated with its licenser. Although the ultimate movement step is feature-driven, there is no need for assuming any look-ahead on the part of the parser.

4.   *Concluding remarks*

I started out by drawing attention to some problems that arise from Chomsky's (1998) phase-based theory of structure-building for successive-cyclic *wh*-movement. A possible alternative to sucessive-cyclic movement suggested by Radford (2001) to the effect that long *wh*-movement proceeds in a single step was rejected on the grounds that it is unable to account for the *wh*-copying phenomenon that is found in languages like German or Romani, which was argued to provide strong evidence that long *wh*-raising does indeed involve a series of local steps. At the

---

[26] From a processing perspective, one might say that it is languages such as Modern Standard English that permit long-distance *wh*-raising *without* providing a means of *wh*-marking intermediate landing sites that should be considered 'odd'.

same time, however, the *wh*-copying phenomen, like other instances of multiple spellout of [wh] features, calls into question the idea that spellout takes place cyclically at the next higher phase level. In other words, we appear to be faced with the paradoxical situation that the assumption of cyclic spellout, while ensuring that long-distance A'-movement takes place in successive-cyclic fashion on the one hand, is actually incompatible with the presence of intermediate *wh*-copies that are still active at the point at which spellout applies. By way of offering a possible solution to this dilemma, I suggested that the notion of 'phase' - in the sense of a semantically and phonologically relatively independent unit - may best be regarded as being relevant primarily to language processing. The operation *Spellout*, on the other hand, is potentially non-local in that it applies only to constitents (or phases) that are convergent. Note that a convergence-based view of spellout presupposes that the derivation has continuous access to the lexicon, as has been argued, among others, by Collins (1997) and Frampton & Gutmann (1999). I do not, at present, see any major problems with the view that lexical access and syntactic structure-building should proceed in parallel. In fact, the idea of dispensing with numerations, or lexical subarrays, would seem to be advantageous from an economy point of view.

Given that it is difficult to identify a plausible *local* trigger for intermediate movement steps, I further proposed that intermediate movement steps are indirectly feature-driven only. That is, while long-distance *wh*-raising as such is formally triggered by the need to check uninterpretable features of both the *wh*-item and the interrogative C head of the matrix clause, non-ultimate movement steps are required essentially for processing reasons, to ensure that a displaced element is retained in working memory during the processing of multiple phases.[27] This idea is in line with Hawkins (1994) claim that grammars are to a large extent performance-driven, and has led to the formulation of the revised Phase Balance condition (37), which does not presuppose any form of look-ahead. Whether or not a given language permits the spelling-out of

---

[27] As Bob Borsley points out, assuming that successive-cyclic movement is motivated by extra-linguistic requirements does not eliminate the need for a formal mechanism to trigger it. One possible way of integrating indirectly feature-driven movement with formal theory would be to re-define EPP features as 'processing' features. Note that the status of EPP features has always been somewhat dubious, given that rather than contributing directly or indirectly (by helping eliminate uninterpretable features though agreement) to sentence interpretation, their presence merely signals to the computational system that movement should apply. Observe, however, that such a move would have far-reaching consequences for the theory of grammar, and would raise a host of new questions about the nature of other types of movement (notably, A-movement).

intermediate PF-copies, though, depends on language-specific properties such as the possibility of separating the operator part from the indefinite part of a *wh*-expression.

References

Atkinson, M. (2000) 'Uninterpretable feature deletion and phases', *Essex Research Reports in Linguistics* 34, 91-122.

Barss, A. (2001) 'Syntactic reconstruction effects', in M. Baltin and C. Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, 670-696.

Bayer, J. (1992) 'Wh-dependencies in syntax and logical form'. Ms., University of Düsseldorf.

Bayer, J. (1996) *Directionality and Logical Form*. Dordrecht: Kluwer.

Berwick, R., and A. Weinberg (1984) *The Grammatical Basis of Linguistic Performance*. Cambridge, MA: MIT Press

Brandner, E. (2000) 'Scope marking and clausal typing', in U. Lutz, G. Müller and A. von Stechow (eds.), 45-75.

Cheng, L. L.-S. (2000) 'Moving just the feature', in U. Lutz, G. Müller and A. von Stechow (eds.), 77-99.

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

Chomsky, N. (1998) Minimalist inquiries: the framework. *MIT Occasional Papers in Linguistics* 15. Cambridge, MA: MITWPL.

Chomsky, N. (1999) Derivation by phase. *MIT Working Papers in Linguistics* 18. Cambridge, MA: MITWPL.

Chomsky, N. (2001) 'Beyond explanatory adequacy'. Ms., MIT.

Clements, G., J. McCloskey, J. Maling and A. Zaenen (1983) 'String-vacuous rule application', *Linguistic Inquiry* 14, 1-17.

Cole, P. (1982) *Imbabura Quechua*. The Hague: North Holland Publishing Company.

Cole, P. and G. Hermon (2000) 'Partial wh-movement: Evidence from Malay', in U. Lutz, G. Müller, and A. von Stechow (eds.), 101-130.

Collins, C. (1997) *Local Economy*. Cambridge, MA: MIT Press.

Dayal, V. (1994) 'Scope marking as indirect wh-dependency', *Natural Language Semantics* 2, 137-170.

Dayal, V. (2000) 'Scope marking: Cross-linguistic variation in indirect dependency', in U. Lutz, G. Müller and A. von Stechow (eds.), 157-193.

De Villiers, J., T. Roeper, and A. Vainikka (1990) 'The acquisition of long-distance rules', in L. Frazier and J. de Villiers (eds.) *Language Processing and Language Acquisition*. Dordrecht: Kluwer, 257-297.

De Vincenzi, M. (1991) *Syntactic Parsing Strategies in Italian*. Dordrecht: Kluwer.

Fanselow, G., and A. Mahajan (2000) 'Towards a minimalist theory of wh-expletives, wh-copying, and successive cyclicity', in U. Lutz, G. Müller and A. von Stechow (eds.), 195-230.

Felser, C. (2001) 'Wh-expletives and secondary predication: German partial wh-movement reconsidered', *Journal of Germanic Linguistics* 13, 5-38.

Felser, C. and L. Rupp (2001) 'Expletives as arguments: Germanic existential sentences revisited', *Linguistische Berichte* 187, 289-324.

Fox, D. (2000) *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.

Frampton, J. and S. Gutmann (1999) 'Cyclic computation, a computationally efficient minimalist syntax', *Syntax* 2, 1-27.

Frazier, L. (1985) 'Syntactic complexity', in D. Dowty, L. Karttunen, and A. Zwicky (eds.) *Natural Language Parsing*. Cambridge: Cambridge University Press.

Frazier, L. (1987a) 'Sentence processing: A tutorial review', in M. Coltheart (ed.) *Attention and Performance XII.* Hillsdale, NJ: Erlbaum.

Frazier, L. (1987b) 'Syntactic processing: Evidence from Dutch', *Natural Language and Linguistic Theory* 5, 519-559.

Frazier, L. and C. Clifton (1989) 'Successive cyclicity in the grammar and in the parser', *Language and Cognitive Processes* 4, 93-126.

Frazier, L. and J. D. Fodor (1978) 'The sausage machine: A new two-stage parsing model', *Cognition* 6, 291-326.

Gibson, E. (1998) 'Linguistic complexity: Locality and syntactic dependencies', *Cognition* 68, 1-76.

Gibson, E. and T. Warren (2000) 'The psychological reality of intermediate linguistic structure in long-distance dependencies'. Ms., MIT.

Hawkins, J. (1994) *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Heck, F. and G. Müller (2000) 'Successive cyclicity, long-distance superiority, and local optimization', *Proceedings of WCCFL* 19, 101-114.

Hiemstra, I. (1986) 'Some aspects of *wh*-questions in Frisian', *North-Western European Language Evolution (NOWELE)* 8, 97-110.

Hinrichs, E. and T. Nakazawa (2001) 'The was-w construction in German: A case study in type coercion'. Ms. University of Tübingen.

Höhle, T. (2000) 'The *w...w* construction: Appositive or scope indicating?', in U. Lutz, G. Müller and A. von Stechow (eds.), 249-270.

Jackendoff, R. (1999) 'The representational structures of the language faculty and their interactions', in C. Brown and P. Hagoort (eds.) *The Neurocognition of Language.* Oxford: Oxford University Press, 37-77.

Just, M. and P. Carpenter (1992) 'A capacity theory of comprehension: New frontiers of evidence and arguments', *Psychological Review* 99, 122-149.

Kaan, E., A. Harris, E. Gibson and P. Holcomb (2000) 'The P600 as an index of syntactic integration difficulty' *Language and Cognitive Processes* 15, 159-201.

Katz, J. and P. Postal (1964) *An Integrated Theory of Linguistic Description.* Cambridge, MA: MIT Press.

Kayne, R. (1994) *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press.

King, J. and M. Just (1991) 'Individual differences in syntactic processing: The role of working memory', *Journal of Memory and Language* 30, 580-602.

Kluender, R. and M. Kutas (1993a) 'Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies', *Journal of Cognitive Neuroscience* 5, 196-214.

Kluender, R. and M. Kutas (1993b) 'Subjacency as a processing phenomenon', *Language and Cognitive Processes* 8, 573-633.

Lasnik, H. (1981) 'Restricting the theory of transformations: A case study', in N. Hornstein and D. Lightfoot (eds.) *Explanation in Linguistics*. London: Longman, 152-173.

Lutz, U., G. Müller, and A. von Stechow (eds.) (2000) *Wh-Scope Marking*. Amsterdam & Philadelphia: John Benjamins.

McCloskey, J. (2000a) 'Resumption, successive cyclicity, and the locality of operations'. Ms., to appear in S. Epstein and D. Seeley (eds.), *Prospects for Derivational Explanation*. Blackwell.

McCloskey, J. (2000b) 'Quantifier float and *wh*-movement in an Irish English', *Linguistic Inquiry* 31, 57-84.

McCloskey, J. (2000c) 'The prosody of quantifier stranding under *WH*-movement in West Ulster English'. Ms. University of Califormia, Santa Cruz.

McCloskey, J. (2001) 'The morphosyntax of wh-extraction in Irish', *Journal of Linguistics* 37, 67-100.

McDaniel, D. (1986) Conditions on wh-chains. PhD dissertation, City University of New York.

McDaniel, D. (1989) 'Partial and multiple *wh*-movement', *Natural Language and Linguistic Theory* 7, 565-604.

McDaniel, D., B. Chiu, and T. Maxfield (1995) 'Parameters for wh-movement types: Evidence from child English', *Natural Language and Linguistic Theory* 13, 709-753.

Nunes, J. (2000) 'Linearization of chains and phonetic realization of multiple copies'. Paper presented at the Workshop on Antisymmetry Theory, Cortona, 15-17 May 2000.

Pafel, J. (2000) 'Absolute and relative: On scope in German *wh*-sentences, was...w-constructions included', in U. Lutz, G. Müller and A. von Stechow (eds.), 333-358.

Pesetsky, D. (2000) *Phrasal Movement and Its Kin*. Cambridge, MA: MIT Press.

Phillips, C. (1996) Order and structure. MIT dissertation.

du Plessis, H. (1977) *'Wh* movement in Afrikaans', *Linguistic Inquiry* 8, 723-726.

Postal, P. (1972) 'On cyclic rules that are not successive cyclic', *Linguistic Inquiry* 3, 211-222.

Radford, A. (2001) 'Probing *Derivation by Phase*'. Ms., University of Essex.

Reis, M. (2000) 'On the parenthetical features of German was...w constructions and how to account for them', in U. Lutz, G. Müller and A. von Stechow (eds.), 359-407.

Rizzi, L. (1990) *Relativized Minimality.* Cambridge, MA: MIT Press.

Roeper, T. (1990) 'How the least effort concept applies to partial grammars: Copying in the acquisition of inversion and cyclic movement'. Ms., University of Massachusetts, Amherst.

Sabel, J. (2000) 'Partial *wh*-movement and the typology of *wh*-questions', in U. Lutz, G. Müller and A. von Stechow (eds.), 409-446.

Saddy, D. (1991) 'Wh scope mechanisms in Bahasa Indonesia', in L. Cheng and H. Demirdache (eds.) *MIT Working Papers in Linguistics* 15, 183-218. Cambridge, MA: MITWPL

Simpson, A. (2000) *Wh-Movement and the Theory of Feature-Checking.* Amsterdam & Philadelphia: John Benjamins.

Stechow, A. von (2000) 'Partial *wh*-movement, scope marking, and transparent Logical Form', in U. Lutz and G. Müller (eds.), 1-36.

Steedman, M. (2000) *The Syntactic Process.* Cambridge, MA: MIT Press.

Svenonius, P. (1994) 'C-selection as feature-checking', *Studia Linguistica* 48, 133-155.

Thornton, R. (1990) Adventures in long-distance moving: The acquisition of complex wh-questions. PhD dissertation, University of Connecticut.

Thornton, R. and S. Crain (1995) 'Successful cyclic movement', in T. Hoekstra and B. Schwartz (eds.) *Language Acquisition Studies in Generative Grammar.* Amsterdam: John Benjamins, 215-252.

Author's Address:

Department of Language & Linguistics
University of Essex
Colchester CO4 3SQ
United Kingdom

Email: *felsec@essex.ac.uk*