On the Treatment of Multiple-*Wh*-Interrogatives in Minimalist Grammars

Hans-Martin Gärtner and Jens Michaelis

1. Introduction

As shown by Michaelis (2001a,b), Stablerian *Minimalist Grammars (MGs)* (Stabler 1997b, 1998, 2008) belong among the class of *mildly context-sensitive grammar* formalisms as, e.g., characterized in Joshi et al. 1991. Crucially, mild context-sensitivity depends on what may—somewhat loosely—be referred to as *resource-sensitivity*, i.e., strictly limited combinability of syntactic objects. Within MGs, resource-sensitivity is implemented in terms of *feature checking*: each structure-building operation must consume (i.e. check and eliminate) features. In addition, the number of features of identical type accessible to an operation at a given stage of a derivation is strictly bounded. The core MG-constraint enforcing such an upper bound on available resources is the so-called *Shortest Move Condition (SMC)*, which—in its most rigid version—allows no more than one feature of a given type to be available for attraction by some category requiring dependency formation. Consider the pair of German *wh*-interrogatives in (1).

(1) a. *Wer vermisst wen?*
 'Who misses who(m)?'
 b. *Wen vermisst wer?*
 lit. 'Who(m) does who miss?'

This structural pattern follows naturally from the assumption that C$^\circ$ possesses an attracting feature `+wh` that requires a constituent with an attractable counterpart `-wh` to move to its specifier, i.e. Spec,CP. The features in question are instances of a *m(ove)-licensor* and a corresponding *m(ove)-licensee*, generally depicted in the way `+x` and `−x`, respectively. (1-a) results if `−wh` is located on *wer*, (1-b) is derived if *wen* possesses `−wh`. Accordingly, the SMC can be (semi-formally) stated as a constraint on the operation *move* as follows. (We will say that a tree τ *displays* feature f, if an instance of f starts the string of *formal features* of the head of τ.)

(2) *Shortest Move Condition (SMC)* [*move*] (semi-formal version)
The operation *move* is applicable to a tree τ iff τ displays m-licensor `+x`, and there is exactly one maximal subtree τ' of τ such that τ' displays m-licensee `−x`.

The SMC was originally designed to guarantee non-violability of *wh*-islands, as illustrated in (3).

(3) a. **Wen fragt sich Maria wer vermisst?*
lit. 'Who(m) does Mary wonder who misses?'
b. **Wer fragt sich Maria wen vermisst?*
lit. 'Who does Mary wonder who(m) misses?'

Clearly, at the stage at which the lower interrogative CP is built, both *wer* and *wen* would have to display `−wh`, one for satisfying `+wh` on the local C°, the other for (later) satisfying `+wh` on the matrix C°. But that is what the SMC rules out and the unacceptability of (3) is correctly predicted.

However, the SMC as stated above faces a number of challenges, two of which are immediately obvious from our discussion. First, classical superiority effects such as the one in (4) are not captured.

(4) **What does who prefer?*

(4) is derivable on a par with (1-b) when *what* displays `−wh` and *who* does not. Secondly, multiple *wh*-movement in languages such as Bulgarian is likewise banned. Consider (5) (Rudin, 1988, p. 449).

(5) *Koj kogo vižda?*
lit. 'Who who(m) sees?'

According to Rudin's analysis both *wh*-phrases have to be moved to Spec,CP *overtly*, which would seem to imply that both *koj* and *kogo* have to display `−wh` at the stage where C° is introduced into the structure.[1] This is ruled out by the SMC. The purpose of the present paper is to deal with the second challenge, i.e. to suggest an SMC-compatible MG-treatment of multiple-*wh*-interrogatives. As already indicated in earlier writings (Gärtner & Michaelis 2005, 2007), we would like to pursue the hypothesis put forward by Grewendorf (2001) and Sabel (2001) that multiple *wh*-movement involves *wh*-clustering. Accordingly, *koj* and *kogo* in (5) do not require individual applications of *wh*-movement but move to Spec,CP as a cluster. The details will be spelled out in Section 2. In Section 3, the system developed in Section 2 will be confronted with some basic constraints governing placement and reordering possibilities in various languages. Section 4 will sharpen the tools employed to deal with *wh*-clustering in MGs. In particular, the nature of overt vs. covert movement will be explored in more depth. Finally, Section 5 deals prospectively with areas for further study: (i) why—pace Grewendorf 2001 and Sabel 2001—*wh*-clustering should be generalized to all languages, and (ii) how MG-style *wh*-clustering can deal with the non-trivial empirical domain of *additional wh-effects*. A formalization of MGs with clustering is given in the appendix.

2. Multiple *Wh*-Movement as *Wh*-Clustering

In this section we will present our assumptions about *wh*-clustering and explore the immediate consequences for a typology of multiple-*wh*-languages. Consider (6), which should be thought of as a "didactic abstraction" from real Bulgarian data.[2]

(6) $[_{\text{CP}}$ [*who* [*what* [*to whom* $]_i$ $]_j$ $]_k$ $[_{\text{C}'}$ C° $[_{\text{IP}}$ t_k $[_{\text{I}'}$ I° $[_{\text{VP}}$ t_j $[_{\text{V}'}$ *gave* t_i]]]]]]

In order to deal with such cases in an SMC-compatible fashion we assume that in addition to *move* there is a structure building operation *cluster*. *cluster* is triggered by a *c(luster)-licensor* ▽wh on (the D°-head of) a *wh*-phrase attracting a *c(luster)-licensee* △wh on (the D°-head of) another. The structures in (7) illustrate the essential steps in deriving (6). (Labels > and < indicate that the heads of the thus labeled trees are found on the right and left branch, respectively.)

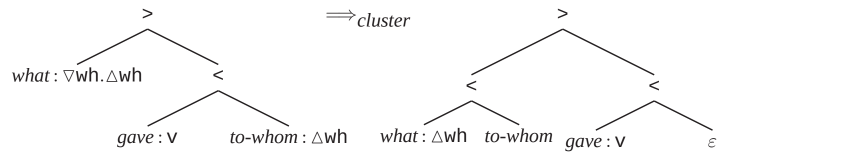
Step (7-a) illustrates a derivational stage at which *what* displays ▽wh and *to-whom* displays △wh. The operation *cluster* applies such that *to-whom* is right-attached to *what* and the features involved are checked and deleted. Traces are marked by ε. At stage (7-b), *who* displays ▽wh and the clustered constituent $[_<$ *what to-whom*]

[1] C° in (5) would have to possess `=i.+wh.+wh.c` as string of formal features: it selects an IP and attracts two *wh*-phrases into Spec,CP—one after the other—yielding a structure of category C(P).

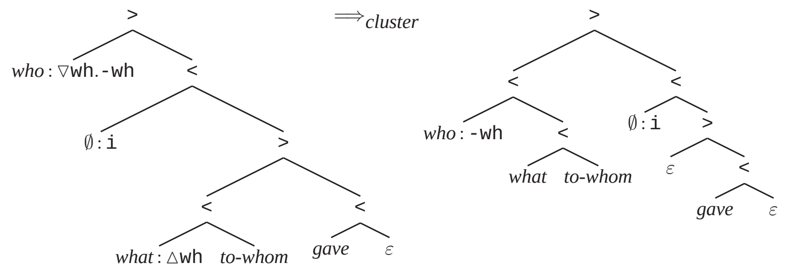
[2] The paper by Rudin (1988) still provides a good starting point for finding out about Bulgarian multiple-*wh*-interrogatives. Further references will be given as we go along.

displays △wh. Again, *cluster* applies and [$_<$ *what to-whom*] gets right-attached to *who*. Finally, at stage (7-c), C° displays +wh while the clustered constituent [$_<$ *who* [$_<$ *what to-whom*]] displays −wh. Now, standard *move* can apply and transport the *wh*-cluster to Spec,CP. The result is (6).

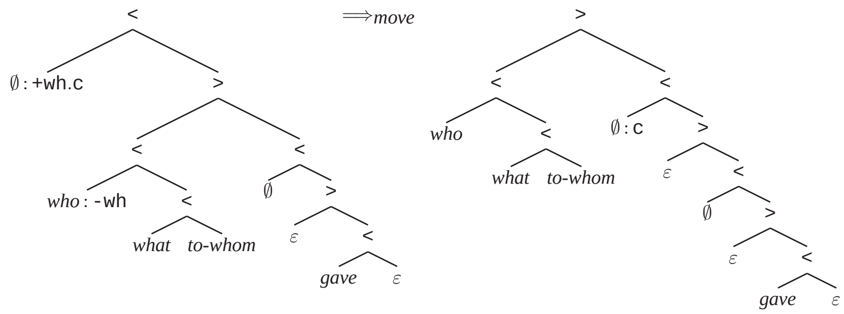
(7) a.



b.



c.



Importantly, at no stage in the derivation of (6) does more than one licensee (of the same type) have to be displayed. Thus, a maximally strictly resource-bounding SMC, covering the displacement operations *move* and *cluster*, can be assumed while at the same time the challenge of (Bulgarian) multiple-*wh*-interrogatives can be met. (8) states the SMC as applied to cluster formation.[3]

3 Allowing τ itself to display ▽wh will be required in cases like *which book about what* where one *wh*-phrase is contained inside another. A further potential application is the domain of head movement and cliticization as discussed by Sabel (2001, Section 2). We have chosen to leave the required modifications for further research.

Note also that in situations where τ displays +x and, at the same time, the highest specifier of τ displays ▽y, application of *cluster* has to precede application of *move* at least in all standard cases

(8) *Shortest Move Condition (SMC)* [*cluster*] (semi-formal version)
The operation *cluster* is applicable to a tree τ iff the highest specifier of τ displays c-licensor $\triangledown\mathtt{x}$, and there is exactly one maximal subtree τ' of τ such that τ' displays c-licensee $\triangle\mathtt{x}$.

Now, it is well-known that multiple-*wh*-interrogatives show a substantial amount of variation among the languages of the world. Let us therefore briefly go over some of the basic means of capturing that variation within our system.

To begin with, Japanese has been argued by Grewendorf (2001) and Sabel (2001) to be exactly like Bulgarian except that *wh*-movement and *wh*-clustering occur *covertly*, i.e. without affecting PF-linearization. Schematically, this can be illustrated by using a counterpart of traces which we will call *shadows* and abbreviate as s. These shadows are moved like standard constituents but they leave PF-visible material behind. Thus, the Japanese counterpart of (6) looks like (9).

(9) $[_{\mathrm{CP}}\ \mathrm{s}_k\ [_{\mathrm{C'}}\ [_{\mathrm{IP}}\ [\ \textit{who}\ \mathrm{s}_j\ ]_k\ [_{\mathrm{I'}}\ [_{\mathrm{VP}}\ [\ \textit{what}\ \mathrm{s}_i\ ]_j\ [_{\mathrm{V'}}\ [\ \textit{to whom}\ ]_i\ \textit{gave}\ ]\ ]\ \mathrm{I}^{\circ}\ ]\ ]\ \mathrm{C}^{\circ}\ ]\ ]$

Formally, the distinction will be captured in the way familiar from work by Chomsky (1995). The set of licensors will be partitioned into strong ones and weak ones (cf. Stabler 1997a,b, 1998), marked henceforth by presence vs. absence of underlining, respectively. Schematically, the difference between Bulgarian and Japanese responsible for the difference between (6) and (9) is determined by the following lexical specifications.

(10) Bulgarian
  a. $\emptyset$ :: $\mathtt{=i.+\underline{wh}.c}$
  b. *who* :: $\mathtt{d.\triangledown\underline{wh}.-wh}$
  c. *what* :: $\mathtt{d.\triangledown\underline{wh}.\triangle wh}$
  d. *to-whom* :: $\mathtt{d.\triangle wh}$

(11) Japanese
  a. $\emptyset$ :: $\mathtt{=i.+wh.c}$
  b. *who* :: $\mathtt{d.\triangledown wh.-wh}$
  c. *what* :: $\mathtt{d.\triangledown wh.\triangle wh}$
  d. *to-whom* :: $\mathtt{d.\triangle wh}$

Thus, while in Bulgarian all *wh*-related m- and c-licensors are strong, in Japanese all of them are weak.

---

where *move* leads to creation of an additional (outer) specifier. Otherwise $\triangledown\mathtt{y}$ would be ineliminable and the derivation could not "converge."

The obvious next question is whether there are mixed systems. Indeed, both Grewendorf (2001) and Sabel (2001) argue that German is like Bulgarian and Japanese, except that only one *wh*-phrase ends up in Spec,CP overtly. From our perspective this means that m-licensors are strong while c-licensors are weak. (12) and (13) show the German counterparts of (6)/(10) and (9)/(11), respectively.

(12) $[_{CP}\,[\,\textit{who}\ s_j\,]_k\ [_{C'}\ C^\circ\ [_{IP}\ t_k\ [_{I'}\ [_{VP}\ [\,\textit{what}\ s_i\,]_j\ [_{V'}\ [\,\textit{to whom}\,]_i\ \textit{gave}\,]\,]\ I^\circ\,]\,]\,]\,]$

(13) German

a. $\emptyset :: \texttt{=i.+}\underline{\texttt{wh}}\texttt{.c}$
b. *who* $:: \texttt{d.}\nabla\texttt{wh.-wh}$
c. *what* $:: \texttt{d.}\nabla\texttt{wh.}\triangle\texttt{wh}$
d. *to-whom* $:: \texttt{d.}\triangle\texttt{wh}$

It is less clear whether there are languages in which m-licensors are weak and c-licensors are strong. Such languages should show *wh*-clustering in a position lower than Spec,CP.[4]

(14) $[_{CP}\ s_k\ [_{C'}\ C^\circ\ [_{IP}\ [\,\textit{who}\ [\,\textit{what}\ [\,\textit{to whom}\,]_i\,]_j\,]_k\ [_{I'}\ I^\circ\ [_{VP}\ t_j\ [_{V'}\ t_i\ \textit{gave}\,]\,]\,]\,]\,]\,]$

(15) unattested

a. $\emptyset :: \texttt{=i.+wh.c}$
b. *who* $:: \texttt{d.}\nabla\underline{\texttt{wh}}\texttt{.-wh}$
c. *what* $:: \texttt{d.}\nabla\underline{\texttt{wh}}\texttt{.}\triangle\texttt{wh}$
d. *to-whom* $:: \texttt{d.}\triangle\texttt{wh}$

(16) summarizes the resulting typology of languages with multiple-*wh*-interrogatives (cf. Grewendorf 2001, p. 105):[5]

(16)

a. $\texttt{+}\underline{\texttt{wh}}$ $\nabla\underline{\texttt{wh}}$ (Bulgarian)
b. $\texttt{+}\underline{\texttt{wh}}$ $\nabla\texttt{wh}$ (German)
c. $\texttt{+wh}$ $\nabla\underline{\texttt{wh}}$ (unattested)
d. $\texttt{+wh}$ $\nabla\texttt{wh}$ (Japanese)

4 Grewendorf (2001, p. 105, fn. 30) cites Russian and Hungarian as candidates for this language type (cf. Richards 2001, p. 33).

5 Languages without multiple-*wh*-interrogatives such as Irish (McCloskey 1979) can be viewed as lacking the means for cluster formation.

3. Some Constraints on Local (Re-)Ordering of *Wh*-Phrases

The simple picture outlined in Section 2 is perturbed by numerous factors as witnessed by the voluminous literature on multiple-*wh*-interrogatives.[6] Most notably, the issue of word order restrictions on multiple *wh*-phrases has been debated extensively. In this section we will first look at (how to capture some basic facts about) the reordering of multiple *wh*-objects as well as the behavior of *wh*-subjects and *wh*-adjuncts (Section 3.1). Secondly, we will discuss whether the previously introduced typology of multiple-*wh*-languages has to be enriched in the light of facts from (languages like) Polish (Section 3.2).

3.1 Subjects, Objects, Adjuncts

It is generally agreed that direct and indirect object can permute in Bulgarian multiple-*wh*-interrogatives. A simple account for this would allow *to whom* to scramble across *what* before *wh*-clustering occurs. Of course, the assignment of c-licensors and c-licensees among these two constituents has to be reversed. (17) shows the essential derivational steps and (18) presents the necessary lexical changes.

(17) a. $[_{\mathrm{VP}}$ [*what*] $[_{\mathrm{V'}}$ [*to whom*] *gave*]]
b. $[_{\mathrm{VP}}$ [*to whom* $]_i$ $[_{\mathrm{VP}}$ [*what*] $[_{\mathrm{V'}}$ t_i *gave*]]]
c. $[_{\mathrm{VP}}$ [[*to whom* $]_i$ [*what* $]_j$] $[_{\mathrm{VP}}$ t_j $[_{\mathrm{V'}}$ t_i *gave*]]]

(18) a. *what* :: $\mathtt{d}.\triangle\mathtt{wh}$
b. *to-whom* :: $\mathtt{d}.{\sim}\mathtt{v}.\triangledown\underline{\mathtt{wh}}.\triangle\mathtt{wh}$

Note that $\sim$v on *to whom* triggers *scrambling* to VP. Technically, a *s(cramble)-licensee*, $\sim$x, checks against a categorial feature x. This leads to *(overt) left-adjunction* of the constituent displaying $\sim$x to the one displaying x and to unilateral checking/elimination of $\sim$x (cf. Frey & Gärtner 2002).[7] Keeping the remaining featural assignments in tact, we can derive (19) as desired.

(19) $[_{\mathrm{CP}}$ [*who* [[*to whom* $]_i$ [*what* $]_j$ $]_k$ $]_l$
$[_{\mathrm{C'}}$ C° $[_{\mathrm{IP}}$ t_l $[_{\mathrm{I'}}$ I° $[_{\mathrm{VP}}$ t_k $[_{\mathrm{VP}}$ t_j $[_{\mathrm{V'}}$ *gave* t_i]]]]]]]

---

6 Among the more recent comprehensive sources we would like to mention the monographs by Pesetsky (2000) and Richards (2001), the collections edited by Boeckx & Grohmann (2003) and Stepanov et al. (2004), as well as the overview articles by Bayer (2006), Dayal (2006), and Cheng (2009). A multitude of further references can be found there.

7 For formal details, see Appendix.

At the same time, it has regularly been assumed that *wh*-subjects must be the initial element in Bulgarian *wh*-clusters. It is quite instructive to study various ways of enforcing such a constraint, so let us dwell on this point for a while. Since scrambling gave rise to permutation among objects, one ingredient in preventing non-initial *wh*-subjects could be a ban on scrambling across subjects in Bulgarian. One way of doing this would be to say that Bulgarian lacks s-licensee ~i, i.e. there is no scrambling to IP. However, it is clear that such an assumption would not be sufficient. Crossing could be effected by some other displacement type, e.g. *wh*-movement. Thus, consider (20), based on the altered lexical specifications in (21).

(20) $[_{CP}$ [[*what* $]_i$ [*who* $]_j$] $[_{C'}$ C° $[_{IP}$ t_j $[_{I'}$ I° $[_{VP}$ *discovered* t_i]]]]]

(21) a. *who* :: d.△wh
b. *what* :: d.-wh.∇$\underline{\text{wh}}$

Here, the direct object *wh*-moves to Spec,CP and then clustering right-attaches the subject to it. Grewendorf (2001, p. 94) introduces a constraint that rules out (20) rather directly, namely, a constraint that strictly forbids *wh*-cluster-formation to take place in Spec,CP. This could be translated into MG-terms by assuming that -wh.∇$\underline{\text{wh}}$ is not a licit substring of formal features. Such a constraint would be part of a theory of *improper movement* (cf. Müller & Sternefeld 1993). Let us record this assumption in (22).[8]

(22) *Improper Movement* - a
* ... -wh.∇$\underline{\text{wh}}$ ...

(20) could alternatively be ruled out on the basis of a theory of *relativized minimality* (Rizzi 1990, 2001). Accordingly, applications of *move* would have to be sensitive to features intervening between m-licensor and m-licensee. In the case at hand, △wh on *who* should prevent +$\underline{\text{wh}}$ in C° to attract -wh on *what* at the derivational stage depicted in (23).

(23) $[_{C'}$ C° $[_{IP}$ *who* $[_{I'}$ I° $[_{VP}$ *discovered what*]]]]

Stabler (2008) has begun working out the details of how to introduce relativized minimality into MGs and we refer readers to that work for further discussion.

A potential alternative to the crossing constraints just discussed would be to disallow *wh*-subjects endowed with feature △wh. To make this less arbitrary, one could link the constraint to (the licensing of) nominative case. If nominative case is licensed via

[8] Obviously, similar restrictions would have to be imposed on topicalization and other types of movements into a "richer" left periphery (Rizzi 1997).

movement—e.g. by taking I° to possess +<u>nom</u> and by base-generating "subjects" with feature -nom in a "low" subject position like Spec,vP—the required constraint could be added to the theory of improper movement as follows.

(24) *Improper Movement* - b
  * ... -nom ... △wh ...

This kind of approach makes the empirical prediction that the ban on non-initial *wh*-subjects in Bulgarian does not only hold locally but persists when *wh*-subjects are extracted into higher interrogative clauses. At this point we will not decide the issue but move on.

A third domain of ordering constraints concerns the placement of *wh*-adjuncts. Famously, the following contrast has been observed for English (Huang 1995, p. 153).

(25) a. *Why did you buy what?*
  b. **What did you buy why?*

Why seems to behave like Bulgarian *wh*-subjects in requiring top position in multiple-*wh*-interrogatives. It is fairly plausible to assume that English lacks scrambling. Instead, (25-b) could arise if cluster formation took place after *what* has been *wh*-moved to Spec,CP. Again, either a well-designed theory of relativized minimality blocks this, the constraint on improper movement in (22) rules it out, or a specific stipulation concerning *wh*-adjuncts is in order. Taking adjuncts to be introduced by the operation *adjoin* that checks an *a(djoin)-selector* ≈x against a categorial feature x and unilaterally eliminates ≈x (Frey & Gärtner 2002), we can formulate another constraint on improper movement to capture the contrast in (25).

(26) *Improper Movement* - c
  * ... ≈x ... △wh ...

(26), however, is clearly too general as (27-b) shows (Fanselow 2004, p. 114).

(27) a. *Where did you see what?*
  b. *What did you see where?*

A lexical stipulation affecting just *why* (namely, * *why* :: ... △wh ...) may thus be preferable.[9] (27-b) also indicates that—unless *where* can be base-generated in a low

[9] We will not be able to do justice to the pecularities of *wh*-adjuncts like *why* here. Stepanov & Tsai (2008) devote a full paper to that subject. Also we will not have anything new to say about the effects of *referentiality* and *D-linking* on multiple-*wh*-interrogatives (cf. Pesetsky 1987, Rizzi 1991, Szabolcsi & Zwarts 1993, Comorovski 1996, Reinhart 1998, Dayal 2006, among many others), except that

position—(22) may be too strong.

Further complications show up once adjuncts in additional languages are taken into account. Consider the following contrast in Japanese (Watanabe 1992, p. 268).

(28) a. ?*Kimi-wa nani-o naze katta no?*
lit. 'You what why bought?'
b. **Kimi-wa naze nani-o katta no?*
lit. 'You why what bought?'

Recall that in Japanese both *wh*-movement and *wh*-clustering are covert. Nevertheless, word order has an effect. Given that Japanese possesses scrambling, the analysis that suggests itself is to say that *naze* ('why') in Japanese may not have $\nabla\texttt{wh}$ but can be licensed by a higher *wh*-phrase like *nani-o* ('what'). (29) shows the required lexical specifications.

(29) a. *naze* :: $\approx\texttt{i}.\triangle\texttt{wh}$
b. *nani-o* :: $\texttt{d}.\sim\texttt{i}.\nabla\texttt{wh}.-\texttt{wh}$

Given this, *naze* will be base-generated in IP-adjoined position and *nani-o* will scramble across it. Then *wh*-clustering and *wh*-movement occur.[10] Grewendorf (2001, p. 94) deems the Japanese kind of pattern important enough to rule out $\nabla\texttt{wh}$ on *wh*-adjuncts. He, thus, argues for a counterpart of (30).

(30) *Improper Movement* - c
a. $* \ldots \approx\texttt{x} \ldots \nabla\underline{\texttt{wh}} \ldots$
b. $* \ldots \approx\texttt{x} \ldots \nabla\texttt{wh} \ldots$

Yet, in view of the English data in (25-a) and (27-a) and the variation reported by Fanselow (2004, p. 95) on German, (31) (cf. Stepanov & Tsai 2008, p. 619, on Serbo-Croatian), a more cautious strategy seems to be advisable.

(31) a. *Wer lachte warum?*
lit. 'Who laughed why?'
b. *Warum lachte wer?*
lit. 'Why laughed who?'

---

it would be compatible with the MG-approach developed here to exempt certain *wh*-phrases from cluster formation and/or *wh*-movement.

10 In addition, it has to be assumed that *kimi-wa* is overtly topicalized.

3.2 Polish

A further well-known pattern of multiple *wh*-phrases arises in languages like Polish (and Serbo-Croatian). These languages are like Bulgarian in that they have overt *wh*-clustering and like German in that only one *wh*-phrase overtly occupies Spec,CP. How can that be so? Assume that Polish is lexically specified like Bulgarian except that the highest *wh*-object possesses an additional feature for scrambling to IP.

(32) Polish

a. $\emptyset$ :: $\texttt{=i.+}\underline{\texttt{wh}}\texttt{.c}$
b. *who* :: $\texttt{d.}\nabla\underline{\texttt{wh}}\texttt{.-wh}$
c. *what* :: $\texttt{d.}\nabla\underline{\texttt{wh}}\texttt{.}\triangle\texttt{wh.}{\sim}\texttt{i}$
d. *to-whom* :: $\texttt{d.}\triangle\texttt{wh}$

This will yield (33).

(33) $[_{\text{CP}}\,[\,\textit{who}\ \text{t}'_j\,]_k\,[_{\text{C}'}\ \text{C}^\circ\ [_{\text{IP}}\,[\,\textit{what}\,[\,\textit{to whom}\,]_i\,]_j\,[_{\text{IP}}\ \text{t}_k\,[_{\text{I}'}\ \text{I}^\circ\,[_{\text{VP}}\ \text{t}_j\,[_{\text{V}'}\ \textit{gave}\ \text{t}_i\,]\,]\,]\,]\,]\,]\,]$

What has happened here is that after *wh*-clustering into Spec,IP, $[_<$ *what to-whom* $]$ moves on via scrambling to left-adjoin to IP. Then the remnant $[_<$ *who* t $]$ *wh*-moves to Spec,CP.

From the perspective of MGs, the derivation just sketched has at least one disadvantage: it violates the *Specifier Island Condition (SPIC)* (Stabler 1999; Gärtner & Michaelis 2005) according to which no proper subtree of a specifier can be moved out of that specifier.[11] Note, however, that the impact of the SPIC on complexity is different from that of the SMC (cf. Gärtner & Michaelis 2005; Kobele & Michaelis 2009; Michaelis 2009). Also, the empirical validity of the SPIC has been questioned (Chomsky 2008). Still, it is useful to study SPIC-compatible alternative derivations of Polish multiple-*wh*-interrogatives. Thus, consider the following variant of (32)/(33) where the top *wh*-phrase possesses a weak instead of a strong c-licensor, i.e. (32-b) is changed to *who* :: $\texttt{d.}\nabla\texttt{wh.-wh}$:

[11] For Grewendorf (2001, p. 112, fn. 40) the derivation would be ruled out because it violates a particular freezing constraint on adjunction according to which once a constituent adjoins to another it cannot move away from that position on its own (cf. Grewendorf & Sabel 1999). This requires that cluster formation counts as adjunction. The MG-formalism presented here would need extra assumptions to enforce this. The freezing constraint is not compatible, among other things, with our analysis of *wh*-object reordering in Bulgarian, (17)-(19). Sabel (2001) draws on parallels between *wh*-clusters and clitic clusters created through *incorporation*. In such a system, constraints on *excorporation* could be invoked in ruling out (33).

(34) $[_{\text{CP}}$ [*who* s_j $]_k$ $[_{\text{C}'}$ C° $[_{\text{IP}}$ [*what* [*to whom* $]_i$ $]_j$ $[_{\text{IP}}$ t_k $[_{\text{I}'}$ I° $[_{\text{VP}}$ t_j $[_{\text{V}'}$ *gave* t_i]]]]]]]

Here, $[_<$ *what to-whom*] clusters into the top *wh*-phrase covertly before it scrambles to IP overtly. Finally, [*Who* s_j] undergoes *wh*-movement to Spec,CP. What may appear strange about (34) is that $[_<$ *what to-whom*] seems to go off in two directions from its VP-internal position. This "forking effect" calls for deeper analysis of what is meant by overt vs. covert movement and how this determines structural assumptions about full constituents, shadows, and (the) traces (they leave). This will be addressed in Section 4.[12]

Let us turn to a second SPIC-compatible derivation of the Polish pattern. Here, scrambling precedes clustering. The required lexical changes are given in (35) and the resulting syntactic structure is illustrated in (35).

(35) a. *who* :: `d.-wh.∇wh`
b. *what* :: `d.∇`$\underline{\texttt{wh}}$`.~i.△wh`

(36) $[_{\text{CP}}$ [[*who* $]_k$ s_j] $[_{\text{C}'}$ C° $[_{\text{IP}}$ [*what* [*to whom* $]_i$ $]_j$ $[_{\text{IP}}$ t_k $[_{\text{I}'}$ I° $[_{\text{VP}}$ t_j $[_{\text{V}'}$ *gave* t_i]]]]]]]

Who has moved to Spec,CP alone and clustering occurs there instead of in Spec,IP. Note that (35-a) is at variance again with the previously discussed constraint on *improper movement* in (22).

3.3 Preliminary Discussion

We have seen in the previous two (sub)sections how MGs are suited to capture variation in terms of the interaction of general and specific constraints. The theory of relativized minimality aside (Stabler 2008), we have stated various candidate constraints on *improper movement*. Apart from exploring their cross-linguistic validity, we must ask about their formal implementation. In particular it is an open question to what extent such a theory is part of a theory of (meta-)lexical specifications regulating possible strings of formal features.

This question in turn raises the issue of the proper treatment of parameterization in MGs and how it relates to "classical" minimalist ideas about confining parameters to functional heads (cf. Chomsky 1995). The more local specifications we have seen concerning subjects, adjuncts and instances of the latter like *why* and *where* feed into the same discussion. Now, at least the SPIC-compatible treatment of Polish in Section

[12] Grewendorf (2001, p. 106f) makes use of such a forking derivation in the analysis of multiple-*wh*-interrogatives in German.

3.2 lends itself to more classical parameterization in terms of feature strength. If such an approach is on the right track it requires differential treatment of top *wh*-phrases in *wh*-clusters. In the Polish case, we postulated a weak c-licensor, $\nabla\mathtt{wh}$, for top *wh*-phrases and a strong one, $\nabla\underline{\mathtt{wh}}$, for all others.[13] If this is generalized, the four way typology in (16) has to give way to the following eight way one. (We mark the top c-licensor with subscripted "T.")

(37)	a.	$+\underline{\mathtt{wh}}$	$\nabla\underline{\mathtt{wh}}_{\mathrm{T}}$	$\nabla\underline{\mathtt{wh}}$	(Bulgarian)
	b.	$+\underline{\mathtt{wh}}$	$\nabla\underline{\mathtt{wh}}_{\mathrm{T}}$	$\nabla\mathtt{wh}$	?
	c.	$+\underline{\mathtt{wh}}$	$\nabla\mathtt{wh}_{\mathrm{T}}$	$\nabla\underline{\mathtt{wh}}$	(Polish)
	d.	$+\underline{\mathtt{wh}}$	$\nabla\mathtt{wh}_{\mathrm{T}}$	$\nabla\mathtt{wh}$	(German)
	e.	$+\mathtt{wh}$	$\nabla\underline{\mathtt{wh}}_{\mathrm{T}}$	$\nabla\underline{\mathtt{wh}}$	(unattested)
	f.	$+\mathtt{wh}$	$\nabla\underline{\mathtt{wh}}_{\mathrm{T}}$	$\nabla\mathtt{wh}$	?
	g.	$+\mathtt{wh}$	$\nabla\mathtt{wh}_{\mathrm{T}}$	$\nabla\underline{\mathtt{wh}}$	?
	h.	$+\mathtt{wh}$	$\nabla\mathtt{wh}_{\mathrm{T}}$	$\nabla\mathtt{wh}$	(Japanese)

4. Refining the MG-Treatment of *Wh*-Clustering

It is now time for being more specific about the distinction between overt and covert dependency formation and the structure of traces and shadows as well as their antecedents and "followers." Note, first of all, that MGs have tended to lack copying, not the least because copying has the potential of "multiplying resources" and thus jeopardizing resource-sensitivity (Stabler 1998, p. 81).[14] Also, MGs have worked on the assumption that lexical items are represented as (feature) triples of type $\langle \pi, \sigma, \iota \rangle$ containing a phonetic component, π, a syntactic one, σ, consisting of the string of formal features, and a semantic, or, *interpreted* subpart, ι. Our lexical specifications so far—e.g. (10) and (11)—showed only π (in graphematic guise) and σ. Also, overt vs. covert displacement was meant to distinguish carrying along vs. leaving behind the π-component of a constituent. What has to be clarified is the character of traces and shadows beyond their π-lessness.

Take the forking derivation in (34). Given that $[_<$ *what to-whom*] underwent scrambling after covert clustering occurred and given that scrambling is driven by syntactic features, $\sim\mathtt{x}$ being part of σ, $[_<$ *what to-whom*] had to have kept its syntactic

[13] Differential treatment of top vs. non-top *wh*-phrases is, of course, what Rudin (1988) proposes to deal with the difference between Polish and Bulgarian. Similar distinctions (must) show up in one form or another in other approaches too. For Richards (2001), for example, differential treatment is implicit in the *Principle of Minimal Compliance*. Accordingly, if one *wh*-phrase—the highest one due to *Shortest Attract*—has satisfied locality constraints all other (lower) ones can violate them.

[14] See, however, alternatives explored by Kobele (2006) and Stabler (2007).

component σ. This leaves two options for the structure of s_j. Either it is the radically empty triple $\langle \varepsilon, \varepsilon, \varepsilon \rangle$ (labeling an empty head) or it contains the interpreted component, which means that s_j represents a full subtree corresponding structurally to [$_<$ *what to-whom*] with heads labeled by $\langle \varepsilon, \varepsilon, \iota \rangle$-triples. We take it that the latter option is the preferred one. *Wh*-clustering and subsequent *wh*-movement has effects on the scope of *wh*-phrases and thus an impact on semantics. In fact, covert *wh*-clustering should be seen as closely related to operations like *quantifier raising (QR)* and "classical" covert *wh*-movement.

However, if shadows carry along the ι-component of a constituent and copying is disallowed, all that is left for scrambling in (34) is an ι-less configuration of $\langle \pi, \sigma, \varepsilon \rangle$-triples. Would that be compatible with the information structural effects commonly assumed to accompany scrambling? In fact, in frameworks like *alternative semantics* (Rooth 1985), a *focus semantic value* of an information structurally marked constituent has to be computed on the basis of its standard denotation, so ι would seem to have to be present in the target position of scrambling.[15] Let us take this dilemma as an argument against allowing forking derivations of the kind shown in (34).

It follows from this very brief discussion that any fully explicit MG needs to sort operations into those that do and those that do not carry along ι and π. Assuming—for the sake of simplicity—an "invariant" semantics, the former decisions will be language-independent while the latter ones are the locus of language-specific parameterization. (38) lists the operations (under discussion in this paper) that carry along ι.

(38) $+\iota$: *move-wh*, *cluster-wh*, *scramble*

For Bulgarian, the set of operations that carry along π is identical, i.e., all displacement is overt:

(39) Bulgarian
$+\pi$: *move-wh*, *cluster-wh*, *scramble*

In fact, we can assume that there is no covert scrambling,[16] so all languages that have scrambling will have it in their "$+\pi$-set." Japanese will have the specifications in (40) whereas German follows the pattern in (41):

[15] The same conclusion can be reached if one adopts the *structured meanings* approach to information structure (von Stechow 1991; Krifka 1993). There, focused constituents induce a split of the formal semantic representation into focus and background component.

[16] Unless (the MG-counterpart of) QR has to be treated in terms of scrambling.

(40) Japanese

a. $+\pi$: *scramble*
b. $-\pi$: *move-wh*, *cluster-wh*

(41) German

a. $+\pi$: *move-wh*, *scramble*
b. $-\pi$: *cluster-wh*

Finally, the SPIC-compatible analysis of Polish—see (37) above—dictates the following specifications. (We mark non-top c-licensors with subscripted "NT."):

(42) Polish

a. $+\pi$: *move-wh*, *cluster-wh*$_{\mathrm{NT}}$, *scramble*
b. $-\pi$: *cluster-wh*$_{\mathrm{T}}$

So, what about formal features, i.e. the σ-component? If forking derivations are strictly banned, it never makes sense to leave σ behind. It follows that "classical" traces denoted by *t* in earlier structures are always σ-less in addition to their π-lessness. This will guarantee the "immobility of traces" often discussed in the literature (Chomsky 1995, p. 304; Stabler 1997a, p. 171). In fact, for the cases discussed so far, *t* can be taken to stand for the radically empty triple $\langle \varepsilon, \varepsilon, \varepsilon \rangle$ (labeling an empty head) (cf. Stabler 1997a, p. 163).[17]

The upshot of these considerations is that virtually—as already envisaged in some detail by Stabler (1997a)—the MG-formalism contains not just one operation *move*, *cluster* etc. but a family of closely related such operators differing in their distribution of the $\langle \pi, \sigma, \iota \rangle$-components among source and target position. (43) lists the possible variants of *move*:

(43) *move-*$\pi\sigma\iota$, *move-*$\pi\sigma$, *move-*$\pi\iota$, *move-*π, *move-*$\sigma\iota$, *move-*σ, *move-*ι,

where *move-*$\pi\sigma\iota$ is taken to displace π, σ, and ι, *move-*$\pi\sigma$ to displace π and σ, and so on.[18] So far we have only appealed to *wh*-movement and we just argued that *move-*

---

[17] For the purpuse of semantic interpretation it may be necessary to leave variables in the base positions of *wh*-expressions (cf. Stabler 1997a; Kobele 2006). Reluctantly, we have to put aside the issue of providing a semantics for *wh*-clustering, as this leads into the intricacies of quantifier composition and/or absorption. Work by Keenan (1989, 1992), van Eijck (2005), and Higginbotham & May (1981) may serve as starting point here. What seems to be clear is that interpretation of derived trees ("logical forms") will have to be an option (cf. Haida 2007).

[18] As far as formalization goes, we have decided to treat *move-*σ not in terms of feature displacement but as a mechanism for pure feature deletion leaving tree structure unaffected otherwise (see Appendix). This operation has become known as *Agree* in the more recent literature (Chomsky 2000).

$\pi\sigma\iota$ applies in Bulgarian, German and Polish, while *move-*$\sigma\iota$ applies in Japanese. Formally, we deal with this by assuming that m-licensors indicate which brand of *move* is triggered (see Appendix). Thus, instead of marking the rough distinction between π-bearing and π-less *move* by underlining, we index m-licensors "with the components" they attract. Bulgarian, German, and Polish have $\texttt{+wh}_{(\pi\sigma\iota)}$ and Japanese has $\texttt{+wh}_{(\sigma\iota)}$ (see Appendix).

Now, a systematic and comprehensive investigation of the *algebra* of possible operation types in MGs is still outstanding.[19] We are certainly not in a position to undertake anything like it here. Let us instead wrap up this section by addressing one of the movement types not considered so far, namely, *move-*π. Staying in the realm of *wh*-movement, consider the following *wh-imperative* discussed by Reis & Rosengren (1992).

(44) *Wieviel schätz mal, dass das kostet?*
how.much estimate once that that costs
'Guess how much that costs!'

This is interpreted as a request asking the addressee to make a guess concerning the question as to how much a contextually specified thing costs. Thus, initial position of the *wh*-expression *wieviel* is misleading. It has "overshot" its mark, i.e. the position it is interpreted in, which is Spec,CP of the interrogative complement of *schätz mal* ('guess'). (44), then, is a good candidate for an application of *move-*π. It would be naïve, however, to endow the matrix C° with $\texttt{+wh}_{(\pi)}$, at least if $\texttt{+wh}_{(\dots)}$ should only occur in interrogatives. So, to the extent that (44) does not have any additional information structural effect—which would require an application of *move-*$\pi\iota$—some purely formal movement trigger $\texttt{+f}_{(\pi)}$ with counterpart $\texttt{-f}$ on *wieviel* is called for in the case of (44).[20] According to Stabler (1997a, p. 163) certain types of movement for case checking may be of the *move-*π type. Otherwise, everything that has been discussed under the label *PF-movement* would seem to belong here too.[21]

5. Further Discussion and Outlook

This final section will be devoted to outlining the directions in which a *wh*-clustering approach to multiple-*wh*-interrogatives has to be developed further and how this relates

[19] Stabler (1997a) sets the stage for such an undertaking. Work by, among others, Pesetsky (2000), Kracht (2001), Bobaljik (2002), and Kobele (2006) also sheds interesting light on this project.

[20] Frey (2004) discusses instances of so-called *formal movement* into the German *Vorfeld*.

[21] Zwart (2001), for example, discusses the contentious issue as to whether verb movement should take place (just) at PF.

to and impacts on the MG-formalism. In particular, we will address the following two issues. First, an argument is given why *wh*-clustering should be generalized to all languages instead of being confined to just a few special languages as envisaged by Grewendorf (2001) and Sabel (2001). Secondly, we will show how MG-style *wh*-clustering accounts for *additional wh-effects*, demonstrating that the theory has considerable staying power even confronted with non-trivial empirical facts.

5.1 Generalized *Wh*-Clustering

We have shown how *wh*-clustering allows an SMC-compatible MG-treatment of multiple-*wh*-interrogatives. Thus, this domain of syntax—contrary to initial appearance—does not threaten the mild context-sensitivity of MGs. As a consequence, adopting the MG-formalism is a motivation for trying to apply the tool of *wh*-clustering to multiple-*wh*-interrogatives in all languages, i.e., a motivation for generalizing *wh*-clustering. This, however, is not what Grewendorf (2001) and Sabel (2001) envisaged. They argued that *wh*-clustering is only found in languages whose inventories of interrogative and (pure) indefinite pronouns either show a considerable overlap or are related by regular morphological processes. As originally discussed in more detail by Cheng (1991; cf. Haspelmath 1997), this holds, among others, for Slavic languages like Bulgarian, Polish, and Russian, eastern Asian languages like Chinese, Japanese, and Korean and Germanic languages like German. However, it does not apply to English. Thus, while German allows interrogative and purely indefinite uses of pronouns like *wer* and *was* and Bulgarian "derives" indefinite pronouns from interrogative ones by prefix *njá*, English uses *who* and *what* as interrogatives and *someone* and *something* as (pure) indefinites the latter, not being morphologically related to the former.[22] Adapting the approach by Cheng (1991), Grewendorf (2001, p. 96) assumes that interrogative pronouns in languages of the former type possess an empty D° endowed with a feature able to attract other *wh*-phrases, i.e., endowed with a counterpart of the MG-feature $\triangledown\underline{\texttt{wh}}$, respectively $\triangledown\texttt{wh}$. The structure is basically the one given by Cheng (1991, p. 86). In MG-terms, an item like Bulgarian *koj* as it figures in (5)/(10-b) would actually be the constituent $[_< \emptyset$ *koj* $]$ resulting from merging the two lexical items given in (45).

(45) a. *koj* :: $\texttt{n}$
b. $\emptyset$ :: $\texttt{=n.d.}\triangledown\underline{\texttt{wh}}\texttt{.-wh}$

Now, although the link between *wh*-clustering and the morphology of *wh*-pronouns is certainly an interesting one worthy of broader typological exploration, we think that

[22] Note the ungrammaticality of **somewho* and the meaning shift resulting for *somewhat*.

the hypothesis of generalized *wh*-clustering is (at least) equally well defensible. We can only argue this point indirectly here by mentioning some worries we have with Grewendorf's (and Sabel's) approach.

First, if $\triangledown\underline{\texttt{wh}}$, respectively $\triangledown\texttt{wh}$, is a feature of a phonetically null D° merging with an otherwise "defective" *wh*-pronoun, how does the theory deal with *wh*-determiners, like German *welch* ('which'), that themselves serve as D° in interrogative DPs, like *welches Buch* ('which book'), as well as purely indefinite ones, like *(da sind) welche* ('there are some')? In fact, Cheng (1991, p. 85) considers items like $\emptyset$ in (45) "to be the null counterpart of *which* in English." The problem is that *welch* never merges with *wh*-cores of category NP (**welchwer*, **welchwas*). Instead it participates in the regular morphological process that turns German *wh*-pronouns into pure indefinites (*irgendwer*, *irgendwas*, . . . , *irgendwelche*).[23] But then, if *welch* can be endowed with $\triangledown\texttt{wh}$ directly, there is no reason why other *wh*-expressions could not. However, if the reliance on null determiners breaks down, the typological distinction based on morphology is no longer independently compelling.

Secondly, we already mentioned that Grewendorf (2001) takes *wh*-adjuncts to be unable to bear $\triangledown\underline{\texttt{wh}}$, respectively $\triangledown\texttt{wh}$. Explicitly, it is stated that "[a]s far as *wh*-adjuncts are concerned, we can assume that they lack a D-head that could be endowed with an uninterpretable feature [Q]" (Grewendorf 2001, p. 102), [Q] being the counterpart of $\triangledown\underline{\texttt{wh}}$, respectively $\triangledown\texttt{wh}$. We already indicated in Section 3.1 that such a rigid approach to *wh*-adjuncts may be empirically inadequate.[24] Also, we are worried about additional assumptions necessary to deal with well-formed multiple-*wh*-interrogatives involving *wh*-adjuncts under such an approach. In particular, for cases like (31-b) it is stipulated that—in addition to base generating *warum* in Spec,CP—D°-movement from *wer* to C° can take place (Grewendorf 2001, p. 117). Also, in order to deal with some intricate facts from Japanese involving the string of *wh*-phrases *dare-ni naze dare-ga* ('whom why who') it is allowed that *dare-ni* attracts both *naze* and *dare-ga* (Grewendorf 2001, p. 104, fn. 29), only to avoid endowing *naze* with $\triangledown\texttt{wh}$. The MG-entry of *dare-ni* would thus have to look as in (46).

(46) a. *dare-ni* :: $\texttt{d.}{\sim}\texttt{i.}\triangledown\texttt{wh.}\triangledown\texttt{wh.-wh}$

Such an analysis is unattractive from the perspective of MGs because, effectively, this will induce an SMC-violation, given that both *naze* and *dare-ga* have to display $\triangle\texttt{wh}$

[23] Following Cheng (1991, p. 85), this would require MG-specifications of the following kind:

(i) a. *welche* :: $\texttt{n}$
b. *irgend* :: $\texttt{=n.d}$

[24] Except for counterparts of *why*, *wh*-adjuncts in German do participate in the morphological process of creating pure indefinites: *irgendwo* ('somewhere'), *irgendwann* ('some time'), *irgendwie* ('somehow') (Grewendorf 2001, p. 105; for Bulgarian, see Haspelmath 1997, p. 267).

at the derivational stage at which *wh*-clustering starts.

In sum, although a full defense of *generalized wh-clustering* has to be left for another occasion, we see no substantial advantage in restricting *wh*-clustering to just a few languages as advocated by Grewendorf (2001) and Sabel (2001).

5.2 Additional *Wh*-Effects

In Section 2 and 3 we looked at constraints on only moderately complicated multiple-*wh*-patterns. Here we give a brief sketch of how *wh-clustering* deals with so-called *additional wh-effects*, an empirical domain that has received considerable attention in the literature. Consider the contrast in (47) (Grewendorf 2001, p. 103, fn. 27).

(47) a. **What did who give to Mary?*
b. *What did who give to whom?*

Famously, an additional *wh*-object can rescue an otherwise ill-formed subject-object inversion of *wh*-phrases, a.k.a. *superiority violation*. Without going into much depth, one can easily see that the two sentences will come out as structurally substantially different under *wh*-clustering. Thus, allowing covert *wh*-clustering to occur in English, the constituent that crosses *who* in (47b) is not just *what* but $[_<$ *what* s], i.e. a complex constituent containing the shadow of *to whom*. This indicates that the required distinction between (47-a) and (47-b) can be made on the basis of the theory of improper movement or relativized minimality.

While in (47) the additional *wh*-phrase occurs in a low position, high position "rescuers" exist in Bulgarian as shown in (48) (Grewendorf 2001, p. 91).

(48) a. **Koja kniga*$_i$ *otreče senatorăt* [*mălvata če iska da zabrani* t_i]*?*
'Which book did the senator deny the rumor that he wanted to ban?'
b. ?*Koj senator koja kniga*$_i$ *otreče* [*mălvata če iska da zabrani* t_i]*?*
'Which senator denied the rumor that he wanted to ban which book?'

Extraction of an object ('which book') from a complex NP island is licensed if the matrix subject is a *wh*-phrase ('which senator'). Obviously, the theory of islands must be sensitive to the difference between *move* and *cluster*, i.e. complex NPs are islands for bearers of `-wh` but not for bearers of `△wh`.[25]

[25] Grewendorf (2001, p. 101) discusses an example from Japanese where an additional "high" *wh*-phrase blocks the otherwise fine extraction of a *wh*-adjunct. Here, at first sight, `-wh` but not `△wh` on the *wh*-adjunct allows extraction. However, as noted by Grewendorf (ibid.), assumptions about intermediate landing sites may also be involved. For the sake of simplicity, we have chosen to leave treatment of successive-cyclic movement out of the picture.

Finally, consider the well-known ameliorating effect an outside *wh*-phrase in English can have on superiority violations like (4).

(49) *Who wonders what who prefers?*

(49) is acceptable if *who* in the subordinate interrogative takes matrix scope. In MG-terms this indicates that *what* can cross *who* when it bears just `-wh` but not when possessing `-wh.∇wh`.

In sum, the clustering approach to multiple-*wh*-interrogatives when developed within MGs looks quite promising in the more involved domain of *additional wh-effects* too.[26]

References

Bayer, J. (2006). Wh-in-Situ. In: Everaert & van Riemsdijk (2006). 376–438, Vol. V.

Bobaljik, J. (2002). A-Chains at the PF Interface. Copies and 'Covert' Movement. *Natural Language and Linguistic Theory* 20. 197–267.

Boeckx, C. & K. Grohmann (eds.) (2003). *Multiple Wh-Fronting*. Amsterdam: John Benjamins.

Cheng, L. L. S. (1991). *On the Typology of WH-Questions*. Ph.D. thesis. Massachusetts Institute of Technology, Cambridge, MA.

Cheng, L. L. S. (2009). Wh-in-Situ, From the 1980s to Now. *Language and Linguistics Compass* 3. 767–791.

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

Chomsky, N. (2000). Minimalist Inquiries. The Framework. In: R. Martin, D. Michaels & J. Uriagereka (eds.) *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press. 89–155.

Chomsky, N. (2008). On Phases. In: R. Freidin, C. Otero & M. L. Zubizaretta (eds.) *Foundational Issues in Linguistic Theory*. Cambridge, MA: MIT Press. 133–166.

Comorovski, I. (1996). *Interrogative Phrases and the Syntax-Semantics Interface*. Dordrecht: Kluwer.

Dayal, V. (2006). Multiple-Wh-Questions. In: Everaert & van Riemsdijk (2006). 275–326, Vol. III.

van Eijck, J. (2005). Normal Form for Characteristic Functions on n-ary Relations. *Journal of Logic and Computation* 15. 85–98.

[26] It would be interesting to study how the "magnetism" of *wh*-clustering relates to the connectedness idea developed by Kayne (1984).

Everaert, M. & H. van Riemsdijk (eds.) (2006). *The Blackwell Companion to Syntax*. Oxford: Blackwell.

Fanselow, G. (2004). The MLC and Derivational Economy. In: Stepanov et al. (2004). 73–123.

Frey, W. (2004). The Grammar-Pragmatics Interface and the German Prefield. *Sprache und Pragmatik* 52. 1–39.

Frey, W. & H.-M. Gärtner (2002). On the Treatment of Scrambling and Adjunction in Minimalist Grammars. In: *Proceedings of the Conference on Formal Grammar (FGTrento),* Trento. 41–52.

Gärtner, H.-M. & J. Michaelis (2005). A Note on the Complexity of Constraint Interaction. Locality Conditions and Minimalist Grammars. In: P. Blache, E. Stabler, J. Busquets & R. Moot (eds.) *Logical Aspects of Computational Linguistics (LACL '05)*, LNAI Vol. 3492. Berlin, Heidelberg: Springer. 114–130.

Gärtner, H.-M. & J. Michaelis (2007). Some Remarks on Locality Conditions and Minimalist Grammars. In: U. Sauerland & H.-M. Gärtner (eds.) *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax and Semantics*. Berlin: Mouton de Gruyter. 161–195.

Grewendorf, G. (2001). Multiple Wh-Fronting. *Linguistic Inquiry* 32. 87–122.

Grewendorf, G. & J. Sabel (1999). Scrambling in German and Japanese. Adjunction versus Multiple Specifiers. *Natural Language and Linguistic Theory* 17. 1–65.

Haida, A. (2007). *The Indefiniteness and Focusing of Wh-Words*. Ph.D. thesis. Humboldt Universität zu Berlin, Berlin.

Haspelmath, M. (1997). *Indefinite Pronouns*. Oxford: Clarendon Press.

Higginbotham, J. & R. May (1981). Questions, Quantifiers, and Crossing. *The Linguistic Review* 1. 41–79.

Huang, J. C.-T. (1995). Logical Form. In: G. Webelhuth (ed.) *Government and Binding Theory and the Minimalist Program. Principles and Parameters in Syntactic Theory*. Oxford: Blackwell. 125–175.

Joshi, A. K., K. Vijay-Shanker & D. J. Weir (1991). The Convergence of Mildly Context-Sensitive Grammar Formalisms. In: P. Sells, S. M. Shieber & T. Wasow (eds.) *Foundational Issues in Natural Language Processing*. Cambridge, MA: MIT Press. 31–81.

Kayne, R. S. (1984). *Connectedness and Binary Branching*. Dordrecht: Foris.

Keenan, E. L. (1989). Semantic Case Theory. In: R. Bartsch, J. van Benthem & P. van Emde Boas (eds.) *Semantics and Contextual Expression*. Dordrecht: Foris. 33–56.

Keenan, E. L. (1992). Beyond the Frege Boundary. *Linguistics and Philosophy* 15. 199–221.

Kobele, G. M. (2006). *Generating Copies. An investigation into structural identity in* *language and grammar*. Ph.D. thesis. University of California, Los Angeles, CA.

Kobele, G. M. & J. Michaelis (2009). Two Type-0 Variants of Minimalist Grammars. In: Rogers (2009). 81–91.

Kracht, M. (2001). Syntax in Chains. *Linguistics and Philosophy* 24. 467–529.

Krifka, M. (1993). Focus and Presupposition in Dynamic Interpretation. *Journal of Semantics* 10. 269–300.

McCloskey, J. (1979). *Transformational Syntax and Model Theoretic Semantics*. Dordrecht: Reidel.

Michaelis, J. (2001a). Derivational Minimalism is Mildly Context-Sensitive. In: M. Moortgat (ed.) *Logical Aspects of Computational Linguistics (LACL '98)*, LNAI Vol. 2014. Berlin, Heidelberg: Springer. 179–198.

Michaelis, J. (2001b). *On Formal Properties of Minimalist Grammars*. (Linguistics in Potsdam 13). Potsdam: Universitätsbibliothek, Publikationsstelle.

Michaelis, J. (2009). An Additional Observation on Strict Derivational Minimalism. In: Rogers (2009). 101–111.

Müller, G. & W. Sternefeld (1993). Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24. 461–507.

Pesetsky, D. (1987). Wh-in-Situ. Movement and Unselective Binding. In: E. Reuland & A. ter Meulen (eds.) *The Representation of (In)definiteness*. Cambridge, MA: MIT Press. 98–129.

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

Reinhart, T. (1998). Wh-in-Situ in the Framework of the Minimalist Program. *Natural Language Semantics* 6. 29–56.

Reis, M. & I. Rosengren (1992). What do *WH*-Imperatives Tell us about *WH*-Movement. *Natural Language and Linguistic Theory* 10. 79–118.

Richards, N. (2001). *Movement in Language. Interactions and Architectures*. Oxford: Oxford University Press.

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

Rizzi, L. (1991). On the Status of Referential Indices. In: A. Kasher (ed.) *The Chomskyan Turn*. Oxford: Blackwell. 273–299.

Rizzi, L. (1997). The Fine Structure of the Left Periphery. In: L. Haegeman (ed.) *Elements of Grammar*. Dordrecht: Kluwer. 281–337.

Rizzi, L. (2001). Relativized Minimality Effects. In: M. Baltin & C. Collins (eds.) *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell. 89–110.

Rogers, J. (ed.) (2009). *Proceedings of FG-MoL 2005: The 10th Conference on Formal Grammar and The 9th Meeting on Mathematics of Language*. Stanford, CA: CSLI Publications.

Rooth, M. (1985). *Association With Focus*. Ph.D. thesis. University of Massachusetts, Amherst, MA.

Rudin, C. (1988). On Multiple Questions and Multiple Wh-Fronting. *Natural Language and Linguistic Theory* 6. 445–501.

Sabel, J. (2001). Deriving Multiple Head and Phrasal Movement. The Cluster Hypothesis. *Linguistic Inquiry* 32. 532–547.

Stabler, E. P. (1997a). Computing Quantifier Scope. In: A. Szabolcsi (ed.) *Ways of Scope Taking*. Dordrecht: Kluwer. 155–182.

Stabler, E. P. (1997b). Derivational Minimalism. In: C. Retoré (ed.) *Logical Aspects of Computational Linguistics (LACL '96)*, LNAI Vol. 1328. Berlin, Heidelberg: Springer. 68–95.

Stabler, E. P. (1998). Acquiring Languages with Movement. *Syntax* 1. 72–97.

Stabler, E. P. (1999). Remnant Movement and Complexity. In: G. Bouma, G.-J. M. Kruijff, E. Hinrichs & R. T. Oehrle (eds.) *Constraints and Resources in Natural Language Syntax and Semantics*. Stanford, CA: CSLI Publications. 299–326.

Stabler, E. P. (2007). Sidewards Without Copying. In: S. Wintner (ed.) *Proceedings of FG 2006: The 11th Conference on Formal Grammar*. Stanford, CA: CSLI Publications. 157–170.

Stabler, E. P. (2008). Computational Perspectives on Minimalism. Unpublished manuscript, University of California, Los Angeles, CA.

von Stechow, A. (1991). Current Issues in the Theory of Focus. In: A. von Stechow & D. Wunderlich (eds.) *Semantik. Ein internationales Handbuch zeitgenössischer Forschung*. Berlin: de Gruyter. 804–825.

Stepanov, A., G. Fanselow & R. Vogel (eds.) (2004). *Minimality Effects in Syntax*. Berlin: Mouton de Gruyter.

Stepanov, A. & W.-T. D. Tsai (2008). Cartography and Licensing of Wh-Adjuncts. A Cross-Linguistic Perspective. *Natural Language and Linguistic Theory* 26. 589–638.

Szabolcsi, A. & F. Zwarts (1993). Weak Islands and an Algebraic Semantics for Scope Taking. *Natural Language Semantics* 1. 235–284.

Watanabe, A. (1992). Subjacency and S-Structure Movement of Wh-in-Situ. *Journal of East Asian Linguistics* 1. 255–291.

Zwart, J.-W. (2001). Syntactic and Phonological Verb Movement. *Syntax* 4. 34–62.

Appendix

Throughout we let $\neg Syn$ and Syn be a finite set of *non-syntactic features* and a finite set of *syntactic features*, respectively, in accordance with (F1)–(F3) below. We take

Feat to be the set $\neg Syn \cup Syn$.

(F1) $\neg$*Syn* is disjoint from *Syn* and partitioned into the sets *Phon* and *Sem*, a set of *phonetic features* and a set of *semantic*, or, *interpreted features*, respectively.

(F2) *Syn* is partitioned into eight sets:[27]

Base			*(basic) categories*	
M-Select	$=$	$\{\, \texttt{=x} \mid \texttt{x} \in Base \,\}$	*m(erge)-selectors*	
A-Select	$=$	$\{\, \approx\texttt{x} \mid \texttt{x} \in Base \,\}$	*a(djoin)-selectors*	
M-Licensors	$=$	$\{\, \texttt{+x}_{(\pi\sigma\iota)} \mid \texttt{x} \in Base \,\}$	*m(ove)-licensors*	[strong]
	$\cup$	$\{\, \texttt{+x}_{(\pi\sigma)} \mid \texttt{x} \in Base \,\}$		
	$\cup$	$\{\, \texttt{+x}_{(\pi\iota)} \mid \texttt{x} \in Base \,\}$		
	$\cup$	$\{\, \texttt{+x}_{(\pi)} \mid \texttt{x} \in Base \,\}$		
	$\cup$	$\{\, \texttt{+x}_{(\sigma\iota)} \mid \texttt{x} \in Base \,\}$	*m(ove)-licensors*	[weak]
	$\cup$	$\{\, \texttt{+x}_{(\sigma)} \mid \texttt{x} \in Base \,\}$		
	$\cup$	$\{\, \texttt{+x}_{(\iota)} \mid \texttt{x} \in Base \,\}$		
M-Licensees	$=$	$\{\, \texttt{-x} \mid \texttt{x} \in Base \,\}$	*m(ove)-licensees*	
S-Licensees	$=$	$\{\, \sim\texttt{x}_{(\pi\sigma\iota)} \mid \texttt{x} \in Base \,\}$	*s(cramble)-licensees*	
	$\cup$	$\{\, \sim\texttt{x}_{(\pi\sigma)} \mid \texttt{x} \in Base \,\}$		
	$\cup$	$\{\, \sim\texttt{x}_{(\pi\iota)} \mid \texttt{x} \in Base \,\}$		
	$\cup$	$\{\, \sim\texttt{x}_{(\pi)} \mid \texttt{x} \in Base \,\}$		
C-Licensors	$=$	$\{\, \triangledown\underline{\texttt{x}} \mid \texttt{x} \in Base \,\}$	*c(luster)-licensors*	[strong]
	$\cup$	$\{\, \triangledown\texttt{x} \mid \texttt{x} \in Base \,\}$	*c(luster)-licensors*	[weak]
C-Licensees	$=$	$\{\, \triangle\texttt{x} \mid \texttt{x} \in Base \,\}$	*c(luster)-licensees*	

(F3) *Base* includes at least the category `c`.

[27] Elements from *Syn* will usually be typeset in `typewriter font`.

Definition 1. An *expression (over Feat)*, also referred to as a *minimalist tree (over Feat)*, is a five-tuple $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau, <_\tau, label_\tau \rangle$ obeying (E1)–(E3).

(E1) $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$ is a finite, binary (ordered) tree defined in the usual sense: N_τ is the finite, non-empty set of *nodes*, and $\triangleleft^*_\tau$ and $\prec_\tau$ are the respective binary relations of *dominance* and *precedence* on N_τ.[28]

(E2) $<_\tau \subseteq N_\tau \times N_\tau$ is the asymmetric relation of *(immediate) projection* that holds for any two siblings, i.e., for each $x \in N_\tau$ different from the root of $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$ either $x <_\tau sibling_\tau(x)$ or $sibling_\tau(x) <_\tau x$ holds.[29]

(E3) $label_\tau$ is the *leaf-labeling function* from the set of leaves of $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$ into $Phon^* \times Syn^*\{\#\}Syn^* \times Sem^*$.[30]

We take $Exp(Feat)$ to denote the class of all expressions over *Feat*.

Let $\tau = \langle N_\tau, \triangleleft^*_\tau, \prec_\tau, <_\tau, label_\tau \rangle \in Exp(Feat)$.[31,32]

For each $x \in N_\tau$, the *head of x (in τ)*, denoted by $head_\tau(x)$, is the (unique) leaf of τ with $x \triangleleft^*_\tau head_\tau(x)$ such that each $y \in N_\tau$ on the path from x to $head_\tau(x)$ with $y \neq x$ projects over its sibling, i.e. $y <_\tau sibling_\tau(y)$. The *head of* τ is the head of τ's root. τ is said to be a *head* (or *simple*) if N_τ consists of exactly one node, otherwise τ is said to be a *non-head* (or *complex*).

An expression $\phi = \langle N_\phi, \triangleleft^*_\phi, \prec_\phi, <_\phi, label_\phi \rangle \in Exp(Feat)$ is a *subexpression of* τ in case $\langle N_\phi, \triangleleft^*_\phi, \prec_\phi \rangle$ is a subtree of $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$, $<_\phi = <_\tau \cap (N_\phi \times N_\phi)$, and

[28] Thus, $\triangleleft^*_\tau$ is the reflexive-transitive closure of $\triangleleft_\tau \subseteq N_\tau \times N_\tau$, the relation of *immediate dominance* on N_τ.

[29] $sibling_\tau(x)$ denotes the (unique) sibling of any given $x \in N_\tau$ different from the root of $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$. If $x <_\tau y$ for some $x, y \in N_\tau$ then x is said to *(immediately) project over* y.

[30] For each set M, M^* is the Kleene closure of M, including ε, the empty string. For any two sets of strings, M and N, MN is the product of M and N w.r.t. string concatenation. Further, # denotes a new symbol not appearing in *Feat*.

[31] Note that the leaf-labeling function $label_\tau$ can easily be extended to a total labeling function ℓ_τ from N_τ into $(Phon^* \times Syn^*\{\#\}Syn^* \times Sem^*) \cup \{<, >\}$, where < and > are two new distinct symbols: to each non-leaf $x \in N_\tau$ we can assign a label from $\{<, >\}$ by ℓ_τ such that $\ell_\tau(x) = <$ iff $y <_\tau z$ for $y, z \in N_\tau$ with $x \triangleleft_\tau y, z$, and $y \prec_\tau z$. In this sense a concrete $\tau \in Exp(Feat)$ is depictable in the way indicated in Figure 1.

[32] For the sake of convenience, we introduce the following convention: in case N_τ is a singleton set, i.e. N_τ is of the form $\{\nu\}$ with ν being the unique node of τ, τ will often be identified with $label_\tau(\nu)$, the label assigned to ν by $label_\tau$.

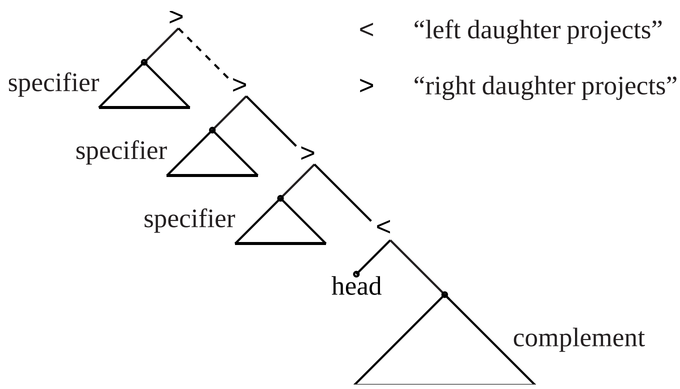


Figure 1: A typical minimalist expression

$\mathit{label}_\phi = \mathit{label}_\tau \restriction_{N_\phi}$.[33] Such a subexpression ϕ is a *maximal projection (in τ)* if its root is a node $x \in N_\tau$ such that x is the root of τ, or such that $\mathit{sibling}_\tau(x) <_\tau x$. $\mathit{MaxProj}(\tau)$ is the set of maximal projections in τ.

$\mathit{spec}_\tau \subseteq \mathit{MaxProj}(\tau) \times \mathit{MaxProj}(\tau)$ is the binary relation defined such that for all $\phi, \chi \in \mathit{MaxProj}(\tau)$ it holds that $\phi\, \mathit{spec}_\tau\, \chi$ iff both $r_\chi = \mathit{sibling}_\tau(x)$ and $x <_\tau r_\chi$ for some $x \in N_\tau$ with $r_\phi \triangleleft^+_\tau x \triangleleft^+_\tau \mathit{head}_\tau(r_\phi)$, where r_ϕ and r_χ are the roots of ϕ and χ, respectively. If $\phi\, \mathit{spec}_\tau\, \chi$ for some $\phi, \chi \in \mathit{MaxProj}(\tau)$ then χ is a *specifier of ϕ (in τ)*. $\mathit{Spec}(\tau)$ is the set $\{\phi \mid \tau\, \mathit{spec}_\tau\, \phi\}$.

Note that, if $\mathit{Spec}(\tau) \neq \emptyset$ then $\mathit{Spec}(\tau)$ is not necessarily a singleton set, but there is a unique specifier υ of τ, which we will refer to as the *highest specifier* of τ, such that the root of υ is immediately dominated by the root of τ.

A $\phi \in \mathit{MaxProj}(\tau)$ is said to *have*, or *display*, *(open) feature f* if the label assigned to ϕ's head by label_τ is of the form $\langle \delta, \beta \# f \beta', \delta' \rangle$ for some $f \in \mathit{Syn}$ and some $\beta, \beta', \delta, \delta' \in \mathit{Feat}^*$.

τ is *complete* if its head-label is in $\mathit{Phon}^* \times \mathit{Syn}^*\{\#\}\{\mathrm{c}\} \times \mathit{Sem}^*$, and each of its other leaf-labels is in $\mathit{Phon}^* \times \mathit{Syn}^*\{\#\} \times \mathit{Sem}^*$. Hence, a complete expression over *Feat* is an expression that has category c, and this instance of c is the only instance of a syntactic feature which is preceded by an instance of # within its local leaf-label, i.e. the leaf-label it appears in.

The *phonetic yield of τ*, denoted by $Y_{\mathit{Phon}}(\tau)$, is the string which results from concatenating in "left-to-right-manner" the first components of the labels assigned via

[33] For any set M, and for any function f from a set A into a set B, $f\restriction_M$ is the restriction of f to the domain $A \cap M$, i.e. $f\restriction_M$ is the function from $A \cap M$ into B such that $f\restriction_M(x) = f(x)$ for all $x \in A \cap M$.

$label_\tau$ to the leaves of $\langle N_\tau, \triangleleft_\tau^*, \prec_\tau \rangle$.

For any $\phi, \chi \in Exp(Feat)$, $[_<\phi, \chi]$ (respectively, $[_>\phi, \chi]$) denotes the complex expression $\psi = \langle N_\psi, \triangleleft_\psi^*, \prec_\psi, <_\psi, label_\psi \rangle \in Exp(Feat)$ for which ϕ and χ are those two subexpressions such that $r_\psi \triangleleft_\psi r_\phi$, $r_\psi \triangleleft_\psi r_\chi$ and $r_\phi \prec_\psi r_\chi$, and such that $r_\phi <_\psi r_\chi$ (respectively $r_\chi <_\psi r_\phi$), where r_ϕ, r_χ and r_ψ are the roots of ϕ, χ and ψ, respectively.

For any $\phi, \chi, \psi \in Exp(Feat)$ such that χ is a subexpression of ϕ, $\phi\{\chi/\psi\}$ is the expression which results from substituting ψ for χ in ϕ. We also simply write $\phi\{\chi/\langle\varepsilon, \varepsilon, \varepsilon\rangle\}$, if ψ is a single-noded tree labeled by $\langle\varepsilon, \varepsilon, \varepsilon\rangle$

Let $\pi \in Phon^*$, $\sigma \in Syn^*\{\#\}Syn^*$ and $\iota \in Sem^*$ such that $\langle \pi, \sigma, \iota \rangle$ is the head-label of some $\tau \in Exp(Feat)$. We denote

the	*phonetic projection (of τ)*	by	$phon(\tau)$,	
the	*syntactic projection (of τ)*	by	$syn(\tau)$,	
the	*semantic projection (of τ)*	by	$sem(\tau)$,	
the	*phonetic-syntactic projection (of τ)*	by	$phon\text{-}syn(\tau)$,	
the	*phonetic-semantic projection (of τ)*	by	$phon\text{-}sem(\tau)$,	and
the	*syntactic-semantic projection (of τ)*	by	$syn\text{-}sem(\tau)$,	

where *phon*, *syn*, *sem*, *phon-syn*, *phon-sem* and *syn-sem* are functions from $Exp(Feat)$ into $Exp(Feat)$ which are recursively defined as follows:

If τ is simple then $phon(\tau) = \langle \pi, \varepsilon, \varepsilon \rangle$, $syn(\tau) = \langle \varepsilon, \sigma, \varepsilon \rangle$, $sem(\tau) = \langle \varepsilon, \varepsilon, \iota \rangle$, $phon\text{-}syn(\tau) = \langle \pi, \sigma, \varepsilon \rangle$, $phon\text{-}sem(\tau) = \langle \pi, \varepsilon, \iota \rangle$ and $syn\text{-}sem(\tau) = \langle \varepsilon, \sigma, \iota \rangle$.[34]

If τ is complex then $\circ(\tau) = [_\bullet \circ(\phi), \circ(\chi)]$ with $\bullet \in \{<, >\}$ and $\phi, \chi \in Exp(Feat)$ such that $\tau = [_\bullet \phi, \chi]$, and with $\circ \in \{phon, syn, sem, phon\text{-}syn, phon\text{-}sem, syn\text{-}sem\}$.

In the following we write *MG* as a shorthand for *minimalist grammar*.

Definition 2. An MG is a 5-tuple of the form $\langle \neg Syn, Syn, Lex, \Omega, \mathrm{c} \rangle$ where Ω is the operator set consisting of the structure building functions *merge*, *adjoin*, *move*, *scramble* and *cluster* defined as in (me), (ad), (mo), (sc) and (cl) below, respectively, and where *Lex* is a *lexicon (over Feat)*, a finite set of simple expressions over *Feat*.

The operators from Ω build larger structure from given expressions by successively checking "from left to right" the instances of syntactic features appearing within the leaf-labels of the expressions involved. The symbol # serves to mark which feature instances have already been checked by the application of some structure building operation.

[34] Recall fn. 32.

(me) *merge* is a partial mapping from $\mathit{Exp}(\mathit{Feat}) \times \mathit{Exp}(\mathit{Feat})$ into $\mathit{Exp}(\mathit{Feat})$. For any $\phi, \chi \in \mathit{Exp}(\mathit{Feat})$, $\langle \phi, \chi \rangle$ is in $\mathit{Dom}(\mathit{merge})$ if for some category $\mathrm{x} \in \mathit{Base}$ and $\alpha, \alpha', \beta, \beta', \gamma, , \gamma', \delta, \delta' \in \mathit{Feat}^*$, conditions (me.i) and (me.ii) are fulfilled:[35]

(me.i) the head-label of ϕ is $\langle \gamma, \alpha \# {=}\mathrm{x} \alpha', \gamma' \rangle$, i.e. ϕ has m-selector $=\mathrm{x}$, and

(me.ii) the head-label of χ is $\langle \delta, \beta \# \mathrm{x} \beta', \delta' \rangle$, i.e. χ has category x.

Then,

(me.1) $\mathit{merge}(\phi, \chi) = [_< \phi', \chi']$ if ϕ is simple, and

(me.2) $\mathit{merge}(\phi, \chi) = [_> \chi', \phi']$ if ϕ is complex,

where ϕ' and χ' result from ϕ and χ, respectively, just by interchanging the instance of # and the instance of the feature directly following the instance of # within the second component of the respective head-label.

(ad) *adjoin* is a partial mapping from $\mathit{Exp}(\mathit{Feat}) \times \mathit{Exp}(\mathit{Feat})$ into the class $\mathit{Exp}(\mathit{Feat})$. A pair $\langle \phi, \chi \rangle$ with $\phi, \chi \in \mathit{Exp}(\mathit{Feat})$ belongs to $\mathit{Dom}(\mathit{adjoin})$ if for some category $\mathrm{x} \in \mathit{Base}$ and $\alpha, \alpha', \beta, \beta', \gamma, \gamma', \delta, \delta' \in \mathit{Feat}^*$, conditions (ad.i) and (ad.ii) are fulfilled:

(ad.i) the head-label of ϕ is $\langle \gamma, \alpha \# {\approx}\mathrm{x} \alpha', \gamma' \rangle$, i.e. ϕ has a-selector $\approx\mathrm{x}$, and

(ad.ii) the head-label of χ is $\langle \delta, \beta \# \mathrm{x} \beta', \delta' \rangle$, i.e. χ has category x.

Then, $\mathit{adjoin}(\phi, \chi) = [_> \phi', \chi]$,

where ϕ' results from ϕ by interchanging the instances of # and $\approx\mathrm{x}$, the latter directly following the former in the second component of the head-label of ϕ.

(mo) *move* is a partial mapping from $\mathit{Exp}(\mathit{Feat})$ into $\mathit{Exp}(\mathit{Feat})$. A $\phi \in \mathit{Exp}(\mathit{Feat})$ is in $\mathit{Dom}(\mathit{move})$ if for some $-\mathrm{x} \in$ *M-Licensees* and $\alpha, \alpha', \gamma, \gamma' \in \mathit{Feat}^*$, (mo.i), (mo.ii) and (mo.smc) are true:

(mo.i) the head-label of ϕ is $\langle \gamma, \alpha \# x \alpha', \gamma' \rangle$ and y is an m-licensor from the set $\{ +\mathrm{x}_{(\pi\sigma\iota)}, +\mathrm{x}_{(\pi\sigma)}, +\mathrm{x}_{(\pi\iota)}, +\mathrm{x}_{(\pi)}, +\mathrm{x}_{(\sigma\iota)}, +\mathrm{x}_{(\sigma)}, +\mathrm{x}_{(\iota)} \}$, i.e. ϕ displays m-licensor y,

[35] For a partial function f from a class A into a class B, $\mathit{Dom}(f)$ is the domain of f, i.e., the class of all $x \in A$ for which $f(x)$ is defined.

(mo.ii) there is a $\chi \in \mathit{MaxProj}(\phi)$ with head-label $\langle \delta, \beta \texttt{\#-x} \beta', \delta' \rangle$ for some $\beta, \beta', \delta, \delta' \in \mathit{Feat}^*$, i.e. there is a $\chi \in \mathit{MaxProj}(\phi)$ displaying $\texttt{-x}$, and

(mo.smc) the existing $\chi \in \mathit{MaxProj}(\phi)$ from (mo.ii) is unique, i.e. there is exactly one $\chi \in \mathit{MaxProj}(\phi)$ displaying $\texttt{-x}$.

(mo-$\pi\sigma\iota$) If $y = \texttt{+x}_{(\pi\sigma\iota)}$ then $\mathit{move}(\phi) = [_{>}\chi', \phi\{\chi/\langle\varepsilon,\varepsilon,\varepsilon\rangle\}']$,

where $\phi\{\chi/\langle\varepsilon,\varepsilon,\varepsilon\rangle\}'$ results from $\phi\{\chi/\langle\varepsilon,\varepsilon,\varepsilon\rangle\}$ by interchanging the instance of # and the instance of $\texttt{+x}_{(\pi\sigma\iota)}$ directly following it within the second component of the head-label of $\phi\{\chi/\langle\varepsilon,\varepsilon,\varepsilon\rangle\}'$. χ' arises from χ by interchanging the instance of # and the instance of $\texttt{-x}$ immediately to its right within the second component of the head-label of χ.

(mo-$\pi\sigma$) If $y = \texttt{+x}_{(\pi\sigma)}$ then $\mathit{move}(\phi) = [_{>}\mathit{phon\text{-}syn}(\chi)', \phi\{\chi/\mathit{sem}(\chi)\}']$,

where $\phi\{\chi/\mathit{sem}(\chi)\}'$ results from $\phi\{\chi/\mathit{sem}(\chi)\}$ by interchanging the instance of # and the instance of $\texttt{+x}_{(\pi\sigma)}$ directly following it within the second component of the second component of the head-label of $\phi\{\chi/\mathit{sem}(\chi)\}$. $\mathit{phon\text{-}syn}(\chi)'$ arises from $\mathit{phon\text{-}syn}(\chi)$ by interchanging the instance of # and the instance of $\texttt{-x}$ immediately to its right within the second component of the head-label of $\mathit{phon\text{-}syn}(\chi)$.

(mo-$\pi\iota$) If $y = \texttt{+x}_{(\pi\iota)}$ then $\mathit{move}(\phi) = [_{>}\mathit{phon\text{-}sem}(\chi), \phi\{\chi/\psi\}']$,

where ψ arises from $\mathit{syn}(\chi)$ by interchanging the instance of # and the instance of $\texttt{-x}$ immediately to its right within the second component of the head-label of $\mathit{syn}(\chi)$, and where $\phi\{\chi/\psi\}'$ results from $\phi\{\chi/\psi\}$ by interchanging the instance of # and the instance of $\texttt{+x}_{(\pi\iota)}$ directly following it within the second component of the second component of the head-label of $\phi\{\chi/\psi\}$.

(mo-π) If $y = \texttt{+x}_{(\pi)}$ then $\mathit{move}(\phi) = [_{>}\mathit{phon}(\chi), \phi\{\chi/\psi\}']$,

where ψ arises from $\mathit{syn\text{-}sem}(\chi)$ by interchanging the instance of # and the instance of $\texttt{-x}$ immediately to its right within the second component of the head-label of $\mathit{syn}(\chi)$, and where $\phi\{\chi/\psi\}'$ results from $\phi\{\chi/\psi\}$ by interchanging the instance of # and the instance of $\texttt{+x}_{(\pi)}$ directly following it within the second component of the second component of the head-label of $\phi\{\chi/\psi\}$.

(mo-$\sigma\iota$) If $y = +\mathrm{x}_{(\sigma\iota)}$ then $move(\phi) = [_{>}syn\text{-}sem(\chi)', \phi\{\chi/phon(\chi)\}']$,

where $\phi\{\chi/phon(\chi)\}'$ results from $\phi\{\chi/phon(\chi)\}$ by interchanging the instance of # and the instance of +x directly following it within the second component of the second component of the head-label of $\phi\{\chi/phon(\chi)\}$. $syn\text{-}sem(\chi)'$ arises from $syn\text{-}sem(\chi)$ by interchanging the instance of # and the instance of −x immediately to its right within the second component of the head-label of $syn\text{-}sem(\chi)$.

(mo-σ) If $y = +\mathrm{x}_{(\sigma)}$ then $move(\phi) = \phi\{\chi/\psi\}'$,

where ψ arises from χ by interchanging the instance of # and the instance of −x immediately to its right within the second component of the head-label of χ, and where $\phi\{\chi/\psi\}'$ results from $\phi\{\chi/\psi\}$ by interchanging the instance of # and the instance of $+\mathrm{x}_{(\sigma)}$ directly following it within the second component of the second component of the head-label of $\phi\{\chi/\psi\}$.

(mo-ι) If $y = +\mathrm{x}_{(\iota)}$ then $move(\phi) = [_{>}sem(\chi)', \phi\{\chi/phon\text{-}syn(\chi)\}']$,

where $\phi\{\chi/phon\text{-}syn(\chi)\}'$ results from $\phi\{\chi/phon\text{-}syn(\chi)\}$ by interchanging the instance of # and the instance of $+\mathrm{x}_{(\iota)}$ directly following it within the second component of the second component of the head-label of $\phi\{\chi/phon\text{-}syn(\chi)\}$. $syn\text{-}sem(\chi)'$ arises from $syn\text{-}sem(\chi)$ by interchanging the instance of # and the instance of −x immediately to its right within the second component of the head-label of $syn\text{-}sem(\chi)$.

(sc) The function *scramble* maps partially from *Exp*(*Feat*) into *Exp*(*Feat*). An expression $\phi \in Exp(Feat)$ is in $Dom(scramble)$ if for some $\mathrm{x} \in Base$ and $\alpha, \alpha', \gamma, \gamma' \in Feat^*$, (sc.i), (sc.ii) and (sc.smc) are true:

(sc.i) the head-label of ϕ is $\langle \gamma, \alpha\#\mathrm{x}\alpha', \gamma' \rangle$, i.e. ϕ has category x,

(sc.ii) there is a $\chi \in MaxProj(\phi)$ with head-label $\langle \delta, \beta\#y\beta', \delta' \rangle$ for some $\beta, \beta', \delta, \delta' \in Feat^*$, and some $y \in \{\sim\mathrm{x}_{(\pi\sigma\iota)}, \sim\mathrm{x}_{(\pi\sigma)}, \sim\mathrm{x}_{(\pi\iota)}, \sim\mathrm{x}_{(\pi)}\}$, i.e. ϕ displays s-licensee y, and

(sc.smc) the existing $\chi \in MaxProj(\phi)$ from (sc.ii) is unique, i.e. there is exactly one $\chi \in MaxProj(\phi)$ displaying y.

(sc-$\pi\sigma\iota$) If $y = \sim\mathrm{x}_{(\pi\sigma\iota)}$ then, $scramble(\phi) = [_{>}\chi', \phi\{\chi/\langle\varepsilon, \varepsilon, \varepsilon\rangle\}]$, where $\chi' \in \mathit{Exp}(\mathit{Feat})$ arises from χ by interchanging the instance of # and the instance of $\sim\mathtt{x}_{(\pi\sigma\iota)}$ immediately to its right within the second component of the head-label of χ.

(sc-$\pi\sigma$) If $y = \sim\mathtt{x}_{(\pi\sigma)}$ then, $\mathit{scramble}(\phi) = [_{>}\mathit{phon\text{-}syn}(\chi)', \phi\{\chi/\mathit{sem}(\chi)\}]$,

where $\mathit{phon\text{-}syn}(\chi)' \in \mathit{Exp}(\mathit{Feat})$ arises from $\mathit{phon\text{-}syn}(\chi)$ by interchanging the instance of # and the instance of $\sim\mathtt{x}_{(\pi\sigma)}$ immediately to its right within the second component of the head-label of $\mathit{phon\text{-}syn}(\chi)$.

(sc-$\pi\iota$) If $y = \sim\mathtt{x}_{(\pi\iota)}$ then, $\mathit{scramble}(\phi) = [_{>}\mathit{phon\text{-}sem}(\chi), \phi\{\chi/\psi\}]$,

where $\psi \in \mathit{Exp}(\mathit{Feat})$ arises from $\mathit{syn}(\chi)$ by interchanging the instance of # and the instance of $\sim\mathtt{x}_{(\pi\iota)}$ immediately to its right within the second component of the head-label of $\mathit{syn}(\chi)$.

(sc-p) If $y = \sim\mathtt{x}_{(\pi)}$ then, $\mathit{scramble}(\phi) = [_{>}\mathit{phon}(\chi), \phi\{\chi/\psi\}]$,

where $\psi \in \mathit{Exp}(\mathit{Feat})$ arises from $\mathit{syn\text{-}sem}(\chi)$ by interchanging the instance of # and the instance of $\sim\mathtt{x}_{(\pi)}$ immediately to its right within the second component of the head-label of $\mathit{syn\text{-}sem}(\chi)$.

(cl) The operator *cluster* is a partial mapping from $\mathit{Exp}(\mathit{Feat})$ to $\mathit{Exp}(\mathit{Feat})$. An expression $\phi \in \mathit{Exp}(\mathit{Feat})$ is in $\mathit{Dom}(\mathit{cluster})$ if there are a c-licensee $\triangle\mathtt{x}$ and $\alpha, \alpha', \gamma, \gamma' \in \mathit{Feat}^*$ such that (cl.i), (cl.ii) and (cl.smc) are true:

(cl.i) there is a $\chi \in \mathit{MaxProj}(\phi)$ such that χ is the highest specifier of ϕ, and the head-label of χ is $\langle\gamma, \alpha\#x\alpha', \gamma'\rangle$ with $y \in \{\triangledown\underline{\mathtt{x}}, \triangledown\mathtt{x}\}$, i.e. ϕ displays the corresponding c-licensor x,

(cl.ii) there is a $\psi \in \mathit{MaxProj}(\phi)$ with head-label $\beta\#\triangle\mathtt{x}\beta'$ for some $\beta, \beta' \in \mathit{Feat}^*$, i.e. $\psi \in \mathit{MaxProj}(\phi)$ exists displaying $\triangle\mathtt{x}$.

(cl.smc) the existing $\psi \in \mathit{MaxProj}(\phi)$ from (cl.ii) is unique, i.e. there is exactly one $\psi \in \mathit{MaxProj}(\phi)$ displaying $\triangle\mathtt{x}$.

(cl-$\pi\sigma\iota$) If $y = \triangledown\underline{\mathtt{x}}$ then $\mathit{cluster}(\phi) = \phi\{\psi/\langle\varepsilon,\varepsilon,\varepsilon\rangle, \chi/[_{<}\chi', \psi']\}$, where χ' results from χ by interchanging the instances of # and $\triangledown\underline{\mathrm{x}}$, the latter directly following the former in the second component of the head-label of χ, while ψ' results from ψ by interchanging the instances of # and $\triangle\mathrm{x}$, the latter directly following the former in the second component of the head-label of ψ.[36]

(cl-ι) If $y = \triangledown\mathrm{x}$ then $\mathit{cluster}(\phi) = \phi\{\psi/\mathit{phon\text{-}syn}(\psi)', \chi/[_<\chi', \mathit{sem}(\psi)]\}$,

where χ' results from χ by interchanging the instances of # and $\triangledown\mathrm{x}$, the latter directly following the former in the second component of the head-label of χ, while *phon-syn*$(\psi)'$ results from *phon-syn*(ψ) by interchanging the instances of # and $\triangle\mathrm{x}$, the latter directly following the former in the second component of the head-label of *phon-syn*(ψ).

The *closure of* G, $\mathit{CL}(G)$, is the set $\bigcup_{k\in\mathbb{N}} \mathit{CL}^k(G)$, where $\mathit{CL}^0(G) = \mathit{Lex}$, and for $k \in \mathbb{N}$, $\mathit{CL}^{k+1}(G) \subseteq \mathit{Exp}(\mathit{Feat})$ is recursively defined as the set

$$\begin{aligned} \mathit{CL}^k(G) \;&\cup\; \{\mathit{merge}(\phi,\chi) \mid \langle\phi,\chi\rangle \in \mathit{Dom}(\mathit{merge}) \cap \mathit{CL}^k(G) \times \mathit{CL}^k(G)\} \\ &\cup\; \{\mathit{adjoin}(\phi,\chi) \mid \langle\phi,\chi\rangle \in \mathit{Dom}(\mathit{adjoin}) \cap \mathit{CL}^k(G) \times \mathit{CL}^k(G)\} \\ &\cup\; \{\mathit{move}(\phi) \mid \phi \in \mathit{Dom}(\mathit{move}) \cap \mathit{CL}^k(G)\} \\ &\cup\; \{\mathit{scramble}(\phi) \mid \phi \in \mathit{Dom}(\mathit{scramble}) \cap \mathit{CL}^k(G)\} \\ &\cup\; \{\mathit{cluster}(\phi) \mid \phi \in \mathit{Dom}(\mathit{cluster}) \cap \mathit{CL}^k(G)\} \end{aligned}$$

The set $\{\tau \mid \tau \in \mathit{CL}(G) \text{ and } \tau \text{ complete}\}$, denoted by $T(G)$, is the *minimalist tree language derivable by* G. The set $\{Y_{\mathit{Phon}}(\tau) \mid \tau \in T(G)\}$, denoted by $L(G)$, is the *minimalist (string) language derivable by* G.

[36] As long as the SMC is obeyed, a proof showing that at least the weak generative capacity is unaffected seems to be straightforward by employing the "usual" reduction methods (cf. Michaelis 2001a).