In Defense of Generalized *Wh*-Clustering

Hans-Martin Gärtner   Jens Michaelis
HAS-RIL Budapest   Bielefeld University

1   Introduction

Following earlier work by i.a. Rudin (1988) and Saito (1994), Grewendorf (2001) and Sabel (2001) argue in favor of treating multiple-*wh*-interrogatives in languages like Bulgarian, Japanese, and German in terms of *wh-clustering*. Abstractly, the Bulgarian counterpart of English (1) would be analyzed as in (2).

(1)   *Who gave what to whom*

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

This involves two applications of *wh*-clustering and one application of *wh*-movement: first *to whom* attaches to the right of *what*. Then the cluster [*what* [*to whom*]] attaches to the right of *who*. Finally, standard *wh*-movement places the cluster [*who* [*what* [*to whom*]]] in Spec,CP. Japanese and German are taken to differ only minimally from Bulgarian as the respective analyses in (3) and (4) show.

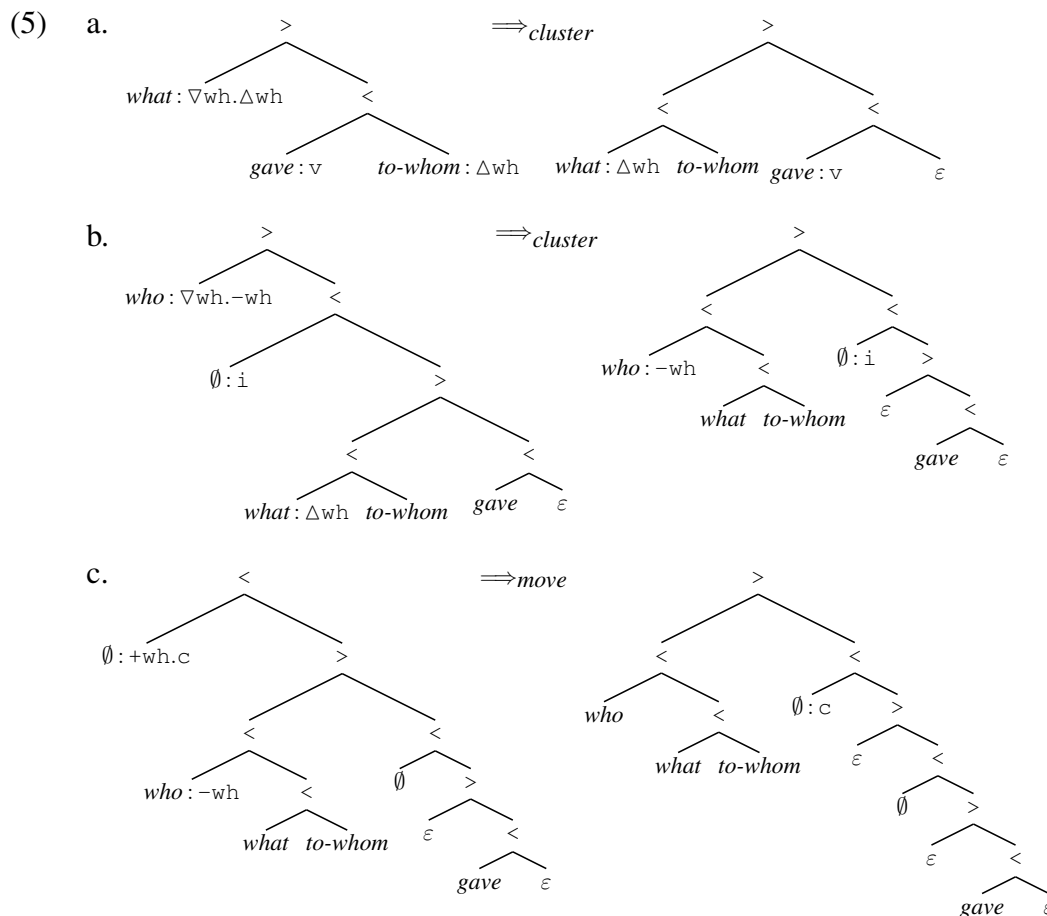(3) $[_{CP} s_k [_{C'} [_{IP} [who\, s_j]_k [_{I'} [_{VP} [what\, s_i]_j [_{V'} [to whom]_i\, gave]] I°]] C°]]$

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

Whereas in Bulgarian all operations are overt, as indicated by the familiar *constituent+trace* dependencies ($XP_n \ldots t_n$), all operations are covert in Japanese, where covert displacement is depicted by what we call *shadow+constituent* dependencies ($s_m \ldots XP_m$). In German, *wh*-clustering is covert and *wh*-movement overt.

In a recent paper (Gärtner and Michaelis 2010), we argue that this kind of clustering approach to multiple-*wh*-interrogatives should be adopted within the framework of Stablerian *Minimalist Grammars (MGs)* (Stabler 1997, 1998, 2008) in order to preserve their *mild context-sensitivity* (Joshi 1985; Joshi, Vijay-Shanker and Weir 1991; Michaelis 2001a, 2001b). Very roughly, the point is that clustering allows checking resources to be strictly bounded. Consider (5), which illustrates the three essential steps of an MG-derivation of (2).

Wh-clustering is triggered by ▽wh on the attracting *wh*-phrase and △wh on the attracted one. *Wh*-movement is triggered by +wh on C° and −wh on the attracted *wh*-phrase. Crucially, at no stage in the derivation does more than one attractable feature of the same kind (△wh or −wh) have to be "active," i.e., have to be leftmost in the string of features on (the head of) a *wh*-phrase. This is in contrast to the situation in "standard" non-clustering theories, where all three *wh*-phrases would have active −wh and get moved to Spec,CP individually. For further detail, i.e.,

an MG-implementation of multiple-*wh*-interrogatives and extensive discussion, we refer the reader to Gärtner and Michaelis 2010.

(5) a.
$$\Longrightarrow_{cluster}$$

what : ▽wh.△wh   *gave* : v   *to-whom* : △wh   *what* : △wh *to-whom* *gave* : v   ε

b.
$$\Longrightarrow_{cluster}$$

who : ▽wh.−wh   ∅ : i   *what* : △wh *to-whom* *gave* ε   *who* : −wh *what* *to-whom* ∅ : i ε *gave* ε

c.
$$\Longrightarrow_{move}$$

∅ : +wh.c   *who* : −wh *what* *to-whom* ∅ ε *gave* ε   *who* *what* *to-whom* ∅ : c ε ∅ ε *gave* ε

Now, as pointed out in Gärtner and Michaelis 2010, adopting the MG-formalism is a motivation for trying to apply the tool of *wh*-clustering to multiple-*wh*-interrogatives in all languages. This, however, is not what Grewendorf (2001) and Sabel (2001) envisage. Building on work by Cheng (1991), they argue 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. Yet, although we can see the initial appeal of limiting *wh*-clustering to languages displaying this kind of *indefinite-interrogative affinity* (cf. Bhat 2000) [henceforth *IIA-languages*], we believe that – in line with an MG-perspective – a defense of generalized *wh*-clustering can be made. For the sake of concreteness, this defense will focus on the most detailed and explicit study of *wh*-clustering so far, i.e., the one by Grewendorf (2001). In particular, we will show that the main empirical domain of application for *wh*-clustering discussed there, namely, additional-*wh* effects, does not warrant any interesting correlation with IIA-languages (Section 4). Before doing so we will briefly point out a potential problem *wh*-determiners in IIA-languages pose

for developing the mechanism of *wh*-clustering in terms of a null determiner (Section 2). Also, we will have a look at simple superiority effects, which Grewendorf (2001) uses to distinguish clustering from non-clustering languages (Section 3).

2   (Null) *Wh*-Determiners

Core ingredient of the *wh*-clustering mechanism developed by Grewendorf (2001) is the null *wh*-determiner postulated by Cheng (1991), which he takes to occur in all IIA-languages (and only there). Examples are languages in which interrogative and (pure) indefinite pronouns are (largely) form-identical – German *wer* ('who'/'someone'), *was* ('what'/'something') etc., and Japanese *dare* ('who'/'someone'), *nani* ('what'/'something') etc. – and languages in which interrogative and (pure) indefinite pronouns are (largely) in transparent morphological relation – German *wer* ('who') vs. *irgendwer* ('someone'), *was* ('what') vs. *irgendwas* ('something') etc., and Bulgarian *koj* ('who') vs. *njakoj* ('someone'), *kude* ('where') vs. *njakude* ('somewhere') etc. The interrogative pronouns of IIA-languages are thus assumed to have the internal structure exemplified in (6) for German *wer* (Grewendorf 2001:95).

(6) $[_{\text{DP}} [_{\text{D}'} [_{\text{D}\circ} \emptyset_{wh}] [_{\text{NP}} wer]]]$

In MG-terms, only $[_{\text{D}\circ} \emptyset_{wh}]$ can bear clustering features like $\triangledown$wh and $\triangle$wh. Equally, the feature attracted under *wh*-movement, $-$wh, would be located there. Non-IIA-languages like English (*who* vs. *someone*, *what* vs. *something* etc.) and French (*qui* vs. *quelqu'un*, *quoi* vs. *quelque chose* etc.) are taken to lack $[_{\text{D}\circ} \emptyset_{wh}]$ and thus lack the means for *wh*-clustering. Their interrogative pronouns are assumed to possess (counterparts of) $-$wh as "inherent lexical property" (Grewendorf 2001:96).[1]

A potential obstacle to the analysis in (6) is the existence in IIA-languages of *wh*-determiner counterparts of English *which*. German *welch*, for example, can be used both in interrogative as well as pure indefinite phrases. But if at the same time $[_{\text{D}\circ} \emptyset_{wh}]$ is "the null counterpart of *which* in English" (Cheng 1991:85), *welch* and $[_{\text{D}\circ} \emptyset_{wh}]$ compete for the same structural position and should thus be in complementary distribution. It seems therefore that elements like *welch* must be licensed inherently. Consequently, if clustering features can be introduced lexically too, the dichotomy between IIA-languages and non-IIA-languages breaks down and the morpholexical argument against generalizing *wh*-clustering to both language types is voided.

---

[1]It would perhaps make sense to hypothesize that languages lacking multiple-*wh*-interrogatives also lack the means for *wh*-clustering, i.e., that such languages are non-IIA-languages. Familiar cases like Irish and Italian would confirm that. However, IIA-languages without multiple-*wh*-interrogatives have been reported in the literature. Examples are the Australian languages Diyari, Martuthunira, and Panyjima (Cheng 1991:110) and Passamaquoddy (Bruening 2007:140).

Opponents of generalized *wh*-clustering could, however, follow Gallmann (1997) and Leu (2008) in assuming that *welch* actually requires a phrasal analysis. One way of spelling this out quite closely in spirit to Leu (2008) is given in (7), where *welch* is treated as an AP introduced by a functional projection between D° and NP.

(7)   [$_{\text{DP}}$ [$_{\text{D}'}$ [$_{\text{D°}}$ $\emptyset_{wh}$] [$_{\text{FP}}$ [$_{\text{AP}}$ *welches*] [$_{\text{F}'}$ F° [$_{\text{NP}}$ *Buch*]]]]]] ('which book')

One would have to equip [$_{\text{D°}}$ $\emptyset_{wh}$] with an additional licensing feature, e.g. +w, that checks one-to-one against a counterpart, −w, on *wh*-pronouns in IIA-languages.[2] It remains to be shown that this kind of analysis can be upheld and extended to *wh*-determiners in other IIA-languages such as *melyik* in Hungarian, *który* in Polish, and *care* in Romanian (cf. Haspelmath 1997).[3]

3 Superiority

Grewendorf (2001:110ff) provides a clustering-based account of the familiar superiority contrast between English and German in (8).

(8)   a.*What did who say?*
       b. *Was sagte wer?*

The IIA-language German possesses the option of (covertly) clustering a *wh*-subject located in its base-position in Spec,vP onto a *wh*-object at the point where the latter has moved to its Case-licensing position (outer Spec,vP or Spec,AgrOP). The cluster [*was* s$_i$] later moves on to Spec,CP leaving *wer*$_i$ inside IP. By contrast, the non-IIA-language English lacks the means of cluster formation so that a derivation parallel to the German one is ruled out. In addition, due to the MLC no other way for *what* to cross *who*, once the latter has moved to Spec,IP, is taken to be possible (Grewendorf 2001:112).

The account just given is immediately challenged by superiority violations in non-IIA-languages, such as the one in (9) from Spanish (Fanselow 2004:93).

(9)   *Qué dijó quién?*

---

[2]The mechanism is related to the treatment of negative indefinites in German by Penka and von Stechow (2001).

[3]Further interesting questions are raised by the assumption (Grewendorf 2001:94, 101, 113, Section 5) that (nonreferential) *wh*-adjuncts lack [$_{\text{D°}}$ $\emptyset_{wh}$]. Thus, it has to be clarified why, e.g., the German *wh*-adverb *wie* ('how') can – like other *wh*-pronouns – be transformed into an indefinite by prefixing *irgend-* (*irgendwie* 'somehow'), a process taken by Cheng (1991:Section 3.2.2) to be indicative of the presence of [$_{\text{D°}}$ $\emptyset_{wh}$]. Grewendorf (2001) motivates this further by the frequent inability of *wh*-adjuncts to serve as cluster-attractors. In MG-terms, this means that they cannot bear ▽wh. Potential counterexamples are dealt with on the basis of an additional licensing mechanism for *wh*-phrases involving head movement (Grewendorf 2001:Section 5). Going into detail here would require comparison of MG-checking theory and the feature licensing mechanisms advocated by Grewendorf (2001) building on the *probe-goal* framework introduced by Chomsky (2000, 2001).

Equally, an extra assumption has to be made to explain the superiority effect in IIA-languages like Bulgarian, shown in (10b) (Rudin 1988:473).

(10) *Kogo koj* vižda?
  lit. 'Whom did who see?'

Grewendorf (2001:112, fn.40) suggests that *wh*-subjects may not undergo overt cluster formation before checking their Case in Spec,IP. This is what distinguishes Bulgarian from German.[4] But, if this is so, the following contrast from Basque, another IIA-language, becomes worrisome (Reglero 2003:194f).

(11) a. *Nork erosi du zer?*
    lit. 'Who has bought what?'
  b.**Zer erosi du nork?*
    lit. 'What has who bought?'

Clearly, covert *wh*-clustering is called for here. Nevertheless, superiority must be obeyed.

 We cannot pursue this issue further since too many additional background assumptions would have to be spelled out in order to make the above arguments more conclusive. However, we believe that no elegant account of simple superiority effects has emanated from the approach restricting *wh*-clustering to IIA-languages. We take this to indicate that the account of superiority is independent and thus equally compatible with generalized *wh*-clustering.

4 Additional-*Wh* Effects

Grewendorf (2001) further motivates the mechanism of *wh*-clustering by showing how to handle *additional-wh effects* (in a broad sense) in Japanese and Bulgarian. Thus consider the contrast in (12) from Japanese (Grewendorf 2001:100; Saito 1994:205).[5]

(12) a.**John-wa* [PP [IP *Mary-ga sono hon-o naze katta*] *kara*]
   John-TOP Mary-NOM that book-ACC why bought since
   okotteru no?
   angry Q
   'Give me the reason such that John is angry because Mary bought that book for that reason'
  b.?*John-wa* [PP [IP *Mary-ga nani-o naze katta*] *kara*] *okotteru no?*
   John-TOP Mary-NOM what-ACC why bought since angry Q
   'Give me things and reasons such that John is angry because Mary bought the respective thing for the respective reason'

---

[4]This account for Bulgarian would seem to carry over to the IIA-language Tlingit (Cable 2007).
[5]Maki (2002) provides a critical assessment of the *wh*-clustering approach to Japanese.

The difference can be explained in terms of a difference between adjunct and argument extraction from an island. While in (12a) the adjunct *naze* would have to undergo *wh*-movement by itself, in (12b) *naze* could possess △wh and cluster onto *nani-o* :: ▽wh.−wh. The cluster is then extracted on the basis of feature −wh on *nani-o* and that extraction can count as argument extraction. Hence the higher acceptability of (12b).

A slightly different kind of contrast is reported for Bulgarian (Grewendorf 2001:91; Richards 2001:50).

(13) 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 *koja kniga* ('which book') from a complex NP via *wh*-movement, i.e., driven by −wh, is unacceptable, as (13a) shows. However, clustering *koja kniga* onto *koj senator* ('which senator') in the matrix clause, i.e., movement driven by △wh, is able to improve acceptability, as (13b) shows.

In a footnote (Grewendorf 2001:103, fn.27), it is pointed out that the perhaps most widely discussed kind of additional-*wh* effect, namely, the one involved in the contrast in (14) (cf. Kayne 1983), cannot be dealt with in terms of *wh*-clustering, due to the non-IIA nature of English.

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

Clearly, (14) would have lent itself to an account in terms of the difference between *who* being crossed by just *what* vs. being crossed by [*what* s$_i$], the shadow s$_i$ being created by covertly clustering *to whom* onto *what*.

Likewise, the often discussed contrast in (15) (Lasnik and Saito 1992:118f) could be considered due to a difference between crossing of *who* by *what* :: −wh.▽wh vs. *what* :: −wh.[6]

(15) a.*Bill wondered what who bought
      b. Who wondered what who bought?

To the extent that more non-IIA-languages show effects of this kind, not generalizing *wh*-clustering in order to account for all (or most) additional-*wh* effects feels like a missed opportunity. Unfortunately, additional-*wh* effects have not been studied to

---

[6]Alternatively, if for some reason the earliest point in the derivation at which *what* could cluster-attract *who* were the point when *what* has reached Spec,CP, the independent principle argued for by Grewendorf (2001:94) according to which clustering cannot take place in Spec,CP would prevent that. Thus, *who* could not check △wh in (15a). In (15b), on the other hand, the additional *wh*-phrase in the matrix could attract the lower *who* before the resulting cluster moves to matrix Spec,CP.

any greater detail across languages. A contrast comparable to the one in (15) is the *wh-triangle effect* reported by Dayal (2006:316f) for the non-IIA-language Hindi (Haspelmath 1997:284f).

(16) a. *jaun jaantaa hai ki   meri-ne kyaa kharidaa*
      John knows   that Mary   what bought
      'John knows what Mary bought'
      NOT: 'What does John know that Mary bought?'
   b. *kaun jaantaa hai ki   meri-ne kahaaN kyaa kharidaa*
      who knows   that Mary   where   what bought
      'Give me persons and things such that the respective person knows where Mary bought the respective thing'

In (16a), *kyaa* :: −wh would have to be extracted while in (16b) *kyaa* :: △wh is involved in long dependency formation. This, of course, could only be the basis of an account if *wh*-clustering were generalized to be available in non-IIA-languages like Hindi as well.[7]

Let's return to the clustering account of (13), according to which movement driven by △wh circumvents an island violation incurred if movement is driven by −wh. The exact same account can, of course, be given for the familiar difference between overt vs. covert *wh*-movement in languages like German. Consider (17).

(17) a.*Was$_i$ mag Franz die Bücher [die t$_i$ kritisieren]*
      lit. 'What does Franz like the books that criticize?'
   b. *Wer mag die Bücher [die was kritisieren]?*
      'Who likes the books that criticize what?'

German is an IIA-language and Grewendorf (2001) discusses *wh*-clustering in German at length (cf. also Section 3 above). Thus, it would feel like a missed opportunity if the account provided for (13) were not available for (17). Movement of *was* driven by −wh in (17a) is ruled out, while *wh*-clustering by *was* :: △wh onto *wer* in (17b) is fine. What's more, as the translations of (17) show, exactly the same contrast holds in English (as it does in many other languages). It would,

---

[7]Persian is an interesting case to look at further here. According to Grewendorf (2001:105, fn.30) Persian may have overt *wh*-clustering but covert (or optionally overt) *wh*-movement, i.e., be of the fourth type of languages predicted by his system (cf. Section 1). The study by Lotfi (2003) appears to confirm that. In particular, Lotfi (2003:184, fn.5) makes explicit appeal to *wh*-clustering in the analysis of a case of fronted adjacent *wh*-phrases. However, Persian seems to have to be counted among the non-IIA-languages if one looks at the paradigms provided by Haspelmath (1997:282ff). It would also count as non-IIA-language if *indefinite-interrogative affinity* is taken in the "narrow" technical sense discussed by Grewendorf (2001). As we've seen in Section 2, the idea is that *wh*-words can be turned into pure indefinites by adding an affix (which may be null). Lotfi (2003:Section 3.2) shows that Persian possesses an inverse process. Pure indefinites like *kas-i* ('someone') or *chiz-i* ('something') can function as *wh*-pronouns if prefixed by *che*. Assuming a broader definition and a way to sort out the DP-internal structure, one may thus consider Persian an IIA-language after all.

thus, feel like an even greater missed opportunity not to generalize *wh*-clustering to non-IIA-languages like English and seek a unified account of such cases.

5 Conclusion

Given the absence of any clear correlation between the IIA-status of languages and their behavior with respect to superiority and additional-*wh* effects, we conclude that these empirical domains constitute no obstacle to generalizing the *wh*-clustering approach to multiple-*wh*-interrogatives.

References

Bhat, D. N. S. 2000. The indefinite-interrogative puzzle. *Linguistic Typology* 4.365–400.

Boeckx, Cedric, and Kleanthes Grohmann (eds.) 2003. *Multiple Wh-Fronting*. Amsterdam: John Benjamins.

Bruening, Benjamin. 2007. *Wh*-in-situ does not correlate with *wh*-indefinites or question particles. *Linguistic Inquiry* 38.139–166.

Cable, Seth, 2007. *The Grammar of Q. Q-Particles and the Nature of Wh-Fronting, as Revealed by the Wh-Questions of Tlingit*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Cheng, Lisa L. S., 1991. *On the Typology of WH-Questions*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Chomsky, Noam. 2000. Minimalist inquiries. The framework. In *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by R. Martin, D. Michaels, and J. Uriagereka, 89–155. Cambridge, MA: MIT Press.

——. 2001. Derivation by phase. In *Ken Hale. A Life in Language*, ed. by M. Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Dayal, Veneeta. 2006. Multiple-wh-questions. In *The Blackwell Companion to Syntax*, ed. by M. Everaert and H. van Riemsdijk, 275–326, Vol. III. Oxford: Blackwell.

Fanselow, Gisbert. 2004. The mlc and derivational economy. In *Minimality Effects in Syntax*, ed. by A. Stepanov, G. Fanselow, and R. Vogel, 73–123. Berlin: Mouton de Gruyter.

Gallmann, Peter. 1997. Zu Morphosyntax und Lexik der w-Wörter. Bericht Nr. 107, Arbeitspapiere des SFB 340, Universität Tübingen.

Gärtner, Hans-Martin, and Jens Michaelis. 2010. On the treatment of multiple-*wh*-interrogatives in minimalist grammars. In *Language and Logos*, ed. by T. Hanneforth and G. Fanselow, Studia grammatica, 339–366. Berlin: Akademie Verlag.

Grewendorf, Günther. 2001. Multiple wh-fronting. *Linguistic Inquiry* 32.87–122.

Haspelmath, Martin. 1997. *Indefinite Pronouns*. Oxford: Clarendon Press.

Joshi, Aravind K. 1985. Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In *Natural Language Parsing. Psychological, Computational, and Theoretical Perspectives*, ed. by D. R. Dowty, L. Karttunen, and A. M. Zwicky, 206–250. New York, NY: Cambridge University Press.

——, K. Vijay-Shanker, and David J. Weir. 1991. The convergence of mildly context-sensitive grammar formalisms. In *Foundational Issues in Natural Language Processing*, ed. by P. Sells, S. M. Shieber, and T. Wasow, 31–81. Cambridge, MA: MIT Press.

Kayne, Richard S. 1983. Connectedness. *Linguistic Inquiry* 14.223–249.

Lasnik, Howard, and Mamoru Saito. 1992. *Move α*. Cambridge, MA: MIT Press.

Leu, Thomas. 2008. The internal syntax of determiners. *Groninger Arbeiten zur germanistischen Linguistik* 47.

Lotfi, Ahmad. 2003. Persian wh-riddles. In (Boeckx and Grohmann 2003), 161–186.

Maki, Hideki. 2002. A note on the wh-cluster hypothesis. *Bulletin of the Faculty of Regional Studies, Gifu University* 11.79–88.

Michaelis, Jens. 2001a. Derivational minimalism is mildly context-sensitive. In *Logical Aspects of Computational Linguistics (LACL '98)*, Lecture Notes in Artificial Intelligence Vol. 2014, ed. by M. Moortgat, 179–198. Berlin, Heidelberg: Springer.

——. 2001b. *On Formal Properties of Minimalist Grammars*. Linguistics in Potsdam 13. Potsdam: Universitätsbibliothek, Publikationsstelle.

Penka, Doris, and Arnim von Stechow. 2001. Negative Indefinita unter Modalverben. In *Modalität und Modalverben im Deutschen*, ed. by R. Müller and M. Reis, 263–286. Hamburg: Buske Verlag.

Reglero, Lara. 2003. Non-wh-fronting in Basque. In (Boeckx and Grohmann 2003), 187–227.

Richards, Norvin. 2001. *Movement in Language. Interactions and Architectures*. Oxford: Oxford University Press.

Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6.445–501.

Sabel, Joachim. 2001. Deriving multiple head and phrasal movement. The cluster hypothesis. *Linguistic Inquiry* 32.532–547.

Saito, Mamoru. 1994. Additional-wh effects and the adjunction site theory. *Journal of East Asian Linguistics* 3.195–240.

Stabler, Edward P. 1997. Derivational minimalism. In *Logical Aspects of Computational Linguistics (LACL '96)*, Lecture Notes in Artificial Intelligence Vol. 1328, ed. by C. Retoré, 68–95. Berlin, Heidelberg: Springer.

——. 1998. Acquiring languages with movement. *Syntax* 1.72–97.

——. 2008. Computational perspectives on minimalism. Unpublished manuscript, University of California, Los Angeles, CA.

Contact address:

Hans-Martin Gärtner
Hungarian Academy of Sciences
Research Institute for Linguistics
Benczúr u. 33
1068 Budapest
Hungary
gaertner@nytud.hu