The clause mate condition: where wh-clusters meet Mereological Syntax

Klaus Abels, UCL, k.abels@ucl.ac.uk

Keywords: multiple wh-fronting, wh-cluster, cyclicity, proper binding condition, Mereological Syntax

A (de-)clustering analysis of multiple wh-movement in Mereological Syntax derives the clause mate condition.

Multiple cross-clausal wh-movement has a curious profile. For example, Abels & Dayal (2023 LI) report that multiple sluicing obeys a clause-mate condition (CMC: all remnants of multiple sluicing must originate in the same finite CP) independently of the wh-movement type (multiple/single/no fronting): the multiple sluice in (1) is but that in (2) is not acceptable (see for experimental confirmation Cortes Rodrigues et al (2024 WCCFL)):

(1) John thinks that every professor spoke with some student but I don't know which professor with which student.

(2) *Every professor thinks that John spoke with some student but I don't know which professor with which student.

At the same time, otherwise illicit long distance construals of certain wh-phrases (principally adjuncts) can become acceptable if accompanied by licit long overt or covert movement of a wh-argument (a.k.a. additional wh effects – see e.g. Grewendorf 2001 LI).

Abels & Dayal (2023) derive CMC for multiple sluicing from the clause boundedness of covert (wh-)movement. They note (p. 434 fn. 3), however, that many multiple wh-fronting languages obey CMC also with overt multiple wh-fronting; this calls the approach based on the locality of covert movement into question. The most well-studied case of overt multiple wh-fronting disobeying CMC comes from Bosnian/Croatian/Montenegrin/Serbian (Lasnik 2014 Syntax and Bošković 2024 WCCFL). However, Georgieva et al. (2025, IGG) show that the counterexamples from BCMS are confounded and that for better controlled examples BCMS does obey CMC after all (both in multiple wh-fronting and in multiple sluicing). (A small handful of other cases from Abels & Dayal (2023) still awaits careful investigation.) The purpose of this talk is to account for CMC effects by updating the clustering analysis from Grewendorf (2001) using ideas presented in Adger's 2025 monograph 'Mereological Syntax'.

Traditional clustering analyses of multiple wh-fronting derive clusters by way of sideward movement, violating the extension condition/proper binding condition on traces. The proper binding condition is, of course, well motivated by minimal paradigms like the following, where the ungrammatical (3c) violates the proper binding condition:

(3) a. I know that [the fact that you know that Fred's asleep] will shock Mary.
b. I know who [the fact that you know that Fred's asleep] will shock ~~who~~.
c. *I know (that) [the fact that you know who Fred's asleep] will shock~~who~~

We concur with Adger that on both empirical and conceptual grounds clusters must not be derived by movement. If clusters are not derived by internal merge but do exist, then they must be generated by external merge (i.e., before the lowest element of the cluster enters the main derivation). A schematic example like (4) with clustering would then receive the rough analysis in (5) (functional heads v, T, and C suppressed for legibility). The cluster is formed by external merge with the subject as an adjunct to the object (see Adger (2025:chapter 5) for details and an analysis of some of the constraints on the order of wh-phrases). The *who-what* cluster is merged into the object's theta position. Here the object *what* but not the adjunct to it (*who*) is thematically licensed. The next step involves derivational dissolution of the cluster, *declustering*. According to Adger, *who* next declusters and moves to its theta-position and subsequently Spec,TP. Finally, the entire cluster moves from the object position to Spec,CP. In

cases of long-distance multiple wh-movement, the cluster would then move on cyclically from Spec,CP to its final landing site.

(4) WHO WHAT BOUGHT?

(5) [$_{CP}$ [WHO WHAT] [$_{TP}$ ~~who~~ [$_{vP}$ ~~who~~ [$_{VP}$ BOUGHT ~~[who what]~~]]]]

The key fact to observe here is that *who* needs a theta role, which it cannot be assigned as an adjunct to *what*. Declustering is necessary and it is necessarily movement of *who* into a theta-position. Movement to theta positions is, of course, employed in the movement theory of control (overview in Hornstein & Polinsky 2010 chapter 1). As noted in Williams (2002: Representation Theory, Abels 2007 Linguistic Variation Yearbook, Abels 2012 Phases: an Essay on cyclicity in Syntax), if movement into theta-position is to be allowed in the grammar at all, it must be integrated into a generalized system of proper and improper movement and ordered before classic A-movement (movement in relation to case and phi-agreement): theta-movement before A-movement before A'-movement. Abels (2007) argues at length that the generalized ban on improper movement applies not only to movement *of* a particular constituent but also to movement *out of* a constituent. For the regular cases of A- and A'-movement this means that an A'-moved constituent cannot itself A-move (classic improper movement) but also that A-extraction from it is impossible (generalized improper movement): A'-movement bleeds A-movement of and out of the moved constituent. Generalizing to movement into theta-positions we derive that A'-movement bleeds movement into theta-positions of and out of the moved constituent. This is key in deriving CMC effects. Consider the derivational possibilities for an example violating the clause-mate condition under Adger's assumption that clusters are formed by external merge. We are here trying to derive the presluice of (2):

(6) a. [$_{CP}$ [which professor with which student] [$_{TP}$ ~~which professor~~ [$_{vP}$ ~~which professor~~ [$_{VP}$ thinks [$_{CP}$ ~~[which professor which student~~ [$_{TP}$ John spoke ~~[which professor with which student]~~]]]

b. [$_{CP}$ [which professor with which student] [$_{TP}$ ~~which professor~~ [$_{vP}$ ~~which professor~~ [$_{VP}$ thinks [$_{CP}$ ~~[which professor with which student~~ [~~which professor~~ [$_{TP}$ John spoke ~~[which professor with which student]~~]]]]

c. [$_{CP}$ [which professor with which student] [$_{TP}$ ~~which professor~~ [$_{vP}$ ~~which professor~~ [$_{VP}$ thinks [$_{CP}$ {~~which professor~~|~~which student~~} [$_{TP}$ John spoke ~~[which professor with which student]~~]]]]]

Derivations (6a-b) and variants violate the generalized ban on improper movement: in (6a), *which professor* moves from an A'-moved constituent in Spec,CP to a theta position; in (6b), *which professor* moves to a theta-position via an A'-position. In addition, (6b) violates the restriction – derived from deeper principles in Adger's framework – that lower specifiers are not escape hatches. Finally any version of (6c) violates the condition that movement from clausal domains must be successive cyclic via the escape hatch position. Again, this is derived from deeper principles in Adger's framework. The same type of declustering from the base position can be given for the additional wh effects discussed in Grewendorf (2001).

BCSM is not traditionally considered a wh-clustering language (unlike Bulgarian). Clustering languages show more rigid ordering of wh-phrases than we find in BCSM direct questions and do not allow the cluster to be interrupted by clitics. Under the current approach, we can capture these differences by assuming that in BCSM direct questions, declustering can happen (again) in the left periphery while this is impossible in Bulgarian, deriving a picture compatible with Rudin's (1988) seminal analysis of the left periphery of the two language types.

The analysis proposed here shares many key ideas with Grewendorf (2001), viz., clustering for multiple long distance movement and the appeal to the ban on improper movement. However, the idea of derivational declustering (Adger 2025) makes the analysis compatible with

the extension condition and the proper binding condition. Moreover, while Grewendorf invokes the ban on improper movement in a rather loose way, it is fully integrated here with the well-motivated generalization of the classic ban from Williams (2002) and Abels (2007).