

# [WLG]

WIENER LINGUISTISCHE GAZETTE

## **Comparing extractions from wh-islands and superiority effects**

*Luigi Rizzi*

Sonderdruck aus: *Wiener Linguistische Gazette* (WLG) 82 (2017): 253–261

Themenheft 11-11-17. *Festschrift für Martin Prinzhorn*  
Hg. v. Clemens Mayr und Edwin Williams

Universität Wien · Institut für Sprachwissenschaft · 2017

**Eigentümer, Herausgeber und Verleger:**

Universität Wien, Institut für Sprachwissenschaft  
Sensengasse 3a  
1090 Wien  
Österreich

**Redaktion:** Christian Bendl, Mi-Cha Flubacher (Angewandte Sprachwissenschaft),  
Stefan Schumacher (Allgemeine und Historische Sprachwissenschaft)

**Kontakt:** [wlg@univie.ac.at](mailto:wlg@univie.ac.at)

**Homepage:** <http://wlg.univie.ac.at>

**ISSN:** 2224-1876

**NBN:** [BL078,1063](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0078-1063-9)

Die *Wiener Linguistische Gazette* erscheint in loser Folge im Open-Access-Format.  
Alle Ausgaben ab Nr. 72 (2005) sind online verfügbar.

## Comparing extractions from wh-islands and superiority effects<sup>\*</sup>

Luigi Rizzi

University of Geneva & University of Siena

### 1. Introduction: the problem

It is instructive to compare the results on extraction from wh-islands arrived at in Villata, Rizzi & Franck (2016) with results obtained on multiple questions with superiority violations by Hofmeister et al (2013). Both sets of results concern the relative acceptability of degraded configurations, and are obtained through formal, controlled techniques (Likert scales with a sizable number of experimental subjects: Sprouse, Schütze & Almeida, 2013). They are parallel in certain respects, and interestingly diverging in others. Here are two representative examples of the two deviant configurations:<sup>1</sup>

- (1) a. Extraction from wh islands: \*What does Mary wonder who read \_\_\_?  
b. Superiority violation: \*Mary wonders what who read \_\_\_

Both configurations involve movement of a wh-element across another one, and are judged as degraded at variable degrees, modulated by the properties of the wh-element(s); they differ in that the two wh-elements have scope over different clauses in cases of extraction (in (1a), *what* has scope over the main clause, and *who* over the embedded clause, the indirect question), while they have scope over the same clause in cases of multiple questions (the indirect question in (1b)).

---

<sup>\*</sup>I worked out the bulk of this analysis in the context of the research on extractions from weak islands which was published as Villata, Rizzi & Franck 2016. In the end, my coauthors and I decided not to put a full-fledged analysis of superiority in the paper, which was mainly focused on wh-islands, and addressed to an interdisciplinary audience also including experimental psycholinguists; so we only hinted at these ideas in the paper, without fully developing them. Still, I believe the detailed comparison between wh-islands and superiority along these lines is intriguing, and worth making accessible. The analysis very much bears on issues of scope at LF, and more broadly of the syntax-interpretation interface. I know how passionate Martin Prinzhorn has always been on these topics in his teaching and in discussions with his students, so, I very much hope he will like this piece.

<sup>1</sup>This paper focuses on differences in relative acceptability between deviant structures, so from now on I will not use the familiar diacritics expressing the absolute status of a configuration.

Hofmeister et al. (2013) compared the relative acceptability of cases of multiple questions in which the lower *wh*-element is moved to the front crossing over the higher *wh*-element (superiority violations, see below for details), and manipulating the bare (*who*, *what*) or lexically restricted (*which NP*) character of the *wh*-element:<sup>2</sup>

- (2) a. Mary wondered what who read \_\_\_ (bare-bare)  
 b. Mary wondered which book who read \_\_\_ (which-bare)  
 c. Mary wondered what which boy read \_\_\_ (bare-which)  
 d. Mary wondered which book which boy read \_\_\_ (which-which)

Hofmeister et al. (2013)'s results can be summarized as follows:

- (3) i. the bare – bare case (2a) is the least acceptable case in this paradigm;  
 ii. the which – which case (2d) is the most acceptable case;  
 iii. the which – bare and bare – which cases (2b,c) are intermediate

In their first experiment, Hofmeister et al. (2013) found a preference for bare-which over which-bare (hence, (2c) more acceptable than (2b)); in their fourth experiment the two cases turned out to be about on a par, both significantly better than bare-bare, and significantly worse than which-which.

In short:

- (3') bare-bare < which-bare ≤ bare-which < which-which

Let us directly compare these results with what Villata, Rizzi & Franck (2016) found in configurations of extraction from indirect questions. While the which-which condition (4d) was the most acceptable in the extraction environment, too, the which-bare condition (4b) was significantly more acceptable than the bare-which condition (4c), which in turn was about on a par with the bare-bare condition:

- (4) a. What did you wonder who read \_\_\_ ? (bare-bare)  
 b. Which book did you wonder who read \_\_\_ ? (which-bare)  
 c. What did you wonder which boy read \_\_\_ ? (bare-which)  
 d. Which book did you wonder which boy read \_\_\_ ? (which-which)

In short, for extraction from the *wh*-island we have:

- (4') bare-bare = bare-which < which-bare < which-which

---

<sup>2</sup>To be compared with a baseline in which the higher *wh*-element is moved and the lower one, the object in these cases, remains *in situ*:

- (2') a. Mary wondered who \_\_\_ read what  
 b. Mary wondered who \_\_\_ read which book  
 c. Mary wondered which boy \_\_\_ read what  
 d. Mary wondered which boy \_\_\_ read which book

So, while the bare-bare case and the which-which case are consistently judged as highly and moderately degraded, respectively, in both (3') and (4'), there is a surprising discrepancy in the acceptability pattern of cases b. and c. in superiority violations and extractions from indirect questions: which-bare is more acceptable than bare-which in extraction contexts (4) (the Villata, Rizzi, Franck results), while the opposite pattern is found in superiority violation contexts (2) in Hofmeister et al.'s first experiment, with bare-which more acceptable than which-bare. In Hofmeister et al.'s fourth experiment the discrepancy is attenuated, with which-bare and bare-which about on a par, but the result in the superiority cases still contrasts with the result in extraction cases, with which-bare significantly better than bare-which in the extraction configuration.

I believe that this discrepancy between (2b,c) and (4b,c) can be amenable to a principled analysis if one takes into account the fundamental distinction between multiple questions and extractions from indirect questions: in extraction cases the two *wh*-elements have scope over two distinct clauses, and are pronounced in the appropriate scope position; in multiple questions the two *wh*-elements have scope over the same clause, and the one *in situ* is not pronounced in its scope position. A classical insight of the analyses of multiple *wh*-questions is that the *in situ* element undergoes covert movement to the left periphery to be assigned the appropriate scope at logical form. No such covert movement takes place in the extraction cases, in which each *wh*-element has been overtly moved to its appropriate scope position. The intuitive idea that I want to develop here is that the reversal in the judgment is a consequence of LF movement, which inverts the two *wh*-phrases in multiple questions.

In order to capitalize on this crucial difference, I will now sketch out an analysis of multiple *wh*-questions and of superiority effects, trying to combine major analytic ideas from the theoretical literature.

## 2. Superiority and multiple *wh*-movement

Consider a simple case of multiple questions illustrating superiority effects:

- (5) a. Who \_\_\_ said what?  
 b. \*What did who say \_\_\_ ?

In a language like English, overt *wh*-movement is restricted to one *wh*-element per clause, hence the second *wh*-element is pronounced *in situ*, in its canonical argument position. The superiority condition (Chomsky 1973) states that the superior, or higher, *wh*-element must move overtly: so (5a) satisfies the condition, while (5b), where the object has been moved, violates it. Plausibly, the logical form of (5) is something like “for what *x* and for what *y*, *x* said *y*”, hence some form of abstract, or “covert”, movement of the *wh*-element pronounced *in situ* must take place to yield the appropriate interpretation (Chomsky 1981). The covert movement idea is immediately supported by the existence of languages in which overt multiple *wh*-movement is possible. In the equivalent of (5) in such languages, both *wh*-elements move to the front (e.g., Rumanian: Alboiu 2002, Soare 2009):

- (6) Cine ce a dat lui Mihai?  
 who what gave to Mihai?

It is tempting to assume that (6) visibly manifests the movement of the second wh-element to the front, which remains covert in English. Notice that in a subset of languages with multiple overt movement the moved elements must appear in a fixed order in the left periphery (see Rudin 1988 for the original typology, and much subsequent work). If they are reversed, as in (7), the structure is ungrammatical:

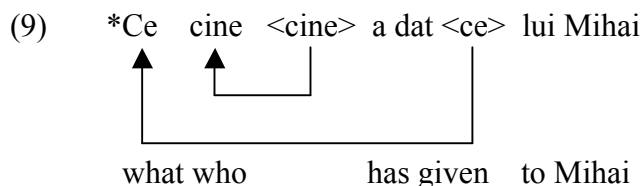
- (7) \*Ce cine a dat lui Mihai?  
 what who gave to Mihai?

It is obviously attractive to consider the hypothesis that the superiority effect in (5) and the ordering constraint in (6)-(7) are the two sides of the same coin (Richards 1997, Pesetsky 2000).

An approach which immediately traces back (6)-(7) to a version of Relativized Minimality (RM) is the one in Krapova & Cinque (2008) for the Bulgarian equivalents. If one considers the two movement steps involved in the derivation of (6), each of them seems to violate RM at the moment in which it takes place (I adopt here the “copy theory of traces”, as in Chomsky 1995, and express the trace as a silent copy of the moved element within angled brackets):

- (8) Cine ce <cine> a dat <ce> lui Mihai  
 who what has given to Mihai
- 

As in standard practice, I will use the term “chain” to refer to the formal object consisting of a moved element and its trace, a silent occurrence of the same element under the adopted theory. The structure that is derived in (8) has crossing chains, as the arrows indicate, and the intervener always is only one member of the relevant chain: a complete chain never intervenes on any link of the other chain. So, Krapova & Cinque observe that this state of affairs suggests that (i) RM applies on the derived representations, rather than on each application of movement (possibly, at the end of each phase: Chomsky 2001), and (ii) “Z intervenes between X and Y” is to be understood as “all the occurrences of Z intervene”: in (8), only one occurrence of *ce*, but not the whole chain, intervenes between *cine* and its trace, and only one occurrence of *cine*, but not the whole chain, intervenes between *ce* and its trace. The ungrammatical order is correctly ruled out by this interpretation:



Here, both occurrences of *cine* intervene between *ce* and its trace, hence the structure is ruled out by RM, under this interpretation. So, intersecting (or crossing) chains are allowed, while nested chains are excluded.

At this point, one may try to link the superiority effect (5) to the account just adopted for (6)-(7). Many approaches have been proposed to express the covert occurrence of movement in (5a,b). Chomsky (1995) proposed that covert movement is in fact movement of the (relevant) features of the *wh*-element to the relevant scope position. For concreteness, I will adopt a variant of this proposal which understands covert movement as a kind of incorporation/cliticisation of (the relevant features of) the *wh*-element *in situ* into the attracting head, here Q, yielding the following LF representation for (5a) (I continue to adopt the copy theory of traces; the chain created by covert movement in bold; “what” is in fact to be understood here as a cover term for whatever features are moved from the *in-situ* position):

(10) Who [ **what** Q ] [ <who> said <**what**> ]

(This proposal may in fact be considered a cliticization/incorporation variant of Richards’ “tucking in” hypothesis, with the features of the unmoved *wh*-element cliticized onto Q). (10) is ruled in under Krapova & Cinque’s interpretation of RM, while the LF derived from (5b) through covert incorporation of (the relevant features of) *who* into Q yields the following:

(11) What [ **who** Q ] [ <**who**> said <what> ]

This is ruled out under Krapova & Cinque’s interpretation of RM, on a par with (9), as the whole chain (*who*, trace) intervenes between *what* and its trace.

Consider now the corresponding sentences, but with lexically restricted *wh*-phrases:

- (12) a. Which student \_\_\_ solved which problem?  
 b. Which problem did which student solve \_\_\_?

(12a) remains non-problematic, on a par with (5a): the relevant features of *which problem* move covertly, yielding an intersecting chain analogous to (10). As for (12b), the assumption has been made elsewhere (Rizzi 2011, Villata, Rizzi & Franck 2016 for arguments and discussion) that lexically restricted *wh*-elements can be attracted both by Q, and by the complex feature conglomerate [Q, N], sitting in a higher position than bare Q in the map of the left periphery. Hence, in such cases we have two attractors available:

(12’) [Q, N] ... [Q] ....

The availability of the second landing site is decisive for permitting a well-formed representation of (12b). Suppose that *which problem* overtly moves to Spec-Q, yielding the following representation at spell-out:

(13) [ which problem Q [ which student solve <which problem> ] ]

At this point the complex higher attractor [Q, N] can be merged and trigger covert movement of (the relevant features of) *which student*, yielding the following (again, with the chain created by covert movement in bold):

(14) **Which student** [Q, N] [which problem Q [ <**which student**> solve <which problem> ] ]

This representation involves intersecting chains, the configuration which is permissible under K&C's interpretation of RM. In a nutshell, the fact that with lexically restricted wh-elements superiority violations are alleviated in part is related to the additional landing site available for such elements through the complex attracting head [Q, N].<sup>3</sup> The same property was also assumed to have a role in determining the (marginal) possibility of extraction of such elements from indirect questions (Villata, Rizzi & Franck 2016). The special properties of lexically restricted wh-elements wrt. superiority and extraction contexts are thus unified in part, and traced back to a single abstract property, the existence of an additional landing site for such elements.<sup>4</sup>

### 3. A fundamental difference between extraction from indirect questions and multiple questions: covert movement

The two environments also differ in part, though. The fundamental insight that covert movement is crucially involved in computing representations for multiple questions is instrumental to understand the surprising discrepancy between Hofmeister et al. (2013)'s results on multiple questions and the results in extraction cases in Villata, Rizzi & Franck 2016, reproduced below from (2) and (4) for ease of reference:

(15) *Extraction from indirect question: (which-bare) > (bare-which)*

(Villata, Rizzi & Franck)

b. (which-bare) Which book did you wonder who read \_\_\_?

c. (bare-which) What did you wonder which boy read \_\_\_?

---

<sup>3</sup>Ur Shlonsky (p.c.) observes that (12b) remains marginal to some extent if compared to (12a), even though it is clearly more acceptable than the bare-bare case *\*What did who read?* So, forming intersecting chains through covert movement must involve an inherent cost (not involved in intersecting chains derived via overt movement: see (6)) responsible for the marginality of the example, for reasons that I will not explore here.

<sup>4</sup>In fact, a related but distinct factor is the set theoretic composition in relevant features of the extractee and of the intervening element. See the discussion of featural Relativized Minimality in Villata, Rizzi & Franck 2016, based on Friedmann, Belletti & Rizzi (2009): see below on the role of this factor in extraction from indirect questions.





Featural Relativized Minimality, as worked out in Friedmann, Belletti & Rizzi 2009 correctly draws the distinction to the advantage of (15b) in this case.

The contrast between Hofmeister et al. (2013) on multiple questions and the result in Villata, Rizzi & Franck 2016 on extraction from indirect questions thus is ultimately related to an independent difference between extraction and superiority environments: the fact that a superiority configuration, but not an extraction configuration, involves covert movement of one wh-element. This critical difference partly reverses the judgment in the superiority case.

To conclude, it is worthwhile to observe that the empirical results obtained through controlled experimental techniques of judgment gathering corroborate the following points of significant theoretical relevance:

1. RM is checked on final representations (logical forms, possibly evaluated at the end of each phase). If it was checked in the course of the derivations, on the individual applications of movement, intersecting chains could never be derived (either by overt or covert movement).
2. The appropriate interpretation of multiple questions in English (and similar languages) is determined by covert movement of (the relevant features of) the wh-element pronounced *in situ* to the appropriate scope position. Covert movement in multiple questions thus plays a critical role in explaining the otherwise surprising reversal of judgment with respect to cases of overt extraction from wh-islands.

### References

- Alboiu, Gabriela. 2002. *The features of movement in Romanian*. Bucharest: Bucharest University Press.
- Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen Anderson & Paul Kiparsky, 232-286. New York, NY: Holt Rinehart and Winston.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.
- Friedmann, Naama, Adriana Belletti, & Luigi Rizzi. 2009. Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua* 119(1):67-88.
- Hofmeister, Philip, T. Florian Jäger, Ivan A. Sag, Inbal Arnon, & Neal Snider. 2013. The source ambiguity problem: distinguishing effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes* 28(1): 48-87.
- Krapova, Ilyiana, & Guglielmo Cinque. 2008. On the Order of wh-Phrases in Bulgarian multiple wh-Fronting. In *Formal description of Slavic languages: the fifth conference, Leipzig 2003*, ed. Gerhild Zybatow, 318-336. Frankfurt am Main: Peter Lang.
- Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge, MA: MIT Press.

- Richards, Norvin. 1997. What moves where when in which language? Doctoral dissertation, MIT, Cambridge, MA.
- Rizzi, Luigi. 2011. Minimality. In *The Oxford handbook of linguistic minimalism*, ed. Cedric Boeckx, 220-238. Oxford, New York, NY: Oxford University Press.
- Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6: 445-501.
- Soare, Gabriela. 2009. *The syntax-information structure interface: a comparative view from Romanian*. Doctoral dissertation, University of Geneva, Geneva.
- Sprouse, Jon, Carson T. Schütze, & Diogo Almeida. 2013. A comparison of informal and formal acceptability judgments using a random sample from *Linguistic Inquiry* 2001-2010. *Lingua* 134: 219-248.
- Villata, Sandra, Luigi Rizzi, & Julie Franck. 2016. Intervention effects and relativized minimality: New experimental evidence from graded judgments, *Lingua* 2016.

Luigi Rizzi  
luigi.rizzi@unisi.it