Available online at www.sciencedirect.com

ScienceDirect

Lingua 179 (2016) 76–96

Lingua

www.elsevier.com/locate/lingua

Intervention effects and Relativized Minimality: New experimental evidence from graded judgments

Sandra Villata [a,*], Luigi Rizzi [b,c], Julie Franck [a]

[a] Département de Psycholinguistique, Développement du Langage et Cognition, Faculté de Psychologie et des Sciences de l'éducation, Université de Genève, Genève, Switzerland

[b] Département de Linguistique, Faculté des Lettres, Université de Genève, Genève, Switzerland

[c] CISCL – Centro Interdipartimentale di Studi Cognitivi sul Linguaggio, Università di Siena, Complesso San Niccolò, Siena, Italy

Received 1 December 2014; received in revised form 31 March 2016; accepted 31 March 2016
Available online 4 May 2016

Abstract

According to Featural Relativized Minimality, the local relation between an extracted element and its trace is disrupted when it crosses an intervening element whose morphosyntactic featural specification matches the specification of the elements it separates. This approach naturally leads to a system able to capture degrees of deviance: the relative acceptability of an intervention configuration will vary as a function of the total, partial or zero featural overlap between the intervener and the target. In a nutshell, configurations involving a lesser degree of featural overlap should be more acceptable than sentences involving a higher degree of overlap. Three acceptability judgment experiments systematically investigated predictions ensuing from Featural Relativized Minimality in extraction from weak islands. Four configurations of feature overlap were systematically tested with different methods of data collection and on a large set of linguistically naïve participants. Results from the three experiments are highly consistent in returning that predictions from Featural Relativized Minimality are globally borne out, except for the configuration involving two lexically restricted wh-elements, for which tentative explanations in terms of grammar or processing are sketched out.

Keywords: Intervention locality; Relativized Minimality; Weak islands; Discourse-linking; Similarity-based interference; Cue-based memory model

1. Introduction

Ever since Ross (1967) it has been claimed that wh-movement respects certain island constraints. For instance, a wh-element like *how* can be extracted from a declarative (1), but not from an indirect question (2).

(1) How do you think John could solve the problem __?
(2) *How do you wonder whether John could solve the problem __?

Indirect questions like (2) illustrate the environments known as *weak islands*. Weak islands are traditionally opposed to strong islands, environments such as complex noun phrases and adjunct islands, in which extractions are assumed to

* Corresponding author at: Uni Pignon, Boulevard du Pont d'Arve 42, 1205 Genève, Switzerland. Tel.: +41 22 379 91 45.
E-mail addresses: sandra.villata@unige.ch (S. Villata), luigi.rizzi@unige.ch (L. Rizzi), julie.franck@unige.ch (J. Franck).

http://dx.doi.org/10.1016/j.lingua.2016.03.004
0024-3841/

give rise to severe ill-formedness not modulated by the nature of the extractee (but see Sprouse et al., 2016, for a more complex picture based on a quantitative definition of island effects made possible by formal experimental methods). Unlike strong islands, the deviance of the extraction from a weak island is modulated by the nature of the extracted wh-element (see Szabolcsi, 2006, for an overview). For instance, while extraction of a bare wh-adverbial like *how* is sharply excluded, extraction of a lexically restricted wh-argument of the form *which NP* (3) is more acceptable, although perceived as deviant to some degree.

(3) ??Which problem do you wonder whether John could solve__ (in this way)?

The deviance of (2) as well as the milder deviance of (3) is typically analyzed as an *intervention effect*, to be captured by Relativized Minimality (Rizzi, 1990) or the corresponding derivational principles introduced in different versions of minimalist syntax, the Minimal Link condition (Chomsky, 1995), or Minimal Search (Chomsky, 2000, 2007). *Intervention locality* partly corresponds to the concept of weak island, while the notion of *impenetrability locality* partly corresponds to the concept of strong islands, as the former solely involves violations of intervention locality in traditional approaches (on the distinction, and the prospects of a unifying approach, see the discussion in Rizzi, 2009).

Informally, the relation between *how* and its trace in (2), and between *which problem* and its trace in (3), is disrupted by the intervention of an element of the same type as the elements it separates, the wh-element *whether* in the embedded complementizer system. The question then arises of why the intervention effect is weakened in (3).

Many factors have been considered as potentially responsible for the contrast between (2) and (3), like the argument/adjunct distinction (Huang, 1982), the distinction between bare and lexically restricted wh-elements (Friedmann et al., 2009), and the distinction between D(iscourse)-linked and non-D-linked elements (Comorovski, 1989; Cinque, 1990; Rizzi, 1990 and much subsequent work building on Pesetsky, 1987). In this paper we will not test classical cases of adjunct extraction from weak islands: on the one hand, they represent the most robust case of weak island violation, but, on the other hand, the judgment may be affected by the possibility of main clause attachment for the adjunct, an option that is not easy to exclude in an experimental setting with naïve informants. So, we will focus on extraction of wh-arguments in which the extraction site is unambiguous, modulating the bare or lexically restricted nature of both the extracted argument and the intervener.[1]

Even if the most robust case of weak island violation is represented by the extraction of an adjunct (as in (2)), the improvement observed for the extraction of a lexically restricted argument (3) is also observable when compared to the extraction of a bare argument (4).

(4) *What do you wonder whether John could solve __ (in this way)?

Moreover, the contrast (3)–(4) is more minimal than the contrast (3)–(2) as it solely involves the presence or absence of the lexical restriction in otherwise identical structures. There may well be an additional argument vs. adjunct factor responsible for making the contrast (3)–(2) particularly sharp, but this factor will not be investigated in this paper. So, here we will restrict our attention to a subset of the cases classically illustrating weak island effects, that is, arguments.

The relevant empirical domain often consists of subtle comparative judgments between different deviant sentences. The facts are typically established through the classical informal methodology of judgment gathering in modern theoretical and comparative syntax. As we systematically deal with graded judgments (B not as good as A, but better than C, etc.) it is worthwhile to also try out more sophisticated methodologies of data gathering, to see how far we can go with controlled judgments on complicated sentences by many linguistically naïve speakers (e.g., Sprouse et al., 2013). Such methods could highlight facets of the empirical domain which otherwise remain unnoticed with a more traditional methodology. Moreover, the subtlety of the judgments on the various types of weak islands makes converging evidence from various empirical sources particularly important before firm conclusions can be drawn. So, one of the purposes of this paper is to explore the heuristic capacity of such methods for enriching the empirical basis of sophisticated formal hypotheses in the study of weak islands and intervention effects.

In section 1 we present featural Relativized Minimality (henceforth, fRM) and illustrate how it accounts for the classical cases of acceptability judgments in extractions from weak islands, and spell out its predictions with respect to the role of

[1] Other factors affect the acceptability of extraction from weak islands. For instance, all other things being equal, extraction from an untensed weak island is more acceptable than extraction from a tensed weak island:

(i) a. ?? Which problem do you wonder how Bill could solve?
b. ? Which problem do you wonder how to solve?

The overall acceptability of a particular case is thus a function of several concomitant factors. In this paper we will only focus on some such factors.

lexical restriction and D-linking in the acceptability of argument extraction. In sections 2, 3 and 4 we report three acceptability judgment experiments (based on scaled and binary judgments) in which we systematically investigated the role of these variables in sentence acceptability. In section 5 we discuss our results against the background of the predictions from fRM, as well as in light of theories of memory interference, and we then conclude with some remarks on the comparison between Relativized Minimality and Superiority.

1.1. Featural Relativized Minimality

While in previous systems the contrast between (2) and (3) was addressed by assuming that wh-elements like *which problem* can use non-local devices connecting them to their traces (Rizzi, 1990 assumed the possibility of a non-local binding relation made possible by the legitimacy of referential indices in such cases, Cinque, 1990 assumed the option of a non-local binding of a null resumptive pronoun), Starke (2001) put forth the idea that a single connecting device, modulated by the nature of the wh-elements at play, could capture the facts. The fundamental idea of fRM can be expressed as follows (building on Starke, 2001; Rizzi, 2001, 2004):

(5) Featural Relativized Minimality:
In . . . X . . . Z . . . Y . . .

A local relation is disrupted between X and Y when

a. Z structurally intervenes between X and Y
b. Z matches the specification in morphosyntactic features of X

Intervention is defined in hierarchical terms through c-command:

(6) Z structurally intervenes between X and Y when Z c-commands Y and Z does not c-command X.

The crucial idea here is that the typology of positions responsible for intervention effects is determined by the morphosyntactic features that define the different positions with respect to the local relation which is checked. As we are looking at local relations determined by movement, the relevant features will be those that participate in the triggering of movement. This is entirely in the spirit of the principle: the locality effect is not absolute, but relativized to the kind of local relation we are looking at. So, we are assuming that the morphosyntactic features that define a position involved in a movement relation are those that trigger movement to that position, and locality is computed in terms of such features: if Z matches the relevant specification of X, the connection between X and Y is disrupted.

Building on this insight, Friedmann et al. (2009) put forth a straightforward set-theoretic approach to the calculation of feature overlap, as is shown in (7), where X is the moved element in the targeted position, Y is the trace of the extracted element, or the source position of movement, and Z is the intervening element; +A and +B are the morphosyntactic features attached to those elements:

(7)	X	Z	Y
a. Identity:	+A	+A	+A
b. Inclusion:	+A,+B	+A	+A,+B
c. Disjunction:	+A	+B	+A

The configurations in (7) represent three possible patterns of relation between the featural specification of the intervener (Z) and the featural specification of the target (X): whenever the featural specification of the intervener is identical to the featural specification of the target (case (7a)), an intervention effect arises; therefore, no local relation can be established between the target X and the source Y, and the sentence is ruled out. The opposite case is Disjunction (7c): if the potential intervener Z has a disjoint specification with respect to the target, no intervention effect arises, and the local relation between X and Y can be successfully established. The Inclusion case (7b) represents an intermediate case: when the featural specification of Z is properly included in the specification of X and Y, we have a partial match. Under the natural assumption that the amount of overlap provides a measure of the degree of deviance, we may expect that this intermediate situation gives rise to a violation which is felt as intermediate between the full acceptability of the disjunction case and the stronger ill-formedness of the identity case. In other words, we may interpret ‘‘Z matches the specification in morphosyntactic features of X’’ in (5b) along the following lines:

(8) a. Full match: strong violation
b. Partial match: weak violation
c. No match: well-formedness

Examples (1), (2), (3) illustrate the three cases. Assume that the relevant featural specification defining the surface position of *how* is [+Q], the feature defining questions and attracting the wh-element; in (2), a [+Q] bearing element, *whether*, intervenes in the stretch of the tree between *how* and its trace:

(2) *How do you wonder [whether John could solve the problem __]?
[+Q] [+Q]
X Z Y

As the relevant featural specification of Z fully matches the specification of X, the local relation between X and Y is disrupted and the structure is severely ill-formed (the same conclusion holds if there are intermediate traces of *how*, for instance in the edge of each phase in the sense of Chomsky, 2001: one link of the chain of *how* will inevitably cross *whether*, and this will disrupt the chain connection). In (1), no other intervening position has the featural specification [+Q] (e.g., whatever feature defines the embedded and main subject position, e.g., [+N], it will not be [+Q]), hence any potential Z in (1) is featurally disjoint from X, no intervention effect is triggered and the connection between *how* and its trace is well-formed:

(1) How do you think John could solve the problem __?
[+Q] [+N]
X Z Y

It is important to stress the difference between the property of being in a structural position which intervenes between two other positions (as defined in (5a) and in (6)) and the property of triggering an *intervention effect* (as defining in (5)): while the former is a pure geometrical property of an element occupying a certain structural position in between two other elements, the latter is the combination of an element being in a certain structural position (5a) *plus* having a certain featural specification (5b). To illustrate, in (1) 'John' intervenes between 'how' and its trace, but it does not cause any intervention effect as its featural specification does not match the featural specification of the extracted element.

As for (3), Friedmann et al. (2009) (fn. 8) assume that the relevant featural specification of the wh-element is complex, [+Q,+N], where [+N] expresses the lexically restricted character of the element. Therefore, in (3) we have a representation like the following:

(3) ?? Which problem do you wonder [whether John could solve __] (in this way)?
[+Q, +N] [+Q]
X Z Y

Here the specification of X properly includes the specification of Z; hence, the latter matches the former only partially, so that a weaker violation is produced, and a milder ill-formedness is expected to arise, under (5)–(8). The system thus accounts for the amelioration of (3) with respect to (2).

The reference to [+N] is crucial in the reference quoted above to capture the difficulties that young children have with (certain) object A'-dependencies. But can one claim on independent grounds that [+N] participates in triggering movement, and as such is taken into account in the computation of fRM? Various languages offer evidence that bare and restricted wh-elements target two distinct landing sites. A straightforward case is offered by some Northern Italian Dialects (see Munaro, 1999), where bare wh-elements end up in a clause-final position, while lexically restricted elements are clause-initial (see Friedmann et al., 2009: fn. 8). Munaro analyses the two cases as both involving leftward movement, but with a lower landing site for the bare wh-element, and a higher landing site for the lexically restricted wh-element. Further remnant movement of the IP to an intermediate position in the C-system determines the surface order in this analysis. Other languages offer other kinds of evidence. For example, in European Portuguese (Ambar, 1988) bare wh-elements require subject inversion, while lexically restricted elements do not. In Northern Norwegian dialects (Westergaard and Vangsnes, 2005), bare wh-elements do not trigger Verb Second, while lexically restricted wh-elements do. In Bavarian (Bayer and Brandner, 2008), lexically restricted elements can co-occur with the complementizer *dass*, while bare elements cannot. Moreover, and perhaps even more relevantly, in a language permitting multiple wh-movement like Romanian, a lexically restricted non-subject wh-phrase can precede a subject bare wh-element (''Cu care candidat cine a votat?'' ''For which candidate who voted?''), an ordering option which is typically excluded for bare non-subject

wh-phrases, which would require the order subject wh - non-subject wh (''Cine cu cine a votat?'' ''Who for whom voted?''). The exceptional ordering option with lexically restricted elements can be readily explained by assuming that they can target a higher position than bare elements in the left periphery (Alboiu, 2002; Soare, 2009; their analyses are in terms of D-linking, but they can be easily transposed in terms of lexical restriction). Clearly, the [+N] feature does not have a role identical to the role of the [+Q] feature, as it cannot trigger movement alone (only [+Q] being a *criterial* feature, but see footnote 2 below on the relevance of this notion); nevertheless, since it participates in the fine identification of the landing site of movement in the left periphery, as the above-mentioned evidence suggests, [+N] is involved in the triggering of movement, and thus it is taken into account in the computation of locality.

To sum up, the three cases in (7) clearly are ranked from a minimum to a maximum of distinctness between X and Z: in the case of featural identity X and Z are minimally distinct, in the case of disjunction they are maximally distinct, and the case of proper inclusion is in between. So, if we think of fRM as defining the optimal case in terms of the maximal distinctness of the target of movement from a potential intervener, we can conceptualize inclusion and identity as partial and complete departures from the optimal case, thus capturing the gradation in the acceptability judgments.[2] In this view, the grammar does not simply determine a bifurcation between ''grammatical'' and ''ungrammatical''; it assigns a more fine-grained degree of grammaticality, with three values, as in (7), or possibly more (if additional morphosyntactic features and set-theoretic relations are taken into account). Notice that the system remains discrete, as only a small number of set-theoretic relations between relevant features are involved, with the traditional binary divide replaced by a more fine-grained discrete gradation.

Finally, before moving further, it is worthwhile to briefly dwell on the conceptual relation between *intervention locality* and *wh-island*. Wh-islands are the most prototypical case of weak islands. As a theory of intervention locality, fRM accounts for wh-islands in terms of intervention effects, which is nothing else than saying that what induces a weak island effect is not the fact of having a particular construction in the island catalogue, the wh-island, but the fact of having an intervention configuration (as defined in (5)). To put it differently, fRM *reduces* the concept of wh-island to the concept of intervention (this is a particular instance of the larger program of reducing properties of grammatical constructions to properties of finer computational ingredients). In this paper we will continue to use both terms, as they span over the same range of phenomena at different degrees of abstraction, but it is worth to keep in mind that if fRM is right, then wh-islands (and possibly other cases of weak islands such as negative islands, factive islands and so forth) are nothing else than particular intervention configurations.

1.2. Plan of the study

The three configurations illustrated in (7) do not exhaust the possible set-theoretic relations between the featural specifications of X and Z. Here we will focus on two more relations: Inverse Inclusion, with the specification of Z properly including the specification of X, and Complex Identity, in which the identity of the featural specification between X and Z involves more than one feature (unlike Bare Identity, in which the identity of features involves a single feature).[3] We will refer to configurations in which a wh-element has been extracted over an intervening wh-element as Intervention

[2] The set-theoretic approach to the calculation of feature overlap is set up in Friedmann et al. (2009) to account for the fact that object relatives are difficult for children, a fact that is traced back to fRM. But consider a headed object relative (with [+R] designating the fact that we are dealing with a relative construction and [+N] expressing its headed character, as opposed to a free relative):

(i) The problem that the student solved __ .
[+R,+N] [+N]

Although (i) instantiates a configuration of inclusion, adults judge object relatives as fully acceptable. The question then arises of why (i) is fully acceptable and 'Which problem do you wonder whether John could solve?' is degraded, as they both instantiate the same set-theoretical relation of inclusion. One possibility would consist in capitalizing on the difference between *criterial* and *non-criterial features* as developed in Rizzi (1997, 2004). A criterial feature is a feature able to trigger movement alone (i.e., +Q, +R(el), +TOP, +FOC); a non-criterial feature (i.e., +N), although contributing to the fine identification of the landing site of movement, is not able to trigger movement on its own. Building on the idea that only features triggering movement are predicted to generate intervention effects, the fRM system could be further modulated by introducing a hierarchy among those features which are able to trigger movement alone (i.e., criterial features) and those able to trigger movement only if accompanied with a criterial feature (i.e., non-criterial features). In this perspective, only an overlap with criterial features would determine the perception of marginality, while an overlap with non-criterial features would be perceived as fully acceptable by adults. This conclusion is consistent with the hypothesis developed in Friedmann et al. (2009) according to which the child system can only deal with the optimal case of disjunction configurations. Sentences like (i) are indeed configurations of inclusion (hence difficult to compute for children), but not of inclusion of a criterial feature, the configuration which makes the structure degraded for adults. We leave refinements of fRM to fully integrate the child and adult pattern for future work.

[3] Another potentially relevant relation is Intersection, whose role is explored in Belletti et al. (2012). The role of the Disjunction configuration, not explored in this study, is investigated in Villata et al. (2015).

conditions (9i). The Intervention conditions were compared to four baseline conditions without intervention (9ii) (see Atkinson et al., 2015 for a similar design in English). The conditions without intervention allowed us to control for any effect of lexical restriction independent of intervention. Note that sentences with intervention involve two other sources of variability that the No extraction conditions did not control for (see Sprouse et al., 2016). One is the presence of a long-distance dependency: Intervention conditions contain a long-distance dependency that No Intervention conditions do not. Hence, the current design does not allow us to determine whether the configuration that creates the weakest intervention effect (Complex identity) still shows any intervention effect or whether it only shows a long-distance dependency effect. Thus, we will capitalize on the relative rates across the four configurations of intervention, under the reasonable assumption that the effect of long-distance dependency is equal across them. The second source of variability in target conditions lies in the animacy of the two wh-elements: the inanimate objects used in the Intervention conditions were paired with animate main clause subjects in the No Intervention conditions. Animate subjects were introduced in the latter in order to avoid introducing an independent source of difficulty (e.g., Traxler et al., 2002; Lowder and Gordon, 2012).

(9) a. Bare Identity:

i. Intervention:
Qu'est-ce que$_j$ tu te demandes [qui$_i$ __$_i$ a résolu __$_j$]?
what do you wonder who solved ?
[+Q] [+Q]

ii. No Intervention:
Qui$_j$ __$_j$ se demande [qui$_i$ __$_i$ a résolu ce problème]?
who wonders who solved this problem?
[+Q] [+Q]

b. Inverse Inclusion

i. Intervention:
Qu'est-ce que$_j$ tu te demandes [quel étudiant$_i$ __$_i$ a résolu __$_j$]?
what do you wonder which student solved ?
[+Q] [+Q, +N]

ii. No Intervention:
Qui$_j$ __$_j$ se demande [quel étudiant$_i$ __$_i$ a résolu ce problème]?
who wonders which student solved this problem ?
[+Q] [+Q, +N]

c. Inclusion

i. Intervention:
Quel problème$_j$ te demandes-tu [qui$_i$ __$_i$ a résolu __$_j$]?
which problem do you wonder who solved ?
[+Q, +N] [+Q]

ii. No Intervention:
Quel professeur$_j$ __$_j$ se demande [qui$_i$ __$_i$ a résolu ce problème]?
which professor wonders who solved this problem ?
[+Q, +N] [+Q]

d. Complex Identity

i. Intervention:
Quel problème$_j$ te demandes-tu [quel étudiant$_i$ __$_i$ a résolu __$_j$]?
which problem do you wonder which student solved ?
[+Q, +N] [+Q, +N]

ii. No Intervention:
Quel professeur$_j$ __$_j$ se demande [quel étudiant$_i$ __$_i$ a résolu ce problème]?
which professor wonders which student$_i$ solved this problem ?
[+Q, +N] [+Q, +N]

In this design, the predictions of fRM bear on the strength of the *intervention effects* defined by comparing target structures with intervention (i.e., the wh-islands) to their corresponding baselines without intervention, rather than on *raw acceptability rates* in the target structures as proposed by the classical approach. If these sentences are presented out of context, the theory predicts stronger intervention effects for Bare Identity, where the intervener fully matches the featural specification of the extracted argument, than for Inclusion where it only partially matches it. As for Complex Identity, an unqualified version of fRM predicts it to be on a par with Bare Identity, since it also involves an intervener

whose features fully match those of the extractee (but see below for an important qualification). Finally, Inverse Inclusion is also predicted to show a strong intervention effect similar to that for Bare Identity since its featural set fully matches that of the extractee (even though it is actually larger). The predictions are summarized in (10).

(10) Predictions of fRM for sentences out of context, if lexical restriction is a formal property defining feature sets, where ''>'' means ''stronger intervention effect than'' and ''='' means ''same intervention effect as'':
Bare Identity = Inverse Inclusion = Complex Identity > Inclusion

The prediction that Complex Identity should pattern alike with Bare Identity and Inverse Inclusion, all being instances of identity of features, is already known to be invalidated by informally gathered judgments suggesting that Complex Identity is better than both Bare Identity and Inverse Inclusion. The problem was noticed on the basis of the Italian equivalents of such structures in Rizzi (2011), and was tackled in the following way. We have seen that there are good reasons to assume that bare and lexically restricted wh-elements have distinct landing sites (see section 1.1), so that the map of the left periphery of the clause involves a pure [+Q] attracting head, and a (higher) attracting head with the complex specification [+Q, +N]. Suppose now that a [+Q, +N] wh-element like *which student* has both movement options: it can be attracted by the [+Q, +N] head, and also by the simple, unqualified [+Q] head, as bare wh-elements (the opposite would not be true: so, a bare wh-element cannot be attracted by a [+Q, +N] head because it does not match its specifications, so that it only has the option of moving to a [+Q] head). In other words, we can assume the following (adapted from Rizzi, 2011: 230, where the option is expressed in a framework assuming the topical feature of D-linked phrases to be the crucial ameliorating factor; the approach is immediately translatable into a system assuming the lexical restriction to be the crucial factor):

(11) *which N* can be attracted both by [+Q, +N] and by [+Q]

We continue to assume that only the attracting features are seen in the computation of fRM. So, Complex Identity (9di) permits a representation like (12), derived by moving *which student* to Spec of [+Q] in the embedded clause, and *which book* to Spec of [+Q, +N] in the main clause:

(12) Which book do you wonder [which student __ could buy __]?
[+Q, +N] [+Q]

Under this analysis, the case of Complex Identity would then reduce to a case of Inclusion. In fact, the judgment on the Italian equivalents assumed in Rizzi (2011) puts the two cases about on a par.[4]

In sum, if proviso (11) is adopted, the set of predictions (10) changes as follows:

(10′) Predictions of fRM for sentences out of context, if lexical restriction is a formal property defining feature sets, and proviso (11) is adopted, where ''>'' means ''stronger intervention effect than'' and ''='' means ''same intervention effect as'':
Bare Identity = Inverse Inclusion > Inclusion = Complex Identity

We would now like to find out what our formal design has to say on this set of predictions.

The study also aimed at addressing experimentally the precise factor(s) that distinguish(es) bare and complex wh-elements (i.e., *what* vs. *which NP*). The approach we have adopted here identifies the crucial factor as a purely formal property, namely the presence of a lexical restriction in the wh-phrase. However, another influential approach advocates an interpretive property, namely the D-linked character of the wh-phrase, i.e., the fact that the range of the variable is a

[4] An anonymous reviewer asks the question why, if proviso (11) holds, a symmetric proviso could not hold for bare wh-elements, as in (11′), which would incorrectly permit bare elements to be (marginally) extracted:

(11′) *What* can be attracted both by [+Q, +N], and by [+Q]

The answer is that the mechanics of the system disallows (11′) while allowing (11): *what* is uniquely specified as [+Q], hence it could not be attracted by the complex head [+Q, +N], requiring also the presence of a lexical restriction. On the other hand, a complex phrase is specified [+Q, +N], hence it meets the attraction conditions of both a complex head [+Q, +N], and of a simple head [+Q].

In fact, (11) can be understood as a particular case of a general and natural condition stating that the requirement for attraction is that the attracted phrase must fully match the specification of the attracting head: so, a featurally complex phrase can be attracted by both a complex and a simple head, as it fully matches both, whereas a featurally simple phrase can only be attracted by a simple head.

presupposed set, pre-established in the discourse and salient. The formal and interpretive properties often co-occur, in that lexically restricted wh-phrases tend to be interpreted as D-linked (and some wh-elements requiring a lexical restriction, such as *which* in English, are assumed to enforce the D-linked interpretation), whereas bare wh-phrases tend to be interpreted as non D-linked. But the two properties do not necessarily coincide: as Pesetsky (1987) pointed out in his seminal work introducing D-linking, a bare wh-element can be made D-linked contextually. For instance, the sentence *Who did John talk to?* uttered out of context does not presuppose a specific set of potential interlocutors, i.e., it is not D-linked. However, it does presuppose such a set in a context like the following: *Under such circumstances, John could have talked to Bill, Peter or Mary; in the end, who did he talk to?*.

In order to disentangle the role of lexical restriction and that of D-linking, we explored whether the presence of a context modulates the intervention effect in the four configurations described in (9). We were interested in two sets of predictions. The first one bears on the comparison between the four configurations. If the feature of D-linking is responsible for the modulation of intervention effects in wh-islands, the four structures, once in context, are expected to be on a par since they would all involve a configuration of Inclusion [+Top, +Q] . . . [+Q], Top being the feature associated with D-linking (under the natural extension of proviso (11) where D-linking is the ameliorative factor, following Rizzi, 2011). If, in contrast, it is the formal feature of lexical restriction that underlies the modulation of intervention effects, then the predictions are those formulated in (10′). The second prediction bears on the effect of context for each of the individual configurations. If the feature of D-linking is responsible for the modulation of intervention effects in wh-islands, then we expect both Bare Identity and Inverse Inclusion to show a stronger intervention effect out of context than in context (given that in context they reduce to a case of Inclusion), while the intervention effect of Inclusion and Complex Identity should remain unchanged (under proviso (11)), since they are cases of Inclusion both in context and out of context. If, in contrast, the feature of lexical restriction is responsible for the modulation of intervention effects, then the ratings for each of the four configurations should be identical in the presence and in the absence of context, since their featural specifications are identical in context and out of context.

In what follows we present three acceptability judgment experiments. In Experiment 1, participants were asked to judge the acceptability of sentences presented out of context on a 7-point Likert scale. In Experiment 2, a large sample of binary acceptability judgments was collected on the same structures, in order to see if results obtained in Experiment 1 were replicable. Finally, in Experiment 3 each sentence was preceded by a context story in order to explore the role of D-linking in modulating acceptability judgments.

2. Experiment 1

2.1. Method

2.1.1. Participants

Forty students at the University of Geneva participated in the experiment for course credit. All participants were between the ages of 18 and 26, and were native speakers of French. All participants were naïve to the purposes of the study.

2.1.2. Materials

We manipulated two variables in a 2 × 4 design: (1) Intervention (Intervention vs. No Intervention), and (2) Structure (Bare Identity, Inverse Inclusion, Inclusion, Complex Identity). The four levels of Structure were obtained by manipulating the lexical restriction of the extracted wh-object and the intervening wh-subject. The two variables were part of a fully crossed design with 8 experimental conditions and were manipulated within participants. Stimuli consisted of 4 groups of sentences distributed across the 8 experimental conditions. All participants saw all stimuli. An example set, translated into English, is reported in Table 1.[5]

In the Intervention condition, the extracted wh-element was always of the form *qu'est-ce que* (*what*) when it was bare, and *quel NP* (*which NP*) when it was restricted. French allows several options for the wh-element corresponding to the English *what*, namely the in situ *quoi* (an option that we discarded as we wanted a long-distance dependency), the moved *que* (which requires subject-verb inversion, e.g., *Que te demandes-tu qui a construit?*) and the moved *que* accompanied by the interrogative form *est-ce que* (pronounced /esk/) which is (synchronically) a kind of overt question marker and prevents the need of the subject-verb inversion. We opted for the *est-ce que* marker for sentences with bare objects (e.g., *Qu'est-ce que tu te demandes qui a lu?*), while we opted for the subject clitic inversion for

[5] For convenience, we used the same labels for the No Intervention and the Intervention configurations. In particular, the terms Inclusion and Inverse Inclusion for the No Intervention condition refer to the fact that the featural specification of the second wh-element is included in the featural specification of first wh-element, even though none of the wh-elements moves across the other one, as it is the case for the Intervention conditions.

Table 1
English translation of the French sentences in the 8 experimental conditions.

Structure	No Intervention	Intervention
Bare Identity	Who wonders who solved this problem?	What do you wonder who solved?
Inverse Inclusion	Who wonders which student solved this problem?	What do you wonder which student solved?
Inclusion	Which professor wonders who solved this problem?	Which problem do you wonder who solved?
Complex Identity	Which professor wonders which student solved this problem?	Which problem do you wonder which student solved?

sentences with lexically restricted objects (e.g., *Quel livre te demandes-tu qui a lu?*). This choice was motivated by informal judgments of naturalness gathered among native French speakers. In the No Intervention conditions we avoided the use of *est-ce que* as it would have involved the use of the ''que to qui'' rule ('que' becomes 'qui' in this context to alleviate the ECP violation in case of subject movement), which for some speakers is not fully natural.

Half of the experimental sentences contained *demander* (*wonder*) as main verb, whereas the other half contained *savoir* (*know*). All sentences containing *demander* were affirmative while only half of the sentences containing *savoir* were affirmative. We thus had 24 affirmative sentences and 8 negative sentences. The extracted wh-element was always inanimate, whereas the intervening wh-element was always animate. Eighty-eight grammatical filler sentences were added to the experimental items in order to introduce some variability in the structures. The fillers were various forms of questions (e.g., extractions from a declarative, simple questions, and additional embedding of the kind *Which problem do you think that Mary thinks that we could solve?*).

2.1.3. Procedure

The experiment was programmed with E-prime. Each sentence was presented on a computer screen one at a time. Participants were tested individually in experimental booths and asked to judge for the acceptability of the sentences on a 7-point Likert scale (1 corresponding to a totally unacceptable sentence and 7 to a perfectly acceptable sentence) by pressing one of the seven numbered buttons on the keyboard. Participants were first shown 3 examples of sentences (one corresponding to a totally unacceptable sentence, one to a nearly acceptable sentence and one to a perfectly acceptable sentence) and their respective ratings (1, 4 and 7), none of which involved islands. They were then presented with 10 training items in order to familiarize themselves with the Likert scale. There was no time constraint on responses. Three pauses were administered during the task. The whole session lasted about 30 min.

2.1.4. Data analyses

The whole data set consisting of 1280 data points was analysed without excluding any outliers. Data were analysed with mixed-effects models estimated with the lmerTest package (http://www.cran.r-project.org/web/packages/lmerTest//lmerTest.pdf) in the R software environment (R Development Core Team, 2011). The 2×4 model has Intervention and Structure as fixed factors and involves random intercepts for subjects and items.[6] The Satterthwaite approximation for degrees of freedom was used to estimate *p*-values. In order to explore the predictions of fRM bearing on the comparison across target structures involving intervention (i.e., wh-islands) controlled for their corresponding baselines without intervention, we ran six additional lmer models with, as fixed factors, Intervention and Structure Comparison, a 2-levels factor representing the six relevant contrasts between structures:

Model 1: Intervention (Intervention vs. No intervention) * Structure comparison 1 (Bare Identity vs. Inverse Inclusion)
Model 2: Intervention (Intervention vs. No intervention) * Structure comparison 2 (Bare Identity vs. Inclusion)
Model 3: Intervention (Intervention vs. No intervention) * Structure comparison 3 (Bare Identity vs. Complex Identity)
Model 4: Intervention (Intervention vs. No intervention) * Structure comparison 4 (Inverse Inclusion vs. Inclusion)
Model 5: Intervention (Intervention vs. No intervention) * Structure comparison 5 (Inverse Inclusion vs. Complex Identity)
Model 6: Intervention (Intervention vs. No intervention) * Structure comparison 6 (Inclusion vs. Complex Identity)

If the key differences between the four structures predicted by fRM are tightly linked to intervention, that is, if they attest to *intervention effects*, these differences should not show up in the corresponding baseline conditions with no intervention, or they should at least be reduced. That is, the predictions of the theory bear on the interactions between Structure Comparison and Intervention: if the theory predicts a stronger intervention effect in structure A than in structure B, this

[6] For the sake of comparability with data analyses from Hofmeister et al. (2013) and Atkinson et al. (2015), the 3-factors model involving Intervention, Lexical restriction of the extracted wh-element and Lexical restriction of the intervening wh-element is reported in the Appendix.

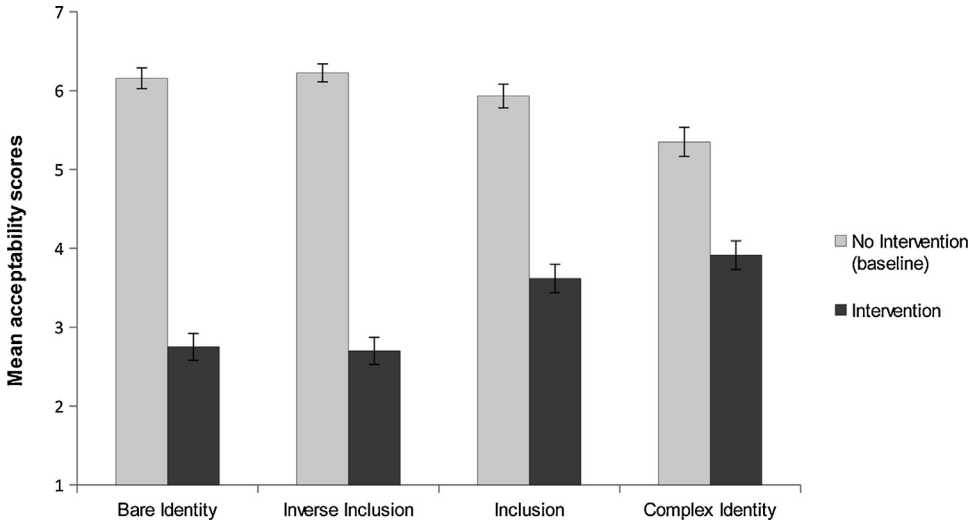

Fig. 1. Plot of the 8 individual structures on the 7-point Likert scale for Experiment 1.

prediction will translate into a bigger difference between A and B in the Intervention condition than in the No intervention, baseline condition. Under prediction (10), Inclusion is expected to show the weakest intervention effect; thus, Models 2, 4 and 6 containing Inclusion are all expected to show an interaction, with a difference between Inclusion and the comparison structure (Bare identity in Model 2, Inverse Inclusion in Model 4, Complex identity in Model 6) being greater in the Intervention condition than in the baseline. If proviso (11) adopted in prediction (10′) is correct, Complex Identity is expected to show a weaker intervention effect than Bare Identity and Inverse Inclusion, which will manifest by way of a significant interaction in Models 3 and 5. Finally, similar intervention effects in Complex Identity and Inclusion should be found under (10′), which should translate in terms of a lack of interaction in Model 6.

2.1.5. Results

Fig. 1 reports the mean acceptability scores and standard errors for the four sets in the target conditions with intervention and the baseline controls. Acceptability rates are significantly higher in the baseline conditions without Intervention ($M = 5.9$) than in the conditions with Intervention ($M = 3.2$) ($F = 1379.84$, $p < .001$), and they significantly differ across the four structures ($F = 4.55$, $p = .003$). The significant interaction between Intervention and Structure ($F = 47.09$, $p < .001$) shows that the effect of Structure is not the same for the conditions with intervention and for their corresponding baselines.

A summary of the interaction tests for the six models with Intervention and Structure Comparison as factors is reported in Table 2. Except for Model 1, all other interactions reached statistical significance, showing that the effect of structure is greater in the conditions with intervention than in their baseline controls. In sum, results show the following pattern of intervention effects (where ‘‘>’’ means ‘‘stronger intervention effect than’’ and ‘‘=’’ means ‘‘same intervention effect as’’):

Bare Identity = Inverse Inclusion > Inclusion > Complex Identity.

Table 2
Summary of the interaction tests for the six models in Experiment 1.

Variable	*F* value	*p*
Model 1: Bare Identity vs. Inverse Inclusion	0.39	.531
Model 2: Bare Identity vs. Inclusion	30.56	<.001
Model 3: Bare Identity vs. Complex Identity	95.02	<.001
Model 4: Inverse Inclusion vs. Inclusion	35.0	<.001
Model 5: Inverse Inclusion vs. Complex Identity	99.98	<.001
Model 6: Inclusion vs. Complex Identity	16.799	<.001

2.1.6. Discussion

Experiment 1 was designed to systematically assess predictions from fRM, recalled here for convenience:

(10′) Predictions of fRM for sentences out of context, if lexical restriction is a formal property defining feature sets, and proviso (11) is adopted (‘‘>’’ means ‘‘stronger intervention effect than’’ and ‘‘=’’ means ‘‘same intervention effect as’’):
Bare Identity = Inverse Inclusion > Inclusion = Complex Identity

Our results thus validate the prediction that Bare Identity and Inverse Inclusion show the same intervention effect, which in turn is stronger than Complex Identity and Inclusion. However, Complex Identity turned out to be less sensitive to intervention effects than Inclusion, which was not expected under fRM. We return to this unexpected result in the General Discussion.

Before extending the discussion of these findings further, it appeared important to determine whether these results are replicable and generalizable. With this aim in mind, we ran the same experiment by increasing the number of items to 320, which allows us to investigate if results from Experiment 1 are replicable, which is always desirable in experimental research, and generalizable to configurations with the same structure but different lexical materials. Moreover, in the next experiment we substituted a binary scale for the 7-point Likert scale, in order to investigate if these effects are captured by binary categorical judgments, which is interesting given that many models of grammaticality are categorical.

3. Experiment 2

3.1. Method

3.1.1. Participants

Twenty students from the University of Geneva participated in the experiment for course credit. All were between the ages of 18 and 26, and were native speakers of French. Participants recruited for this experiment had not taken part in Experiment 1.

3.1.2. Materials

The same variables manipulated in Experiment 1 were manipulated in the same design: (1) Intervention (Intervention vs. No Intervention), and (2) Structure (Bare Identity, Inverse Inclusion, Inclusion, Complex Identity). Stimuli consisted of 40 groups of sentences distributed across the 8 experimental conditions. We added 320 fillers to the experimental items, created by replacing the transitive embedded verb of experimental sentences (e.g., solve) with an intransitive verb (e.g., sleep). These fillers served as experimental sentences for another experiment not reported here. We thus had 320 experimental sentences and 320 fillers. The 640 items were split into two between-subjects lists containing 320 items each.

3.1.3. Procedure

The experiment was programmed with LimeSurvey (http://www.limesurvey.com/), a free and open-source software that makes it possible to create online surveys. The questionnaire was administered individually in experimental booths. Participants were asked to judge the acceptability of each sentence on a binary scale by clicking on ‘yes’ if the sentence was acceptable and on ‘no’ if it was unacceptable. Participants were first shown 9 example sentences (5 acceptable and 4 unacceptable) and their respective ratings, none of which involved islands. LimeSurvey displayed all sentences on the computer screen, one after the other. The whole session lasted about 30 min.

3.1.4. Data analysis

The whole data set consisting of 6400 data points was analysed without excluding outliers. Generalized mixed effects models for binomial distribution (GLME) were estimated in the R software environment (R Development Core Team, 2011). The fixed factors of the model are Intervention and Structure. The model involves random intercepts for subjects and items. Likelihood ratio tests were run in order to test for main effects and interactions. The additional theory-based analyses involving structure comparison were the same as described in Experiment 1.

3.1.5. Results

Fig. 2 illustrates the mean acceptability scores and standard errors for the four configurations in the target conditions with intervention and the baseline controls without intervention. The acceptability rates are significantly higher in the No Intervention conditions ($M = 0.95$) than in the Intervention conditions ($M = 0.28$) ($\chi^2(1) = 2007.2$, $p < .001$), and they significantly differ across the four structures ($\chi^2(3) = 218.92$, $p < .001$). The significant interaction between Intervention and Structure ($\chi^2(3) = 178.87$, $p < .001$) reveals that the effect of Structure is not the same for the conditions with intervention and for the baselines.

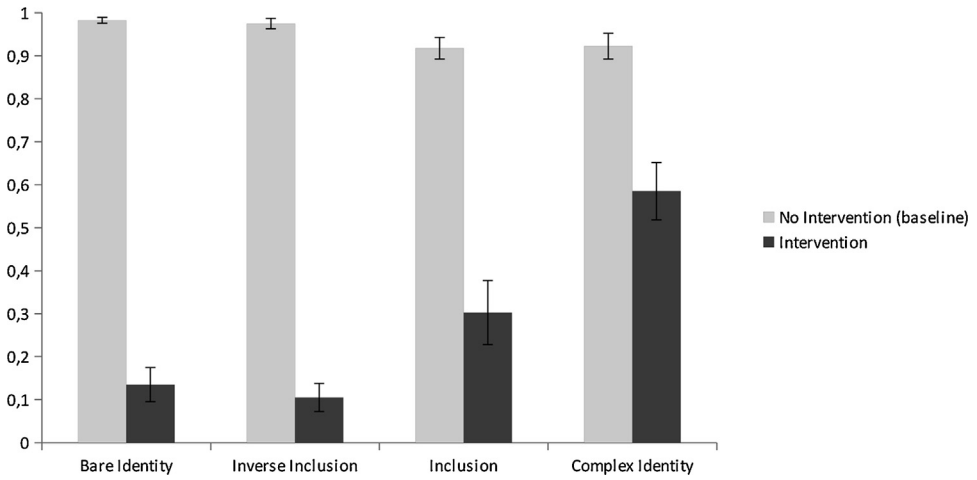

Fig. 2. Plot of the 8 individual structures on the binary scale for Experiment 2.

Table 3
Summary of the interaction tests for the six models in Experiment 2.

Variable	Estimate	Std. error	*z*	*p*
Model 1: Bare Identity vs. Inverse Inclusion	0.073	.564	0.130	.896
Model 2: Bare Identity vs. Inclusion	3.137	.487	6.436	<.001
Model 3: Bare Identity vs. Complex Identity	4.509	0.493	9.137	<.001
Model 4: Inverse Inclusion vs. Inclusion	3.248	0.460	7.056	<.001
Model 5: Inverse Inclusion vs. Complex Identity	4.591	0.462	9.922	<.001
Model 6: Inclusion vs. Complex Identity	1.716	0.339	5.052	<.001

A summary of the interaction tests for the six models with Intervention and Structure Comparison as fixed factors is reported in Table 3. Except in Model 1, all interactions were statistically significant, replicating results from Experiment 1. In sum, results show the following pattern of intervention effects (where ''>'' means ''stronger intervention effect than'' and ''='' means ''same intervention effect as''):

Bare Identity = Inverse Inclusion > Inclusion > Complex Identity

3.1.6. Discussion

Results from the second experiment replicate those from Experiment 1 with a different task involving binary judgments and with an increased number of items, attesting to the robustness of the data. In particular, we replicated the finding that configurations involving a lexically restricted extractee (Inclusion and Complex Identity) are less sensitive to intervention effects than configurations involving a bare extractee (Bare Identity and Inverse Inclusion), as predicted by fRM under proviso (11).

As discussed in section 1, two independent (but often co-occurring) factors could be responsible for the decrease of the intervention effect found for sentences involving lexically restricted extractees: one factor is lexical restriction, a structural feature, the other is the D(iscourse)-linked character of the extractee. In order to assess the potential role of D-linking, we ran a third experiment in which each sentence was preceded by a context story, in order to make bare wh-phrases linked to the discourse (i.e., D-linked).

4. Experiment 3

4.1. Method

4.1.1. Participants

Forty-nine students from the University of Geneva participated in the experiment for course credit. All were between the ages of 18 and 26, and were native speakers of French. Participants recruited for this experiment had not taken part in Experiment 1 or in Experiment 2.

Table 4
Example of context story and the corresponding questions for Experiment 3.

	Story	Questions
No Intervention	In the Physics department there is a blackboard on which each month a physicist writes a problem to be solved by a student. This month a student has solved the problem only 2 h after the problem was written on the blackboard. The professors are astonished! Among the physicists, Arnold knows that Frederic has solved the problem, because he told him. But Pascal, another physicist, is not aware of this. According to him, the two students likely to have solved the problem are Frederic and Nicolas. Pascal would really like to know who the author of the solution is and he will inquire!	1. Who is wondering who has solved this problem? 2. Who is wondering which student has solved this problem? 3. Which physicist is wondering who has solved this problem? 4. Which physicist is wondering which student has solved this problem?
Intervention	You are a student in Mathematics. Last week your professor assigned to your class 5 problems to be solved. You have solved the first four problems, but the fifth one was so difficult that you did not hand it out. When you arrived in the class this morning, the professor has announced that only one student has been able to solve the fifth problem. You would really like to know who the genius that has been able to do it is!	5. What did you wonder who solved? 6. What did you wonder which student solved? 7. Which problem did you wonder who solved? 8. Which problem did you wonder which student solved?

4.1.2. Materials

For this experiment the very same material as Experiment 1 was used, except that each question was preceded by a short context story. The context story was designed to promote D-linking by introducing a set of entities to which the wh-elements refer to. For each group of sentences we created two stories: one for the four No Intervention configurations and another one for the corresponding four Intervention configurations. Examples of context stories for one set of sentences are given in Table 4. Four lists were created such that each participant read a story only once. Each list contained 30 sentences in total (8 experimental sentences and 22 fillers).

4.1.3. Procedure

The procedure was the same as the one used in Experiment 1, except that each context story was presented on the computer screen before the presentation of the question. Participants pressed the space bar to switch from the story to the question.

4.1.4. Data analyses

One subject was excluded from the analysis since 75% of his scores were 2 standard deviations away from the mean of the population. Analyses were thus run on the 384 remaining data points. The same data analyses as Experiment 1 were conducted, with a first model involving Intervention and Structure followed by six models involving Intervention and Structure comparison. As stated in section 1.2, if the key factor underlying intervention effects is D-linking, all configurations will show the same intervention effect since, once in context, they all reduce to a configuration of Inclusion (i.e., [+Top,+Q]. . .[+Q]). This will translate into a lack of interaction between Intervention and Structure comparison in all six models.

4.1.5. Results

Fig. 3 reports the mean acceptability scores and standard errors for the four configurations in the target conditions with intervention and the baseline controls without intervention. The acceptability rates are significantly higher in the conditions without Intervention ($M = 5.9$) than in the conditions with Intervention ($M = 2.9$) ($F = 549.80$, $p < .001$), and they significantly differ across the four structures ($F = 4.49$, $p = .004$). The significant interaction between Intervention and Structure ($F = 12.88$, $p < .001$) shows that the effect of Structure is not the same for the conditions with intervention and for the baselines.

A summary of the interaction tests for the six models with Intervention and Structure Comparison as fixed factors is reported in Table 5. The interactions in Models 3–6 were significant, showing that the intervention effect is stronger in Inverse Inclusion than in Inclusion and Complex Identity, that it is stronger in Inclusion than Complex Identity, and that it is stronger in Bare Identity than in Complex Identity. The lack of interaction in Model 1 shows that the intervention effect in Bare Identity does not differ from that in Inverse Inclusion. The interaction in Model 2 did not reach significance level,

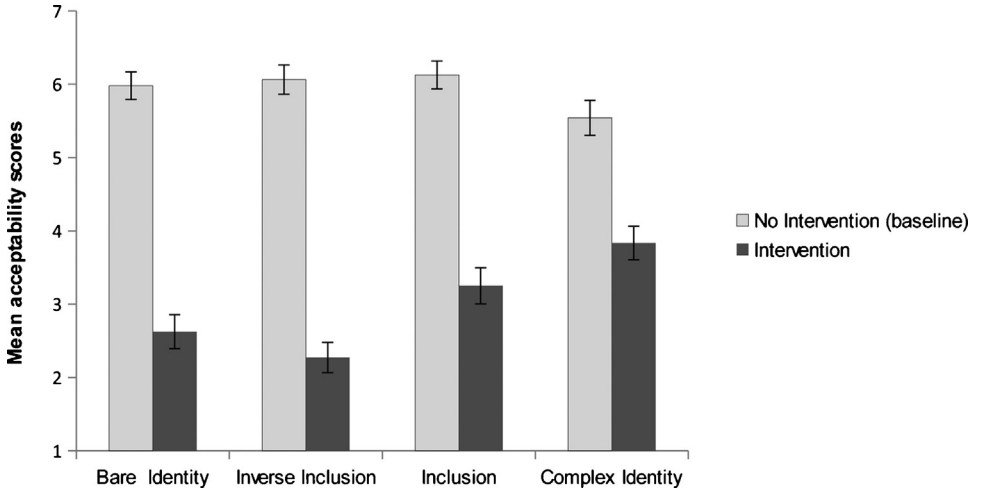

Fig. 3. Plot of the 8 individual structures on the 7-point Likert scale for Experiment 3.

Table 5
Summary of the interaction tests for the six models in Experiment 3.

Variable	*F* value	*p*
Model 1: Bare Identity vs. Inverse Inclusion	1.63	.203
Model 2: Bare Identity vs. Inclusion	1.794	.188
Model 3: Bare Identity vs. Complex Identity	19.252	<.001
Model 4: Inverse Inclusion vs. Inclusion	6.75	.01
Model 5: Inverse Inclusion vs. Complex Identity	8.349	.004
Model 6: Inclusion vs. Complex Identity	8.776	.003

suggesting that the intervention effect in Bare Identity is not statistically different from Inclusion, although a trend was found towards a stronger effect in Bare Identity. In sum, results from individual models show the following pattern (where ''>'' means ''stronger intervention effect than'' and ''='' means ''similar intervention effect as''):

Inverse Inclusion > Inclusion > Complex Identity
Inverse Inclusion = Bare Identity > Complex identity
Bare Identity = Inclusion

4.1.5.1. Effect of context: Comparison between Experiments 1 and 3.

We now focus on the second prediction discussed in section 1.2, which bears on the effect of context for each of the individual structures. As we have seen, if the feature of D-linking is responsible for the modulation of intervention effects in wh-islands, then we expect both Bare Identity and Inverse Inclusion to show a stronger intervention effect out of context than in context, while no effect should be found for Inclusion and Complex Identity. If, in contrast, the feature of lexical restriction is the crucial factor in the modulation of intervention effects, then the ratings for each of the four configurations should be identical in the presence and in the absence of context. With the aim of testing these predictions, we ran various models on the results of Experiments 1 and 3 merged together with Context as a between-participants factor.[7]

Fig. 4 summarizes results from Experiments 1 and 3. A first general model with Structure, Intervention and Context as fixed factors was run, in order to test for the interactions involving Context. Results showed a marginally significant interaction between Context and Intervention ($F = 3.09$, $p = .07$), attesting to a tendency for stronger intervention effects for structures in context (No intervention: $M = 5.92$; Intervention: $M = 2.99$) than for structures without context (No intervention: $M = 5.91$; Intervention: $M = 3.24$). The interaction between Context and Structure was not significant

[7] Note that the design obtained by merging experiments 1 and 3 is unbalanced, in that there are more data points in the No Context condition than in the Context condition. However, mixed-effects models are robust for unbalanced designs (e.g., Baayen et al., 2008).

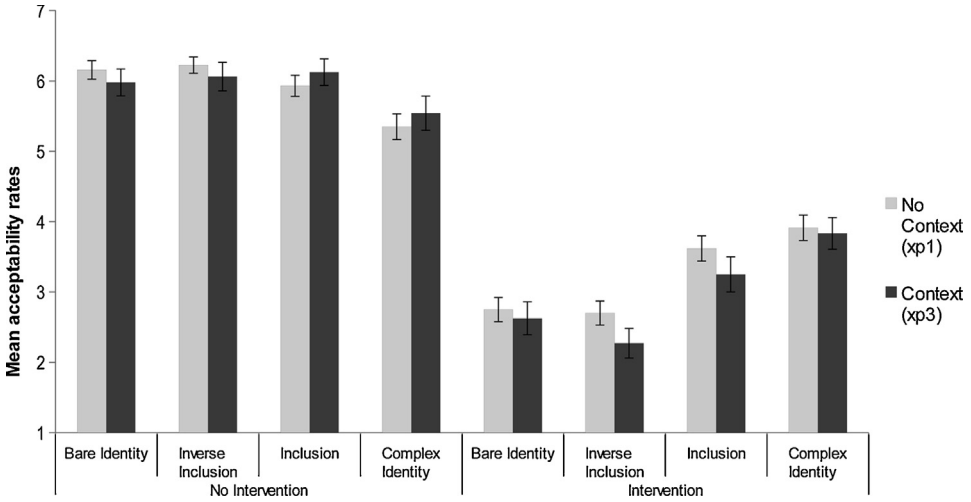

Fig. 4. Plot of the 8 individual structures on the 7-point Likert scale for Experiments 1 and 3. The first four structures on the left are conditions with No Intervention, while the four structures on the right are conditions with Intervention.

Table 6
Summary of the interaction tests for the four models in Experiments 1 and 3.

Variable	*F* value	*p*
Model 1: Bare Identity out of context * Bare Identity in context	0.03	.853
Model 2: Inverse Inclusion out of context * Inverse Inclusion in context	0.83	.362
Model 3: Inclusion out of context * Inclusion in context	3.208	.074
Model 4: Complex Identity out of context * Complex Identity in context	0.732	.393

($F = 0.98$, $p = .40$), revealing that the effect of structure is not different when there is context from when there is no context. The three-way interaction between Structure, Intervention and Context was not significant ($F = 0.71$, $p = .55$).

In order to test our predictions about the effect of context for each structure, four separate models were run with Intervention and Context as factors. If Context modulates the intervention effect in a particular structure, an interaction is expected between Context and Intervention. Model 1 tested the interaction between Intervention and Context for Bare Identity, Model 2 for Inverse Inclusion, Model 3 for Inclusion and Model 4 for Complex Identity. Results from the four models are reported in Table 6. None of the models reached statistical significance, suggesting that the presence of context does not significantly modulate intervention effects. Model 3 is nevertheless marginally significant ($p = .074$), showing that the intervention effect for the Inclusion configuration is actually stronger in the presence of a context (No intervention: $M = 6.125$; Intervention: $M = 3.25$) than in the absence of context (No intervention: $M = 5.93$; Intervention: $M = 3.61$).

4.1.6. Discussion

Results from Experiment 3 show some differences in the intervention effects of the four individual structures when sentences were presented in context. This is unexpected under the hypothesis that D-linking is the relevant factor in modulating intervention effects, since this hypothesis predicted that no difference should be found amongst the four structures (see section 1.2). Nevertheless, even though some of the structures showed different intervention effects with and without context, others did not, in contrast to what was found in Experiments 1 and 2, and in line with the D-linking hypothesis. We found that the intervention effect for Bare Identity did not differ from that for Inverse Inclusion nor from that for Inclusion. However, notice that the lack of difference between Bare Identity and Inverse Inclusion was also predicted by the lexical restriction hypothesis (10′). The finding that Bare Identity showed a similar intervention effect as Inclusion, however, was predicted by the D-linking hypothesis, but not by the lexical restriction hypothesis, which predicted a weaker intervention effect for Inclusion (10′). Yet, closer inspection of the effect of context on the configuration of Inclusion shows that the lack of difference between Bare identity and Inclusion in the presence of context is actually not due to a reduced intervention effect for Bare Identity in context, as predicted by the D-linking hypothesis, but to an increased intervention effect for Inclusion in context. Given that none of the hypotheses predicted that context should increase the intervention effect in Inclusion, we suggest that this data point is noise.

Finally, it is important to note that the finding that context did not modulate intervention effects (or at least not in any relevant way) bears on a null result, which may also be due to the fact that the context stories we used were not able to induce D-linking, a possibility that we cannot discard (see Sprouse's dissertation (2007) for a similar null result in Superiority violations).

5. General discussion

A series of three experiments explored the predictions from featural Relativized Minimality on the acceptability of wh-islands, with an experimental design allowing us to control for variables independent of the critical variables assumed by the theory. The theory assumes that when two elements that should enter into a local relation are separated by an element matching the featural specification of the elements it separates, an intervention effect arises and the sentence is perceived as degraded. According to the version of the principle we have adopted here, the features that are taken into account by the system are morphosyntactic features attracting movement. When the intervener fully matches the morphosyntactic featural specification of the target (configurations of Identity and Inverse Inclusion), the structure is expected to be severely ill-formed; when the intervener matches the specification of the target only in part (configuration of Inclusion) a milder deviance should arise.

Capitalizing on theoretical work suggesting that lexical restriction is a feature triggering movement and on experimental evidence showing that it plays a role in modulating intervention effects in object relative clause comprehension in acquisition (Friedmann et al., 2009, Munaro, 1999), we tested four set theoretic configurations: Bare Identity (both wh-elements are bare), Complex Identity (both wh-elements are lexically restricted), Inclusion (the extracted wh-element is lexically restricted but the intervening one is not) and Inverse inclusion (the extracted wh-element is not lexically restricted but the intervener is). Experiment 1 involved a 7-point scale acceptability judgment procedure, Experiment 2 involved a binary acceptability procedure, and Experiment 3 was identical to Experiment 1 except that sentences were preceded by a discourse context. Results from the three experiments are highly consistent, replicating major effects with different methods of data collection.

5.1. Featural RM and weak islands

In line with the predictions of fRM, results from Experiments 1 and 2 show that when the feature match is complete, as in Bare Identity and Inverse Inclusion, the intervention effect is stronger than when it is partial, as in Inclusion and Complex Identity. As discussed in section 1, the weaker intervention effects observed for complex wh-extractees could either be due to their lexical restriction or to their D-linked character. Adding a short context story before the sentences allowed us to tease apart the role of these two factors in Experiment 3. A formally bare wh-element can indeed be made contextually D-linked if uttered in a context presupposing a set of entities which constitute the range of the wh-variable. Results from Experiment 3 show that the presence of a context did not contribute to reduce the intervention effects in any of the four configurations (it even tended to increase it in the Inclusion configuration, a result which remains unexplained), and the context did not fundamentally modify the gradient of intervention effects observed in Experiments 1 and 2. On that basis, we conclude that D-linking is not the relevant factor in modulating intervention effects, which rather appear carried by the formal syntactic property of lexical restriction of complex wh-elements.

Another aspect of the data concerns the status of Complex Identity. All experiments showed a weaker intervention effect in the configuration of Complex Identity. Proviso (11) involved the assumption that Complex Identity can always admit a structural analysis with the lower complex wh-phrase attracted by a bare [+Q] head, as in (12). Therefore, if the featural specification which is taken into account is the one actually attracting movement, the case of Complex Identity can reduce to Inclusion, and thus display a weaker intervention effect than Bare Identity and Inverse Inclusion. The experimental data confirm these predictions. Nevertheless, the three experiments consistently show an even weaker intervention effect for Complex Identity than for Inclusion, in line with what was also observed in English (Atkinson et al., 2015). One way to address these data would be to amend the system of fRM in order to capture this case as well. One possibility is to broaden the class of features taken into account in the calculation of the constraint, currently taken to be morphosyntactic features triggering movement, to also include the lexical features expressed by the lexical restriction. Consider whatever featural property that defines the lexical noun *problem* and differentiates it from the lexical noun *student* in *Which problem do you wonder which student solved?*. If such features are taken into account in the calculation of locality, the set-theoretic relation between *which problem* and *which student* would become one of Intersection: the two phrases would have in common [+Q] and [+N], and would differ in all the lexical features which differentiate *problem* from *student*. This possible revision gains some plausibility in view of the fact that the set-theoretic relation of Intersection was shown to enhance comprehension of object relatives in children in Belletti et al. (2012) in comparison to Inclusion. Nevertheless, the extension of the class of relevant features beyond the class of morphosyntactic features triggering

movement raises problems elsewhere, for reasons discussed by Belletti and colleagues who argued in detail for a restrictive view of the feature calculus for locality.

The possibility that a wider range of features enters into the metric of similarity between the moved element and the intervener is actually closely aligned with recent theories of memory interference in sentence comprehension. In the next section, the relevant characteristics of these theories are sketched out in order to evaluate the extent to which they may account for our findings.

5.2. Cue-based memory models and similarity-based interference

A core feature of weak islands is that they contain two wh-elements that have to be integrated with the verb. Hence, when the relevant verb is reached, the representations of the arguments must be retrieved from memory in order to assign them their thematic roles and derive the proper interpretation. Content-addressable memory models provide a detailed analysis of how difficulties in parsing long-distance dependencies arise from constraints from memory retrieval mechanisms (e.g., Gordon et al., 2001, 2004; Lewis et al., 2006; McElree, 2000; Van Dyke and McElree, 2006). According to this framework, memory retrieval is driven by *cues*, which inform the parser about the features identifying the element that it is looking for and distinguishing it from other irrelevant representations in memory. Cues enable direct access to the relevant representation, i.e., the to-be-retrieved element, without the need to search through all the stored representations, which ensures a rapid recovery of relevant information stored in memory. However, rapidity comes at a cost, and there are well-understood deficiencies of this type of retrieval mechanism. A cue-driven operation can fail to recover the intended representation if the retrieval cues do not sufficiently overlap with the features with which the representation has been encoded into memory. Even when retrieval cues do sufficiently overlap with the representation, retrieval may still fail if those cues also match, even partially, the contents of other items in memory. This condition is known as *similarity-based interference* (e.g., Gordon et al., 2001; Nairne, 2002; Lewis et al., 2006; McElree, 2006). Early reports that the similarity between an element to be retrieved from memory and an intervening element caused increased reading times and increased comprehension errors came from studies on object relative clauses. These studies showed that reducing the similarity between the subject and the object, like when the subject is pronominalized while the object is a full NP, reduced processing times at the verb (e.g., Gordon et al., 2001, 2004). Similarity-based interference has also been attested in the processing of subject-verb long-distance dependencies: lower acceptability and longer reading times at the verb have been observed for cases in which the syntactic relation between the subject and the verb was interrupted by a relative clause containing an intervening grammatical subject, while no such effects have been observed when the interpolated material contained an intervening NP occupying an object position (e.g., Van Dyke and Lewis, 2003). Interestingly, other studies showed that semantically similar distractors could also generate interference (e.g., Van Dyke, 2007).

The finding that both syntactic and semantic cues play a role in determining interference provides us with a possible explanation for the weaker intervention effect observed for Complex Identity as compared to Inclusion. This could be due to the greater semantic distinctiveness, and therefore the lower similarity of lexically restricted wh-elements. Indeed, although in Complex Identity both wh-elements are identical regarding their syntactic features (i.e., [+Q, +N]), they are endowed with rich bundles of semantic features which all contribute to increasing their distinctiveness. If we consider a sentence like *Which problem do you wonder which student solved?* it is highly plausible that the retrieval of the extracted wh-element (*which problem*) is eased by the fact that *problems* can be solved, therefore providing a good semantic match to the object retrieval process triggered by the verb, whereas *students* cannot be solved, therefore providing a poor match, and reducing interference.

This hypothesis is supported by a series of studies conducted by Hofmeister and colleagues (e.g., Hofmeister et al., 2007, 2013), in which they found that increasing the syntactic and semantic complexity of the extracted element eases its retrieval and increases the acceptability of the sentence. For instance, using cleft constructions (e.g., *It was a(n) (alleged Venezuelan) communist who the members of the club banned from ever entering the premises*), they observed faster reading times after the embedded verb *banned* in sentences requiring the retrieval of complex objects (*an alleged Venezuelan communist*) as compared to those requiring the retrieval of simple objects (*a communist*) (Hofmeister et al., 2007). Interestingly, faster reading times after the verb coincided with slower reading times at the NP region (length being controlled). The authors took this result as indicative of a deeper encoding of complex NPs (*an alleged Venezuelan communist* vs. *a communist*), and argued that deeper encoding allowed for easier retrieval by providing distinctive retrieval cues.

In sum, psycholinguistic evidence supports the hypothesis that semantic richness plays a role in the processing of sentences involving memory retrieval, by reducing cue overlap and therefore similarity-based interference. The weaker intervention effect found for wh-islands containing lexically restricted wh-extractees may derive from the same principle of cue-based, direct access memory retrieval. In this view, interference is a function of the degree of similarity between the extractee and the intervener, where both syntactic and semantic features come into play in the similarity metric (see Villata

and Franck, 2015). We leave these possible developments within or outside the system of fRM for future work (on the comparison and the prospects for an integration between fRM and memory-based models see also Belletti and Rizzi, 2013; Santos, 2011).

5.3. Relativized Minimality and Superiority

It has been suggested that Superiority could be reduced to a case of intervention (ultimately RM or analogous principles in minimalist approaches), in which case the ungrammaticality of a sentence of the kind of (14a) lies in the movement of *what* across a similar element *who* (Chomsky, 1995). If so, we would expect lexical restriction to affect superiority violations configurations in the same way it affects configurations of intervention. However, this does not seem to be the case, as some recent findings by Hofmeister et al. (2013) attested. In four experiments (two acceptability judgment and two self-paced reading experiments), the authors explored the role of lexical restriction in structures involving Superiority violations (Pesetsky, 2000). They contrasted 4 conditions, illustrated in (14).

(14) a. Mary wondered what who read.
b. Mary wondered which book who read.
c. Mary wondered what which boy read.
d. Mary wondered which book which boy read.

In line with our results, the condition of Complex Identity (14d) always gave rise to the highest acceptability scores, whereas the condition of Bare Identity (14a) generated the lowest scores.[8] However, the two intermediate conditions showed a different profile to ours. In particular, if we focus only on Hofmeister et al.'s acceptability judgment experiments (leaving aside the self-paced reading experiments which are less directly comparable to our acceptability judgments experiments), their first experiment showed a reverse profile to ours, with what appears like Inverse Inclusion (14c) showing higher scores than what appears like Inclusion (14b), while their fourth experiment showed the two conditions to be on a par. Interestingly, an increase in acceptability ratings for condition (14c) as compared to condition (14b) was also attested in Hofmeister et al. (2007) as well as in German (Featherston, 2005). Hence, experimental evidence in English and German suggests the following pattern for Superiority violations (where ''<'' means ''lower acceptability rates than''): *Bare Identity* < *Inclusion* < *Inverse Inclusion* < *Complex Identity*, with Inverse Inclusion being more acceptable than Inclusion, in contrast to the higher acceptability of Inclusion found in weak islands.

So, the issue arises of understanding the reversed effect of lexical restriction found for cases of multiple wh-questions (Inclusion < Inverse Inclusion in Hofmeister et al., 2013) and for wh-islands (Inclusion > Inverse inclusion in our acceptability results). In that regard, it is interesting to explore the consequences of the classical analysis of multiple questions in Chomsky (1981), according to which the wh-element which is *in situ* at surface structure moves to the left periphery covertly in the syntax of Logical Form, in order to be placed in the appropriate scope position as binder of a wh-variable on the appropriate level of representation at the interface with semantic interpretation. The respective Logical Forms of (14) are thus the following, after covert movement:

(15) a. Mary wondered who_j $what_i$ [___$_j$ read ___$_i$].
b. Mary wondered who_j which $book_i$ [___$_j$ read ___$_i$].
c. Mary wondered which boy_j $what_i$ [___$_j$ read ___$_i$].
d. Mary wondered which boy_j which $book_i$ [___$_j$ read ___$_i$].

In this analysis, the locality violation determining the low acceptability of (14a) is determined by the covert movement of *who* crossing *what* moved to the left periphery in the overt syntax. If some version of this approach is on the right track, the reversal observed in Hofmeister et al.'s results (as well as in Featherston's study on German) with respect to ours is immediately understandable because covert movement inverts the polarity between Inclusion and Inverse Inclusion: (14c) superficially looks like Inverse Inclusion, but if locality is computed on the covert movement of *which boy*, the relevant representation is (15c), a case of Inclusion; similarly, (14b) looks superficially like a case of Inclusion, but if the relation is computed on (15b) after covert movement of *who*, the case really instantiates Inverse Inclusion. The higher acceptability of (14c) over (14b) is thus expected, and in line with our results on extraction from weak island.

[8] Notice that these results expressed in terms of acceptability rates easily translate in terms of intervention effect: high acceptability rates correspond to a weak intervention effect, while low acceptability rates correspond to a strong intervention effect. Hence, the sign '<' is used in this section to express low acceptability rates, while the corresponding sign '>' has been used in Experiments 1, 2 and 3 as expressing strong intervention effect.

6. Conclusion

The use of a formal methodology for the gathering of grammaticality judgments shows a considerable heuristic value and potential for the enrichment of the empirical basis in a complex domain like the violation of weak islands. On the one hand, this methodology has provided detailed evidence confirming basic predictions of fRM, thus offering further support for this approach to intervention locality; on the other hand, it has opened new questions, namely in connection with the relatively high acceptability of Complex Identity and the analysis of Superiority violation constructions, which suggests that new ideas and lines of inquiry are worth exploring to further increase the empirical adequacy of the theory of intervention effects.

Acknowledgements

We would like to thank Akira Omaki, who provided invaluable contribution to this work, as well as Martin Corley and Shravan Vasishth for their helpful comments. We would also like to thank Julien Chanal, Olivier Renaud and Pyeong Whang for helpful advices on statistical analyses, and Andres Posada for his technical support. We would also like to thank two anonymous reviewers as well as Jon Sprouse for his invaluable work of revision. We take complete responsibility for the content of the paper.

Appendix

Here we report results of a $2 \times 2 \times 2$ linear mixed effects models using the fixed factors Intervention, Lexical restriction of the extracted element (Wh1) and Lexical restriction of the intervening element (Wh2). For the sake of brevity (and since results obtained with this alternative method perfectly replicate those obtained in the 2×4 design), we report the alternative analysis for Experiment 1 only. The analyses for the remaining two experiments with a $2 \times 2 \times 2$ design are available from the authors.

Data were analysed with mixed-effects models estimated with the lmerTest package (http://www.cran.r-project.org/web/packages/lmerTest//lmerTest.pdf) in the R software environment (R Development Core Team, 2011). The fixed factors of the model are Intervention, Wh1 and Wh2. The model involves random intercepts for subjects and items.

Results from mixed models analyses revealed a main effect of Intervention, attesting to significantly higher rates for sentences in the No Intervention condition than for sentences in the Intervention condition ($\beta = -2.67$, $t = -37.146$, $p < .001$). Results also attested to a significant triple interaction between Intervention, Wh1 and Wh2 ($\beta = 0.993$, $t = 3.456$, $p < .001$). Thus, subsequent models were conducted separately for the No Intervention and the Intervention conditions.

Intervention conditions. A summary of the fixed effects is reported in Table A1.

Results showed a main effect of Wh1, attesting to a significant improvement in acceptability scores when Wh1 is lexically restricted as compared to when it is bare, showing that Complex Identity and Inclusion are globally rated higher than Bare Identity and Inverse Inclusion together. No main effect of Wh2 and no interaction between Wh1 and Wh2 were found. Subsequent models were conducted in order to provide a more fine-grained investigation of the relative acceptability of the 4 structures in the Intervention condition. A significant effect of Wh2 was attested when Wh1 is restricted ($\beta = 0.294$, $t = 1.898$, $p = .058$), showing that Complex Identity is rated higher than Inclusion, while no effect of Wh2 is attested when Wh1 is bare ($\beta = -0.050$, $t = -0.358$, $p = .721$), showing that Bare Identity and Inverse Inclusion are on par. In sum, results from individual models show the following pattern (where ''>'' means ''more acceptable than'' and ''='' means ''on a par with''): *Complex Identity > inclusion > Bare Identity = Inverse Inclusion*.

No Intervention conditions. Results are illustrated in Fig. 1, and a summary of the fixed effects is reported in Table A2.

Results reveal a main effect of Wh1, attesting to a significant improvement in acceptability scores when Wh1 is bare as compared to when it is lexically restricted, showing that Bare Identity and Inverse Inclusion are globally rated higher than Inclusion and Complex Identity. A main effect of Wh2 was also found, attesting to higher acceptability ratings when Wh2 is

Table A1
Fixed effects summary for Experiment 1 in the Intervention conditions.

Variable	Estimate	Std. error	*t*	*p*
Intercept	3.245	0.202	16.050	<.001
Wh1	1.041	0.105	9.900	<.001
Wh2	0.122	0.105	1.159	.247
Wh1 * Wh2	0.343	0.210	1.635	.103

Table A2
Fixed effects summary for Experiment 1 in the No Intervention condition.

Variable	Estimate	Std. error	*t*	*p*
Intercept	5.916	0.140	42.098	<.001
Wh1	−0.550	0.086	−6.397	<.001
Wh2	−0.256	0.086	−2.980	.002
Wh1 * Wh2	−0.650	0.172	−3.780	<.001

bare than when it is lexically restricted, showing that Bare Identity and Inclusion are globally rated higher than Inverse Inclusion and Complex Identity. A significant interaction between Wh1 and Wh2 was also found, with a significant effect of Wh2 when Wh1 is restricted ($\beta = -0.581$, $t = -4.475$, $p < .001$), attesting to higher acceptability for Inclusion with respect to Complex Identity, but no effect of Wh2 when Wh1 is bare ($\beta = 0.068$, $t = 0.668$, $p = .504$), showing that no significant difference is attested between Bare Identity and Inverse Inclusion. In sum, results from individual models show the following pattern (where ''>'' means ''more acceptable than'' and ''='' means ''on a par with''): *Bare Identity = Inverse Inclusion > Inclusion > Complex Identity*.

References

Alboiu, G., 2002. The Features of Movement in Romanian. Bucharest University Press, Bucharest.

Ambar, M., 1988. Para uma sintaxe da inversão sujeito-verbo em português. Estudos Lingüìsticos, Lisbon.

Atkinson, E., Aaron, A., Kyle, R., Omaki, A., 2015. Similarity of wh-phrases and acceptability variation in wh-islands. Front. Psychol. 6, 2048. http://dx.doi.org/10.3389/fpsyg.2015.02048

Baayen, R.H., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modelling with crossed random effect for subjects and items. J. Mem. Lang. 59, 390–412.

Bayer, J., Brandner, E., 2008. On wh-head-movement and the Doubly-Filled-Comp Filter. In: Chang, C.B., Haynie, H.J. (Eds.), Proceedings of the 26th West Coast Conference on Formal Linguistics. Cascadilla Proceedings Project, Somerville, pp. 87–95.

Belletti, A., Rizzi, L., 2013. Intervention in grammar and processing. In: Caponigro, I., Cecchetto, C. (Eds.), From Grammar to Meaning. Cambridge University Press, Cambridge, pp. 294–311.

Belletti, A., Friedmann, N., Brunato, D., Rizzi, L., 2012. Does gender make a difference? Comparing the effects of gender on children's comprehension of relative clauses in Hebrew and Italian. Lingua 122, 1053–1069.

Chomsky, N., 1981. Lectures on Government and Binding. Foris, Dordrecht.

Chomsky, N., 1995. The Minimalist Program. MIT Press, Cambridge, MA.

Chomsky, N., 2000. Minimalist inquiries: the framework. In: Martin, R., Michaels, D., Uriagereka, J. (Eds.), Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik. MIT Press, Cambridge, MA, pp. 89–155.

Chomsky, N., 2001. Derivation by phase. In: Kenstowicz, M. (Ed.), Ken Hale: A Life in Language. MIT Press, Cambridge, MA, pp. 1–52.

Chomsky, N., 2007. Approaching UG from below. In: Sauerland, U., Gärtner, H.-M. (Eds.), Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-semantics. Mouton de Gruyter, New York, pp. 1–29.

Cinque, G., 1990. Types of Ā-Dependencies. MIT Press, Cambridge, MA.

Comorovski, I., 1989. Discourse linking and the wh-island constraint. In: Proceedings of the Nineteenth Annual Meeting, NELS, .

Featherston, S., 2005. Magnitude estimation and what it can do for your syntax: some wh-constraints in German. Lingua 115, 1525–1550.

Friedmann, N., Belletti, A., Rizzi, L., 2009. Relativized relatives: types of intervention in the acquisition of A-bar dependencies. Lingua 119 (1), 67–88.

Gordon, P.C., Randall, H., Marcus, J., 2001. Memory interference during language processing. J. Exp. Psychol. Learn. Mem. Cogn. 27, 1411–1423.

Gordon, P.C., Randall, H., Marcus, J., 2004. Effects of noun phrase type on sentence complexity. J. Mem. Lang. 51 (1), 97–114.

Hofmeister, P., Florian, J.T., Sag, I.A., Arnon, I., Snider, N., 2007. Locality and accessibility in wh-questions. In: Featherston, S., Sternefeld, W. (Eds.), Roots: Linguistics in Search of its Evidential Base. Mouton de Gruyter, Berlin, pp. 185–206.

Hofmeister, P., Florian, J.T., Sag, I.A., Arnon, I., Snider, N., 2013. The source ambiguity problem: distinguishing effects of grammar and processing on acceptability judgments. Lang. Cognitive Proc. 28 (1), 48–87.

Huang, J., 1982. Logical relations in Chinese and the theory of grammar (Doctoral Dissertation). MIT Press, Cambridge, MA.

Lewis, R.L., Vasishth, S., Van Dyke, J., 2006. Computational principles of working memory in sentence comprehension. Trends Cogn. Sci. 10 (10), 447–454.

LimeSurvey Project Team/Carsten Schmitz, 2012. LimeSurvey: An Open Source Survey Tool. LimeSurvey Project, Hamburg, Germany. http://www.limesurvey.org

Lowder, M.W., Gordon, P.C., 2012. The pistol that injured the cowboy: difficulty with inanimate subject–verb integration is reduced by structural separation. J. Mem. Lang. 66 (4), 819–832.

McElree, B., 2000. Sentence comprehension is mediated by content-addressable memory structures. J. Psycholinguist. Res. 29, 111–123.

McElree, B., 2006. Accessing recent events. In: Ross, B.H. (Ed.), The Psychology of Learning and Motivation, vol. 46. Academic Press, San Diego.

Munaro, N., 1999. Sintagmi interrogativi nei dialetti italiani settentrionali. RGG Monographs, Unipress, Padova.

Nairne, S.J., 2002. The myth of the encoding-retrieval match. Memory 10, 389–395.

Pesetsky, D., 1987. Wh-in-situ: movement and unselective binding. In: Reuland, E., Meulen, A.T. (Eds.), The Representation of (In)definiteness. MIT Press, Cambridge, MA, pp. 98–129.
Pesetsky, D., 2000. Phrasal Movement and Its Kin. MIT Press, Cambridge, MA.
R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. The R Foundation for Statistical Computing, Vienna, Austria Available online at http://www.R-project.org/
Rizzi, L., 1990. Relativized Minimality. MIT Press, Cambridge, MA.
Rizzi, L., 1997. The fine structure of the left periphery. In: Haegeman, L. (Ed.), Elements of Grammar. Kluwer, Dordrecht, pp. 281–337.
Rizzi, L., 2001. On the position ''Int(errogative)'' in the left periphery of the clause. In: Cinque, G., Salvi, G. (Eds.), Current Studies in Italian Syntax, Essays offered to Lorenzo Renzi. Elsevier, Amsterdam, pp. 287–296.
Rizzi, L., 2004. Locality and the left periphery. In: Belletti, A. (Ed.), Structures and Beyond: The Cartography of Syntactic Structures, vol. 3. Oxford University Press, Oxford, pp. 223–251.
Rizzi, L., 2009. Movement and concepts of locality. In: Piattelli-Palmarini, M., Uriagereka, J., Salaburu, P. (Eds.), Of Minds and Language – The Basque Country Encounter with Noam Chomsky. Oxford University Press, Oxford, New York, pp. 155–168.
Rizzi, L., 2011. Minimality. In: Boeckx, C. (Ed.), The Oxford Handbook of Linguistic Minimalism. Oxford University Press, Oxford, New York, pp. 220–238.
Ross, J.R., 1967. Constraints on variables in syntax (Doctoral Dissertations). MIT Press, Cambridge, MA.
Santos, I.O., 2011. On Relativized Minimality, memory and cue-based parsing. IBERIA 3 (1).
Soare, G., 2009. The Syntax-Information Structure Interface: a comparative view from Romanian (Doctoral Thesis). University of Geneva.
Sprouse, J., 2007. A program for experimental syntax (Dissertation). University of Maryland, College Park, MD.
Sprouse, J., Schütze, C., Almeida, D., 2013. A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. Lingua 134, 219–248.
Sprouse, J., Caponigro, I., Greco, C., Cecchetto, C., 2016. Experimental syntax and the cross-linguistic variation of island effects in English and Italian. Nat. Lang. Linguist. Theory.
Starke, M., 2001. Move dissolves into merge: a theory of locality (Doctoral Dissertation). University of Geneva.
Szabolcsi, A., 2006. Strong vs. weak islands. In: Everaert, M., van Riemsdijk, H. (Eds.), The Blackwell Companion to Syntax, vol. 4. Blackwell Publishing, Malden, MA, USA, pp. 479–532.
Traxler, M.J., Pickering, M.J., McElree, Brian, 2002. Coercion in sentence processing: evidence from eye-movements and self-paced reading. J. Mem. Lang. 47 (4), 530–547.
Van Dyke, J.A., 2007. Interference effects from grammatically unavailable constituents during sentence processing. J. Exp. Psychol. Learn. Mem. Cogn. 33, 407–430.
Van Dyke, J.A., Lewis, R.L., 2003. Distinguishing effects of structure and decay on attachment and repair: a cue-based parsing account of recovery from misanalyzed ambiguities. J. Mem. Lang. 49, 285–316.
Van Dyke, J.A., McElree, B., 2006. Retrieval interference in sentence comprehension. J. Mem. Lang. 55, 157–166.
Villata, S., Franck, J., 2015. Does semantics similarity affect weak islands acceptability? In: Proceedings of IGG41. (in press).
Villata, S., Rizzi, L., Franck, J., 2015. Intervention effects in wh-islands and non-islands: experimental evidence. In: IGG41, .
Westergaard, M., Vangsnes, Ø., 2005. Wh-questions, V2, and the left periphery of three Norwegina dialect types. J. Comp. Ger. Lingusit. 8, 117–158.