# On Relativized Minimality, memory and cue-based parsing[1]

Iván Ortega-Santos
rtgsntos@memphis.edu
*University of Memphis*

**Abstract:** It is argued that Relativized Minimality (RM, Rizzi 1990 and Starke 2001, a.o.) is a conventionalized property of the grammar that is functionally grounded as a response to memory. In particular, it is shown that a cue-based retrieval parser (Van Dyke & Lewis 2003, Lewis & Vasishth 2005), according to which the integration of incoming words into existing interpretation is limited by retrieval interference and decay, can explain a number of features of RM. This analysis can capture both the properties that RM and general cognitive phenomena share (e.g., the similarity-based interference) and the features of RM which are specific to the grammar, (e.g., the role of c-command). Furthermore, issues such as crosslinguistic variation in RM effects and a number of arguments that have been put forward against coding islands in the grammar (e.g., Kluender 2004 and Hofmeister & Sag 2010) are addressed.

*Keywords:* Relativized Minimality, cue-based parsing, wh-islands, memory interference

**Resumen:** En este estudio se argumenta que la Minimidad Relativizada (MR; Rizzi 1990 y Starke 2001, entre otros) es una propiedad gramatical convencionalizada cuyo origen se encuentra en las propiedades de la memoria. En concreto, se muestra cómo el analizador sintáctico propuesto por Van Dyke y Lewis (2003) y Lewis y Vasishth (2005), de acuerdo al cual la integración de unidades lingüísticas en la interpretación existente se encuentra limitada por la interferencia en su recuperación y el desvanecimiento de la representación en la memoria con el paso del tiempo, puede explicar las características de la MR. En concreto, este análisis permite explicar las propiedades que la MR comparte con otros fenómenos

cognitivos en general, (por ejemplo, la interferencia en función de la similitud) y también las propiedades que son específicas de la gramática, (por ejemplo, mando-c). Adicionalmente, se discuten varios temas como la variación interlingüística de la MR o algunos argumentos existentes en contra de codificar las islas sintácticas en la gramática (por ejemplo, Kluender 2004 y Hofmeister y Sag 2010).

*Palabras clave:* Minimidad Relativizada, analizador sintáctico, islas-qu, interferencia en la memoria

**Resumo:** É defendido que Minimalidade Relativizada (MR, Rizzi 1990 e Starke 2001, entre outros) é uma propriedade convencionalizada da gramática que surge funcionalmente como uma resposta para a memória. Em particular, é demonstrado que um parser de recuperação baseado em pistas (Van Dyke & Lewis 2003, Lewis & Vasishth 2005), segundo o qual a integração das palavras de entrada numa interpretação existente é limitada por interferências na recuperação e deterioração, pode explicar várias propriedades da MR. Esta análise capta quer as propriedades que a MR e os fenómenos cognitivos em geral partilham (por exemplo, a interferência baseada em semelhanças) quer as propriedades da MR que são específicas da gramática (por exemplo, o papel de c-comando). São ainda abordadas questões como a variação interlinguística em efeitos de Minimalidade Relativizada e vários argumentos que têm sido apresentados contra ilhas de código na gramática (por exemplo, Kluender 2004 e Hofmeister & Sag 2010).

*Palavras-chave*: Minimalidade Relativizada, parsing baseado em pistas, ilhas wh-, interferência da memória.

One key feature of displacement or movement in natural language is the fact that this operation is subject to certain limitations. In particular, Chomsky (1964) and Ross (1967) showed that movement out of certain domains referred to as *islands*, e.g., sentential subjects, adjunct phrases, coordinate phrases or embedded interrogative clauses, is highly restricted, if possible at all. This observation has been crucial to the development of generative grammar (see Boeckx 2008 and references therein for discussion). Attempts to provide a syntactic treatment of these opaque domains (e.g. Chomsky's 1973 Subjacency) found an answer in research devoted to the relationship between the grammar and the parser. Specifically, researchers like Pritchett (1991), Kluender (2004) or Hofmeister & Sag (2010) have claimed that the opacity of islands follows from properties of the parser and are not part of the grammar (see Phillips 2006 for a critical review of these approaches). In turn, another line of research took grammatical constraints like Subjacency to have a functional motivation, found in the structure of the parser (Berwick & Weinberg 1984 and Weinberg 1988). Specifically, according to these researchers, parsing theory benefits from the constraints given by grammars, e.g., if it were not for locality principles, parsing would be difficult, if not impossible. On the other side, they claim that the pattern of locality constraints seems best explained as a fact about parsing, not a

fact concerning the grammar. In particular, parsing considerations impose a requirement that left context be literally finite when resolving long-distance dependencies. Ultimately, this requirement becomes grammaticized.

The purpose of this research is to argue for a theory of the latter kind, for a set of constraint on movement known as Relativized Minimality (RM; Rizzi 1990, 2001 and Starke 2001, a.o.). Following Ortega-Santos (2007b), it will be argued that RM is grammaticized as a real constraint that is functionally grounded as a response to memory. In particular, recent research has provided evidence for so-called cue-based retrieval parsing (e.g. Van Dyke & Lewis 2003 or Lewis & Vasisht 2005, a.o.). The main features of this framework are that the integration of words into the existing interpretation is limited by retrieval interference and decay. As will be discussed, this parser can explain a number of features of RM. Recent developments in research on language acquisition (Adani et al. 2010) and aphasia (Grillo 2008) support this interpretation. Furthermore, a number of arguments against reducing RM effects to parsing constraints which dismiss the role of the grammar (e.g., Kluender 2004 and Hofmeister & Sag 2010) are put forward.

Section 1 presents the relevant theoretical linguistics background, section 2 introduces the cue-based retrieval parser and evidence for it, and section 3 develops the proposal. Section 4 discusses a number of issues that provide evidence for the grammaticized nature of RM, (e.g., the relativization of RM to c-command paths or constraints on covert movement) in opposition to reductionist approaches that reduce islands to parsability considerations without the islands being coded in the grammar. Section 5 further compares reductionist theories of islands to the present approach.

## 1. Theoretical linguistics background

Within the biolinguistic perspective, it is argued that three factors exist that interact to determine (I-) languages attained: 'genetic endowment (the topic of Universal Grammar), experience, and principles that are language- or even organism-independent' (Chomsky 2005: 1). The latter factors, so called third factors, are principles not specific to the faculty of language - language-independent principles of data processing, structural architecture, and computational efficiency- which, nonetheless, affect the growth of language in the individual. A recurrent research topic within this framework, e.g., in the Minimalist Program (Chomsky 1995, 2005, etc.), is 'the extent to which apparent principles of Language […] are unique to this cognitive system or whether similar ''formal arrangements'' are found in other cognitive domains in humans or other organisms' (Chomsky 2005: 1-2). Still another more basic questions is how much of language can be given a principled explanation. Within this line

of research, RM is particularly interesting as it stands out as a computationally efficient principle: It helps reduce the number of possible structural relations that transformations may take and it has a 'least effort' flavor (see Chomsky & Lasnik 1995: 89-90). Similarly, RM is also relevant to study the relationship between linguistic theory and research on cognition. In the words of Rizzi (2004: 224):

> 'RM has desirable properties and appears to be a natural principle of mental computation. *It is the kind of principle that we may expect to hold across cognitive domains*: if locality is relevant at all for other kinds of mental computation, we may well expect it to hold in a similar form' (my emphasis).

In keeping with this view, it will be argued that language-independent properties of a so-called cue-based retrieval parser (e.g. Van Dyke & Lewis 2003 or Lewis & Vasisht 2005, a.o.) are responsible for the emergence of RM.

Originally put forward in Rizzi (1990), RM can be defined in the following way:

(1) *Relativized Minimality*

A movement operation cannot involve X and Y over a Z which is relevantly identical to Y in the configuration ...[X...[Z...[Y...]…]…]… if Z c-commands Y (Hornstein 2009: 35). [2, 3]

Early work on RM included the 'relevantly identical' part to capture the fact that heads were interveners for heads, A specifiers for A specifiers and A-bar specifiers for A-bar specifiers. Still, later developments relativized RM to features (Starke 2001 and Rizzi 2004). To simplify the discussion, I will assume the early definition.

RM explains the following contrasts found in (2). In such data, one can see that *how* can undergo long-distance wh-movement in English (see (2)a) and that one can have two wh-elements move to separate CPs, but only as long as the RM condition is met (cf. the contrast between (2)b and (2)c, respectively; data taken from Hornstein et al. 2005: 137):

(2)  (a)  [ how$_i$ did you say [ t$_i$ John fixed the car t$_i$ ] ]?
     (b)  [ who$_k$ [ t$_k$ wondered [ how$_i$ you fixed the car t$_i$ ] ] ]?

---

[2] I adopt this definition as it is succinct and consistent with current theoretical assumptions. In this sense I abstract away from irrelevant details, e.g., the fact that the original formulation involved the notion of government (Rizzi 1990), a theoretical construct abandoned in later research. See Chomsky's Minimal Link Condition (Chomsky 1995: 311) and Attract Closest (Chomsky 1995: 297), and the proposals of Chomsky (2000), Starke (2001) and Rizzi (2004) for other definitions. See also Fitzpatrick (2002) for an overview of similarity-based locality in syntax.

[3] See Lohndal (2008) for discussion of the relation between Minimality and Multiple Agree. In cases of Multiple Agree, a Probe agrees with all matching Goals simultaneously, in a configuration that resembles the ...X....Y...Z... of Minimality.

(c) *[ how$_i$ do you wonder [ who$_k$ [ t$_k$ [ fixed the car t$_i$ ] ] ] ]?[4]  (Hornstein et al. 2005: 136)

A similar argument can be made on the basis of head-movement and superraising, (3) and (4), respectively, though for the purposes of the discussion below, I will be concerned mostly with wh-movement:

(3)  (a)  They could have left.
     (b)  Could$_i$ they t$_i$ have left?
     (c)  *Have$_k$ they could t$_k$ left?                    (Rizzi 1990: 11)

(4)  (a)  It seems that John$_i$ is likely t$_i$ to win.      (Rizzi 1990 : 10)
     (b)  *John$_i$ seems that it is likely t$_i$ to win.

With regard to the role of c-command in the definition of RM, it is worth noting that the contrast in grammaticality of the following sentences can only be attributed to the different c-command relations between the wh-elements involved:

(5)  (a)  John wondered who books about what impressed.[5]

     (b)  John wondered what whose mother said.

(6)  (a)  *John wondered who what impressed.

     (b)  *John wondered what who said. (Hornstein 2009: 35)

*What* in (5)a and *whose* in (5)b do not c-command the base position of the other wh-element found in their respective sentences, and, as a consequence, no RM violation is found. In (6), the c-command condition in the definition of RM is met and, as a consequence, an RM violation occurs.

Within a derivational approach to syntax, RM is captured by means of the notion Attract Closest:

(7)  K attracts F if F is the closest feature that can enter into a checking relation with a sub-label of K (Chomsky 1995: 297).[6]

In this system, as well as in recent phased-based implementations (e.g.,

---

[4] As seen in this example, RM can explain the existence of certain islands, e.g., wh-islands. Needless to say, RM is not a theory put forward to account for *all* syntactic islands, e.g., Subject Islands. This simply reflects the fact that syntactic islands form a heterogeneous group (though see Starke 2001 for an attempt to treat all islands uniformly in terms of RM). See Boeckx (2008) for relevant discussion.

[5] Some informants find this sentence slightly less than perfect. The crucial point, though, is that (6a) is worse, N. Hornstein, p.c.

[6] The advantage of implementing RM in this way is that the effects of the RM constraint emerge locally as the derivation unfolds. For instance, given data like (i), there is no need to take the whole sentences into account to assess whether RM is violated (Reinhart 2006: 20-22):

   i. a.   [ It seems that [ Max$_i$ is certain [t$_i$ to arrive]]]

     b. *[ Max$_i$ seems that [ it is certain [t$_i$ to arrive]]]

   See Boeckx (2007:138-143) for further relevant discussion.

Chomsky 2000 and subsequent work), RM does not exist as an independent constraint, rather its effects are captured in other components of the system. Whereas an evaluation of these alternative implementations to enforce RM effects is beyond the scope of this research, it should be noted that RM effects are also conventionalized in theoretical approaches rejecting the existence of an RM constraint or filter and, therefore, the existence of this conventionalization calls for an explanation. For the sake of presentation I will refer to RM as being part of the grammar irrespective of whether it is coded as a constraint.

To sum up, plenty of evidence in favor of the role of RM in the grammar exists. Still, it is not clear what RM or its features (e.g., the similarity-based component of this constraint) ultimately follow from, an issue that will be addressed below.

## 2. Cue-based parsing and memory interference

Within the framework of cue-based retrieval parsing (e.g., Van Dyke & Lewis 2003, a.o.), the integration of incoming words into existing interpretation is limited by interference and decay. With regard to interference, it is assumed that the retrieval mechanism of the parser is content-addressable. This means that retrieval is carried out via a direct and parallel matching of all possible candidates against the retrieval cues. The more candidates match the retrieval cue, the less likely the cue is effective and the more difficult processing becomes, a phenomenon known as cue-overload. As far as the decay factor is concerned, every unit/item being processed is assumed to have an activation value that fluctuates over time as a function of usage history and time-based *decay*. The activation affects the probability and latency of retrieval, in that more active items show an advantage when compared to less active items. The advantage of this view of the parser is that its features, cue-based retrieval and decay, are well established in other domains of cognition (see Lewis 1998 and Lewis & Vasishth 2005 and references therein).[7] An example of this is found not only in interference effects in language but also in people's memory for visual stimuli (Chandler 1991) and motor skills (Adams 1987) and for facts accessed during mental skills such as mental arithmetic, (Campbell 1991). Furthermore, interference effects are also found in animal learning that employ Pavlovian conditioning paradigms (Bouton 1993).

---

[7] With regard to the issue of how these psychological hypotheses are instantiated in the brain, see Jones & Polk (2002), where an attractor network model which exhibits both time-based decay and similarity-based interference is put forward.

## 2.1. Evidence in favor of cue-based retrieval parsing

Van Dyke & Lewis (2003) put forward the cue-based retrieval parser and provide experimental evidence for it. Specifically, they conducted a self-paced reading task and a grammaticality judgment task involving the following structures:

(8) (a) The secretary forgot the student who was waiting for the exam WAS standing in the hallway.

(b) The secretary forgot the student who knew the exam was important WAS standing in the hallway.

These sentences show a garden path effect in that the DP [*the student* {*who was waiting for the exam / who knew the exam was important*}] is understood as the direct object of the verb *forget*, until the disambiguating point is reached, namely, *WAS*. At this point, reanalysis takes place so as to make the DP [*the student* {*who was waiting for the exam / who knew the exam was important*}] the subject of an embedded clause which is the direct object of *forget*. These conditions differ from one another in the number of items that cause interference at retrieval. In particular, two retrievals are required: one to break the object link between *forgot* and *student*, and another one to reattach the (temporal) object of *forgot* as the subject of *WAS*. In the case of the first cue, a verb taking a sentential complement (*forgot*) is to be retrieved, whereas in the case of the second cue, a nominative singular noun (*student*) is to be retrieved. Despite the fact that both sentences contain the same number of words and the same number of nouns in the ambiguous region, retrieval is subject to interference in (8)a and (8)b to a different extent. In (8)a, when retrieving *forgot*, the features of the cue partially overlap with the first *was*. In turn, when retrieving *student*, the features of the cue partially overlap with those of *who* and *exam*. This partial overlap contrasts with the situation in (8)b. Here, the cue to retrieve *forgot* partially matches the first *was* as in the previous case and, in addition, it perfectly matches *knew*, a verb that subcategorizes for a complement clause. In turn, the cue to retrieve *student* partially matches *who* as before and, in addition, it perfectly matches *exam*, a nominative singular noun. As a consequence, the retrieval interference theory predicts that the low interference condition in (8)a should be easier to repair than the high interference condition in (8)b. As Van Dyke & Lewis show, this is reflected in a higher accuracy of grammaticality judgments and shorter reading times for (8)a, the low interference condition, than for (8)b, the high interference condition.

With regard to the role of decay in this paradigm, assuming time-based decay of elements in the derivational space, in the examples above the activation of *the student* has decayed in contrast to the activation of *the exam*, which is more highly activated than *the student* due to the recency effect.

Consequently, *the exam* is more readily available (everything else being equal), a fact that adds to the processing difficulty of the structure.[8]

Subsequent work by Adani, van der Lely, Forgiarini & Guasti (2010) showed that matching gender and number features, that is to say, grammatical features, also affect performance of children speaking Italian (see also Adani 2008 for the same result concerning number in English; see Friedmann, Belletti & Rizzi 2009 for earlier discussion of RM effects in child language). In particular, children showed more comprehension errors when gender and number features were matched in center-embedded object relative clauses (see (9)a and (10)a vs. (9)b and (10)b, respectively):

(9) (a) Il leone      che il gatto     sta toccando è seduto per terra.
        The lion-SG that the cat-SG is touching   is sitting on the ground

     (b) Il leone       che i coccodrilli stanno toccando è seduto per terra.
         The lion-SG that the crocs-PL are     touching is sitting on the ground

(10) (a) Il gatto     che il   topo      sta lavando è   salito    sullo sgabello.
         The cat-M that the mouse-M is   washing has climbed onto the stool

     (b) Il gatto     che la   capra sta lavando è    salito    sullo sgabello
         The cat-M that the goat-F is   washing has climbed onto the stool

Adani et al. (2010), following Friedmann, Belletti & Rizzi (2009), capture the effect of matching grammatical features in performance by claiming that RM applies when processing these sentences. In particular, this constraint entails the computation of subset relations involving the features of the object and the subject in the relative clause, a costly computation (see Reinhart 2006). Children would apply a stricter version of RM due to immature computational resources, as opposed to lack of competence.[9] Adani et al.'s results are crucial from the current point of view as they show that interference effects are triggered by grammatical features of the kind syntax is assumed to manipulate, thus underscoring the plausibility of linking RM with memory interference. Specifically, in the spirit of Ortega-Santos (2007b), Adani et al. (2010: 2164)

---

[8] The high interference condition in (8b) of Van Dyke & Lewis (2003) is complex in a way that the low interference condition in (8a) is not (M. Wagers, p.c.). Specifically, the high interference condition includes a(n additional) clausal boundary that the low interference condition does not. Clausal boundaries can affect performance (e.g., see Kluender 1998 for some discussion). This criticism, nonetheless, does not apply to the experiments of Adani et al. (2010) to be introduced next.

[9] In addition, number dissimilarities had a stronger impact than gender in reducing intervention effects. This asymmetry between number and gender is interpreted as sensitivity to DP-internal structure: external and syntactically active features such as number reduce intervention in contrast to internal and (possibly) lexicalized features such as gender.

briefly discuss the close relation between cue-based parsing and RM in the following paragraph:

> (…) we would like to mention that [the] notion of *cue* is, in several ways, analogous to what theoretical linguists call *feature*. Hence, a formal model such as the one elaborated by Van Dyke and colleagues could also conceivably predict the effects presented in this paper. Therefore, a way to reconcile a syntactic locality principle such as RM and a formal computational model of sentence processing is foreseeable (author's emphasis). [10]

## 3. RM: a grammatical convention that is functionally grounded as a response to memory

There are at least 3 different possibilities concerning the relationship between cue-based parsing and RM: (i.) RM is part of UG and the similarities between RM and the dynamics of the cue-based parser are just a coincidence; (ii.) RM effects are processing effects not coded in the grammar; (iii.) RM is a grammatical convention that is functionally grounded as a response to cue-based parsing. Below, I will argue for a theory of the latter kind.

The dynamics of the cue-based retrieval parser and RM effects bear a striking resemblance to one another: both enforce similarity-based interference and the configurations where these effects are found, respectively, overlap to some extent. In particular, for a cue-based retrieval parser, the following linear sequence / linguistic input where $\beta$ is expected to trigger the retrieval of $\alpha1$ on the basis of some linguistic feature present in both $\alpha1$ and $\alpha2$ is predicted to be problematic (Lewis & Vasisht 2005: 19):

(11)   $\alpha1 \ldots \alpha2 \ldots \beta$

Why? If the content of $\alpha1$ and $\alpha2$ is similar, then given that retrieval is content-addressable, both $\alpha1$ and $\alpha2$ are possible candidates – they interfere with each other. All things being equal, memory retrieval will succeed for whichever of $\alpha1$ or $\alpha2$ has the highest activation value, namely, the most recent element due to role of time-based decay in memory, $\alpha2$.

Still, the RM configuration, (1), is a subset of the memory interference configuration, (11), in that the RM configuration is relativized to c-command

---

[10] An extensive review of the evidence in favor of cue-based parsing is beyond the scope of this section. See also McElree, Foraker & Dyer (2003), Lewis & Vasishth (2005), Van Dyke & McElree (2006) and Gordon et al. (2001, 2004, 2006) for further evidence in this regard. Furthermore, work by Wagers (2008) and Lau (2009) reveals that (a.) interference effects are most likely to be found when triggered by inherent properties of the units involved or in scenarios where reanalysis is necessary; and (b.) the predictive nature of the parser might be instrumental in reducing interference effects in a content-addressable memory.  The wh-property of wh-elements can be considered inherent and, consequently, I believe that interference is relevant for wh-islands as will be discussed below.

paths. In this sense, a cue-based retrieval parser gets close to deriving RM effects in that the standard RM configuration (...X...Z...Y...) is predicted to be difficult to parse, a processing difficulty that can be conventionalized in the grammar in the spirit of Berwick and Weinberg (1984) (see also Hawkins 1999 and Christiansen & Chater 2008 for relevant discussion).[11]

In keeping with this line of thought, research on acquisition and aphasia provides further support for linking RM and the properties of cue-based parsing. As stated before, Friedmann et al. (2009) and Adani et al. (2010) showed evidence for abnormally strict RM effects in child language. In a similar vein, Grillo's (2008) recent research on Broca's aphasia underscores the validity of RM as a psychologically valid constraint. Grillo develops an RM-based approach to comprehension deficits with movement-derived sentences in agrammatic Broca's aphasia: agrammatic aphasics cannot represent the full array of morphosyntactic features (in particular, scope and discourse features) associated with syntactic categories. This gives raise to Minimality effects in contexts where non-brain-damaged adult speakers represent the intervening element (Z) as distinct from the attractor/goal (X, Y) in the ...[X...[Z...[Y...]…]…]… configuration. This explains the comprehension deficits with movement-derived sentences in agrammatic Broca's aphasia. If the present approach linking memory interference (cue based overlap and time-based decay) is on the right track, it predicts that children and aphasics should be more sensitive to memory interference (cue overlap) than non-brain-damaged adult speakers. Indeed, both predictions are borne out, a fact that underscores the relationship between RM and the features of the cue-based parser. Specifically, Kail (2002) and Dempster (1992) provide evidence for the acute sensitivity to interference of young children when compared to adults.[12]

---

[11] Similarity-based interference can take place at encoding, storage or retrieval. Ortega-Santos (2007a) explored these first two components by linking RM and neural accommodation, a fatigue-like phenomenon whereby an exposure to an element of a certain kind inhibits the processing of another element of a similar kind (see Huber & O'Reilly 2003, a.o.). The present proposal follows Ortega-Santos (2007b) in putting the emphasis on retrieval interference (and time-based decay) instead, as memory research has provided evidence in favor of the pervasive effects of this kind of interference as opposed to the two other kinds (see Tehan & Humphreys 1996 and Suprenant & Neath 2009).

[12] As far as the acquisition of RM is concerned, following Hawkins (1999: 281-282) it is posited that: a. children will comprehend their input and postulate grammars in accordance with ease of processing; b. *positive evidence* will allow the children to go beyond such initial restrictive hypotheses. Note that unusually high sensitivity to interference effects would cause children to be conservative at initial stages of acquisition in accordance to the Subset Principle (Berwick and Weinberg 1984). See

Thompson-Schill, Jonides, Marshuetz, Smith, D'Esposito, Kan, Knight & Swick (2002) provide evidence that the left inferior frontal gyrus, which includes Broca's area, plays a key role in resolving memory interference (see also Kuhl & Wagner 2009 and reference therein), whereas Novick, Trueswell & Thompson-Schill (2005) and Schnur, Schwartz, Brecher & Hodgson (2006) present evidence that patients with damage to Broca's area show abnormally high interference effects in areas of cognition different from sentence processing, e.g., picture naming tasks. [13]

Note that the present view takes for granted that a full-fledged syntactic component including RM exists and that this system influences comprehension and production. Townsend & Bever (2001)'s analysis-by-synthesis provides this kind of framework for comprehension (see Bever & Poeppel 2010 for recent discussion). According to these researchers, humans basically understand everything *twice*, once on the basis of perceptual templates which assign likely interpretations to sentences by using a pattern completion system, and once again by the assignment of syntactic derivations. Specifically, the preliminary analysis of the sentence, as elements become available, yields a numeration used to construct a syntactic derivation of the sentence. This derivation is compared to the preliminary analysis. It is within this derivation that RM effects may arise in comprehension.[14] In turn, for the purposes of production, I assume that a full-fledged syntactic component feeds the parser/producer.

## 4. RM is part of the grammar

The view that RM is part of the grammar and that it is functionally grounded captures: (i.) the similarities between RM and the properties of the cue-based retrieval parser, which entails giving a principled explanation to a number of features of RM; (ii.) a number of features that a priori do not seem to follow from the properties of the parser. This section shows the advantages of this point of view by discussing the following properties of RM: the need for c-command in the conventionalization of RM, the existence of those Minimal Domains, where RM does not seem to apply (cf. Chomsky's 1995 Equidistance), the crosslinguistic variation in RM effects, Rizzi 's (1978) analysis of bounding

---

Omaki (2010) and references therein for discussion of conservatism in child language and the working memory capacity of children.

[13] Grodzinsky (2005) argues in favor of the role of Broca's area in the processing of syntactic movement. Note that this specialization is a priori not incompatible with the view developed below which links RM with general cognitive properties, while maintaining a degree of specificity of RM due to its grammatization.

[14] See also Phillips (1996) for an alternative framework in terms of left-to-right syntax where there is no distinction between the parser and the grammar.

nodes in Romance and its relation to said crosslinguistic variation, argument/adjunct asymmetries in wh-movement and the existence of constraints on covert movement (cf. Huang 1982). These properties are shown to follow from the grammar or the grammar in combination with the parser.

## 4.1. On the definition of RM

### 4.1.1. C-command: the parser vs. the grammar

As seen in the discussion around (5)-(6), c-command is a hallmark of RM, in contrast to the interference effects discussed in the parsing literature. In the spirit of Berwick and Weinberg's (1984) work, this sensitivity of RM to c-command can be derived as long as RM is taken to be part of the speakers' linguistics competence. Once interference and decay effects become part of the grammar, they are stored in the vocabulary of the grammar. Under standard assumptions, the vocabulary of the grammar does not include any notions of closeness in terms of linear order, but rather in terms of hierarchical closeness (e.g., Spec, Head relations, Head, Complement relations, etc.). Because of this, the sole manner the grammar has to express a configuration like ...X...Z...Y... is in terms of a structural notion, namely, c-command. [15] Note that this explanation is unavailable to pure processing theories of islands which deny the relevance of the grammar to capture island effects. As a consequence, such theories cannot capture contrasts as those in (5) and (6).

Given that interference and decay effects in (language) processing are not restricted by the c-command condition in contrast to RM, we expect a certain degree of modular independence between c-command and interference effects in the realm of wh-movement.[16] Inasmuch as such cases of interference do not violate a grammatical constraint (RM) (as stated before, the grammar cannot code an interference effect unless it is relativized to c-command) such effects are predicted to exist, but to be rarely reported as having an effect on the acceptability of the sentences. Indeed, such highly infrequent cases have been documented in the literature. An example of this is found in Aoun & Li (2003) who have pointed out that the following configuration is judged deviant in Lebanese Arabic, even though the c-command condition on RM is not met (see

---

[15] Arguably, dominance is still another relevant grammatical relation, in which case the present view could be extended to the A-over-A constraint. I leave this issue for future research. See Fitzpatrick (2002) and Hornstein (2009) for discussion of the relationship between the A-over-A constraint and RM.

[16] See also Rizzi (2004) for dissociation of the basic components of RM, namely, locality, c-command and identity, in human language. According to Rizzi, locality without c-command is found in gapping and in phonology. See section 5 for further discussion of these kinds of dissociations.

also Müller 2004 for some German and English cases):

(12)   *[$_{CP}$ $wh_1$ ... [$_{IP}$ ... [$_{island}$ ... $wh_2$ ...] ... $x_1$ ...]]

As stated, this pattern of interference rarely affects grammaticality in natural languages in keeping with the view that this is a pure processing effect that cannot even be grammaticized.[17]

## 4.1.2. Grammars do not count

Still another peculiarity of the RM definition is that just one *RM violation leads to ungrammaticality*, as suggested by the RM effects seen so far. One would like to know why this is the case, that is to say, why the RM constraint takes the form it does, given that the processing burden caused by time-based decay and cue-overlap does not force this kind of encoding. This feature of RM effects follows straightforwardly under the standard assumption that the grammar cannot count. A constraint stating that three or five RM violations lead to ungrammaticality cannot be encoded in the grammar (cf. Berwick and Weinberg's 1984 discussion of Subjacency). From the present perspective, the RM constraint takes the form it does because that is the only way the grammar can code it.

## 4.2. On Minimal Domains

A hallmark of RM is that its effects seem to be absent when elements in the same 'Minimal Domain' (MinD) –equidistant elements– are involved. Chomsky argues that the notion of closeness should be relativized to MinDs in the following way:

(13)   if β c-commands α and π is the target of raising, then β is closer to K than α unless β is in the same MinD as α or π.

In turn, the MinD of a head H is defined as the set of terms immediately contained in projections of H.[18] MinDs are needed to explain, for instance, (i.) object raising for case checking purposes (either overtly or covertly, depending on the language), where the object moves past the subject given the VP-Internal Subject Hypothesis (e.g., Koopman & Sportiche 1991); and (ii.) subsequent

---

[17] Still, under very specific circumstances, processing factors might affect the perception of grammaticality, as in the case of the classic garden path effects in (i):

    i.  The horse raced past the barn fell. (Bever 1970)

See Townsend & Bever (2001) for relevant discussion.

[18] The MinD is extended by head movement, though these extensions can be eliminated from the system (see Hornstein, Grohmann & Nunes 2005: ch.5 for detailed discussion).

subject raising to Spec,TP, e.g., in Icelandic.[19] As Hornstein (2009: ch.2) points out, the problem with the MinD domain view is that, given conventional assumptions, multiple Specs of the same projection are in c-command configurations, with one Spec c-commanding the other. As a consequence, Equidistance is an ad hoc stipulation (or else the c-command condition on Minimality is wrong). In this context, Hornstein puts forth an approach to Minimality that actually derives Equidistance while maintaining the c-command condition. According to Hornstein, c-command is relevant to Minimality because in order for grammars to implement Attract Closest / RM, they must be able to measure path lengths (lengths of movements). Unfortunately, as discussed before, grammars do not count under standard assumptions. Still, path lengths can be measured without counting by using Boolean measures: the relative size of two sets is fixed if one is a proper subset of the other. In the words of Hornstein (2009: 39):

> (...) what is so measured are paths, the set of Maximal Projections that dominate the launch site and the target. Grammars prefer those moves with the shortest "Boolean" paths. To be so comparable, the paths being compared must involve elements that c-command one another for failure to c-command results in paths that are not in subset relations and so are neither longer nor shorter than each other using a Boolean measure.

Hornstein illustrates the discussion with the following English and Icelandic raising structures and their derivations, respectively, where $T_0$ is the target of the movement of DP1:

(14)    a. John seems to Mary to be tall.
    a'. [TP2 T0 [VP seem [PP P DP2] [TP1 DP1 ….    *English – no RM violation*

    b. *Hestarnir virdast mer vera seinir.
        the-horses seem me-Dative to-be slow
    b'. *[TP2 T0 [VP seem DP2 [TP1 DP1 ….    *Icelandic – RM violation*

In (14)a', the path of DP1 is {TP1,VP,TP2}, whereas the path of DP2 is {PP,VP,TP2}. Neither is a subset of the other and so neither path is shorter than the other and no RM violation ensues. In (14)b', the path of DP1 is {TP1,VP,TP2}, whereas the path of DP2 is {VP, TP2}. Clearly, the path of DP2 is a subset of the path of DP1 and thus, it is shorter. As a consequence, the experiencer does not block the movement of the DP1 in (14)a' (English), but it does block the movement of this very DP in (14)b' (Icelandic).

As Hornstein notes, one consequence of his analysis is that MinD or Equidistance "exceptions" to Minimality immediately follow. Given the

---

[19] See, nonetheless, Koizumi (1995) for another plausible approach to these data, where Equidistance plays no role.

structure in (15), where T is the target, the paths of XP and YP are identical, {TP,BP}.

(15)    [TP T ... [BP XP [B′ YP [ B…]]]]

Thus, we can account for why Minimality does not hold for elements in the same MinD: such elements always traverse equivalent paths. Why? Given this rationale, multiple specifiers of a common head are elements in the same domain. Such elements always traverse equivalent paths and Minimality effects should not arise between them, irrespective of whether we look at the landing sites or launch sites or both.

To conclude: Whereas an evaluation of the analyses that have been put forward in the literature to measure distance is beyond the scope of this paper (see Fitzpatrick 2002 in this regard), Hornstein's approach is particularly relevant because, if correct, the existence of MinDs does not have to be stipulated under the view that RM is coded in the grammar. In contrast, these MinDs remain unexplained in a pure processing account of Minimality.

## 4.3. On the parameterization of RM effects [20]

RM effects show a certain degree of crosslinguistic variation. It is argued that this degree of variation is not due to the presence of RM in certain languages and lack of thereof in others, but rather the choice of the elements that count as interveners is parameterized. That is to say, in the spirit of Hawkins (1999: 267), it is considered that 'the conventions of grammars have been set differently [...] in response to relative complexity'. In the present proposal, cue overload and decay are based on the linguistic properties of the stimuli. Cues, that is to say, linguistic properties, are subject to parameterization. It is predicted that those strategies helping to make the retrieval cues unambiguous will help prevent similarity-based interference and allow for a more relaxed grammatization of interveners. Particular attention is paid to a classic example of parameterization, namely, Rizzi's parameterization of wh-islands in English and Romance Null Subject Languages. The current view allows for a reinterpretation of generalizations concerning the relationship between RM effects, on the one hand, and overt agreement marking (see Rizzi 1990 a.o.), on the other.

In particular, it is predicted that languages including overt verbal agreement or overtly case-marked wh-elements would have an advantage in processing terms (due to the richness of the corresponding retrieval cues, see

---

[20] As stated before, this research focuses on wh-movement. Uriagereka (forthcoming) notes that the present approach predicts the existence of crosslinguistic variation in RM effects in head movement and superraising as well. The prediction is borne out (see Vicente 2007 and Ura 2007 for details, respectively).

also Lewis 1996 and Arnon et al.'s 2006 related claims) and, thus, grammatization of what counts as an intervener might be less strict than in languages lacking such advantage, e.g., English.[21] The prediction is fulfilled: e.g., Rizzi (1978) notes that Null Subject Languages have some strategies to void islands/RM effects in contrast to strict word order languages as English (see also Jaeggli 1982: 154-156, Torrego 1984: 114 and Bosque & Rexach 2009: 467 for Spanish), a fact that he links to the rich agreement/flexible word order of these languages (cf. also Chomsky 2001:30 for the speculations that rich agreement may play a role in circumventing intervention effects of potential goals, as necessary to extract an element out of an island). For instance, both case-marking and agreement (or something close to resumption, given the availability of object clitics in Spanish as opposed to English) would play a role in explaining the following contrast in the existence of RM effects in English and Spanish (I abstract away from irrelevant details):

(16)  (a) *Who$_i$ do(n't) you know when$_k$ you saw t$_i$?
   (a') A quién$_i$ (no) sabes        cuándo$_k$ le viste t$_i$?
       to whom (not) know.2.SG when    him-saw.2.SG?

   (b) *Who$_i$ do(n't) you know how much$_k$ t$_i$ weights t$_k$?

   (b') Quién$_i$ (no) sabes        cuánto$_k$ t$_i$ pesa t$_k$?
       Who   (not) know.2.SG how much weights? (Torrego 1984: 114)

   (c) ?*What do(n't) you know how$_k$ to fix t$_x$?
   (c') Qué$_x$ (no) sabes        cómo$_k$ reparar t$_x$?
       What  (not) know.2.SG how   to-repair?

Whereas the exact characterization of the retrieval cues in both languages might be subject to debate, the intuition is clear: Generally speaking, Spanish has more information available at the point of retrieval, both because of the richer verbal agreement system and because of the richer case morphology which marks animate objects as in (16)a'. Arguably, this makes a difference in terms of the success of the corresponding retrieval cues - there would be less ambiguity at retrieval in this language than in English, a view that has received support from Arnon et al.'s (2006) crosslinguistic experimental research. This difference in linguistic input would affect the grammatization of what counts as an intervener.[22] If this view is on the right track, this provides further support for the current hypothesis.

---

[21] F. Adani (p.c.) further notes that in the wh-island cases under discussion not only input frequency may play a role but rather the different morpho-syntactic statuses of features/cues could be at stake. See fn. 9 for discussion of this kind of perspective.

[22] Other areas of the syntax of Romance which lack such advantages may show Minimality effects; e.g., see Rizzi (2004) for Italian cases concerning the distribution of adverbs. See also the discussion on further complexity of wh-movement in Romance in section 4.4.

The well-known insensitivity to islands that D-linked wh-phrases show can be tentatively viewed in this light as well. For example., one could claim that D-linked wh-phrases can be extracted out of wh-islands without a RM violation, due to their richness; that is to say, due to their having features that make their retrieval easy (see also Arnon et al. 2006 and Hofman & Sag 2010). Be that as it may, in terms of the grammar, the difference between D-linked and bare wh-phrases is a difference in c-command, as suggested by the semantics of D-linking, and this has an effect on the computation of RM (Grohmann 2000; see Lasnik & Uriagereka 2005:ch. 3 for discussion).[23, 24]

## 4.4. Bounding nodes in Romance

As discussed in the previous subsection, Null Subject Romance languages and rich agreement/case marking languages generally seem to be less sensitive to wh-islands than English. However, a close look at the theoretical literature on islands reveals a slightly different picture. Indeed, Romance Null Subject Languages, like Italian or Spanish, have been argued to tolerate the crossing of one wh-island, but not two. To be more precise, according to Rizzi's (1978) parameterization of bounding nodes in English and Romance Null Subject Languages, S is a bounding node in English, in contrast to S' in Italian or Spanish (see also Torrego 1984 for Spanish). Given that Subjacency (Chomsky 1973) forbids movement from crossing two bounding nodes, the presence of a wh-element in any of the intermediate Comps in (17) in English would trigger ungrammaticality (Comps that have an influence on the grammaticality of wh-movement are indicated in bold):

(17)    *English*

[Comp3  [IP3 ..[**Comp2**  [IP2..... [**Comp1**  [IP1 ....wh-.....]]]]]]

---

[23] The fact that resumption helps avoid RM / interference effects, as stated in the theoretical literature (e.g., Boeckx 2003 and references therein) can be understood as still another strategy that helps reduce interference and decay effects. Still, the experimental literature has called those ameliorating effects into question (e.g., Alexopoulou and Keller 2002), possibly because the processing of resumptive pronouns adds also complexity to the structure in contrast to gaps (Hawkins 1999:265).

[24] Note that further crosslinguistic variation may result from independent properties at the level of the grammar or the parser, e.g., across languages wh-movement may target CP or IP (see Richards 1997 for discussion). Furthermore, cue overlap and decay are not the only processing costs associated with filler-gap dependencies, e.g., see Kluender (2004), Gibson (1998) or Boston (2010) for discussion. There is no implication that those other factors do not play a role in the grammatization of RM, but rather that the role of cue overlap and decay is particularly relevant.

In turn, in Romance Null Subject Languages the presence of a wh-element in the first Comp is irrelevant. The presence of a wh-element in subsequent Comps would trigger ungrammaticality:

(18)   *Romance Null Subject Languages*

[Comp3  [IP3 ..[**Comp2**  [IP2..... [Comp1  [IP1 ....wh-.....]]]]]]

The Spanish data can be illustrated in (19)a and (19)b. In the former sentence, the first Comp that the moved wh-element encounters is filled by another wh-element. In contrast, in the latter sentence the first and second Comps that the moved wh-element encounters are filled by another wh-element:

(19)   (a) A quién$_i$ (me dijiste    que) (no)  sabes [Comp1 cuándo$_k$  le    viste t$_i$]?
           to whom$_i$ (me told.2SG that) (not) know.2SG    when     him$_i$ saw.2SG?

       (b) (?)A quién$_i$  me dijiste [**Comp2** por qué (no) sabes [Comp1 cuándo le viste t$_i$]]?
           to whom$_i$ me told.2SG         why   (not) you-know    when  him saw.2SG?

The contrast in (19) is fairly subtle. Furthermore, Bosque & Rexach (2009: 460) include the following example, which illustrates the configuration in (18) and yet is perfect:

(20)   Ese  es el  profesor del que no  sé           cómo no se te ocurre    qué  decir.
        That is the teacher   of  that not know.1SG  how   not to-you-occurs what to-say

Why should Romance languages show such subtle contrasts as in (19), contrary to what the current approach would seem to predict? The answer might lie again in the properties of the parser. As stated previously, the success of the processing procedure depends on the activation value of the representations involved. This activation value is subject to time-based decay. Wh-elements undergoing fairly *long* long-distance wh-movement, skipping intermediate comps as in (19)b, might be subject to decay to a greater extent than wh-elements which stop at intermediate comps as in (19)a.[25] As a consequence, wh-movement as in (19)b correlates with higher processing costs (see also Gibson 1998), a fact that might affect the perception of the grammaticality of the resulting structures.[26] In this sense, the apparent counterexamples that motivated Rizzi's (1978) analysis of bounding nodes in

---

[25] See Franck, Lassis, Frauenfelder & Rizzi (2006) for evidence concerning the relevance of intermediate traces or copies in language processing.

[26] Torrego (1984) focuses on the interaction between subject-verb inversion and successive cyclic wh-movement providing evidence for Rizzi's parameterized theory of bounding nodes (see above). Nonetheless, those effects show variability within the Spanish language (e.g., see Baković 1998). See also Jaeggli (1985) and Fitzpatrick (2002) for further discussion on the nuances of wh-movement in this language.

Romance languages can be reconciled with the present proposal, at least tentatively.

## 4.5. Argument/adjunct asymmetries with regard to wh-movement

It is well-known that both wh-adjuncts and wh-arguments interfere with one another, due to their common wh-feature. Still, wh-arguments and wh-adjuncts pattern slightly differently in that wh-adjuncts have a more restricted distribution than wh-arguments (see Rizzi 1990 and Lasnik and Saito 1992 for an overview on this topic). This can be seen in the following contrast in Spanish, where in (21)b *cómo* 'how' cannot be interpreted in the embedded clause:

(21)  (a) Qué  no  sabes       cómo responder?
          What not know-you how   to-answer?

      (b) Cómo no  sabes       qué   responder?
          How   not know-you what to-answer?

This does not seem to be related to the properties of a cue-based retrieval parser and, hence, it needs to be explained. Under the assumption that adjuncts head their own projection above the position where arguments are generated, Superiority/Minimality forces the adjunct to move first, hence explaining these contrasts (Eguren & Fernández Soriano 2004 and references therein). Be that as it may, there are some asymmetries between arguments and adjuncts for the parser: arguments correlate with richer retrieval cues than adjuncts because of their close relation with predicates. Displaced arguments, in contrast to adjuncts, also provide indirect evidence for the retrieval site in the form of a missing obligatory constituent (Hofmeister & Sag 2010: 33; see also Lee & Thompson 2010 for evidence that there are greater processing costs associated with adjuncts than with arguments). Therefore, both the grammar and the parser conspire so that adjuncts are predicted to have a more restricted distribution than arguments.[27]

## 4.6. On constraints on covert movement – Lasnik's (1999) challenge

Research on wh-in situ languages has revealed that at least a subset of wh-in situ languages are sensitive to wh-islands. The purpose of this section is

---

[27] In turn, subject / complement asymmetries with regard to wh-movement remain mysterious under the present approach. According to Kluender (2004) subjects and in general elements at clausal edges are associated with high processing costs. At the same time, crosslinguistically, subjects are more frequently associated with morphological agreement than objects, meaning the retrieval cues for subjects would be richer.

to discuss the relevance of these facts in establishing whether RM is grammaticized or not. [28]

Building on this observation that certain wh-in situ languages are subject to locality constraints (e.g., Japanese), Lasnik (1999) challenged reductionist processing theories of islands: If an overt wh-island violation gives rise to ungrammaticality because it is hard to relate the moved element to its canonical surface position, e.g., in (22), why should the same constraint hold true in wh-in situ languages, e.g., in the Japanese counterpart found in (23), where the wh-element remains in situ and its integration into the structure is easy?

(22)   ??What do you remember where we bought?

(23)   ??[nani-o     doko-de   katta    ka] oboete-iru        no?
        What-Acc where-At  bought  Q   remember-Prog Q [29]   (Watanabe 2001: 205)

In (22), limited memory resources to access intervening referents at clause edges or to keep the fillers active could affect the perception of grammaticality and explain the data without recourse to the grammar (see Kluender 2004 and Hofmeister & Sag 2010 and references therein), but (23) remains mysterious in this light. Under the present view, data as (23) can be explained by adopting the independently developed proposal that in situ wh-elements move covertly (Huang 1982). The LF representation of (23) is (24) under the intended reading:

(24)   [CP nani$_i$-o [IP pro$_{you}$ [CP doko-de$_j$ [IP pro$_{we}$ t$_i$ t$_j$ katta]   ka] oboete-iru]        no
        what-Acc                 where-at                 bought Q   remember-Prog Q

Following that line of thought, speakers reconstruct the LF with its covert movement operations within the anlysis-by-synthesis framework that has been adopted and RM effects arise. In a similar vein, Watanabe's (2001) proposal allows for a more up-to-date version of this analysis in keeping with current syntactic theorizing. According to Watanabe, languages allowing for in situ wh-elements involve movement in overt syntax, though what would move in such cases is not the whole wh-element but a null operator. Accordingly, (25) is represented as (26) in overt syntax:

(25)   Boku-wa [CP [IP John-ga     nani-o     katta]  ka] shiritai.
        I-top            John-Nom what-Acc bought Q   want-to-know
        'I want to know what John bought.'

(26)   Boku-wa [CP Op$_i$ [IP John-ga [t$_i$ nani]-o katta] Q] shiritai.

If Watanabe's independently motivated analysis is correct, at some point in the processing of wh-in situ elements, speakers of Japanese-type of languages

---

[28] A detailed evaluation of proposals dealing with wh-in situ is beyond the scope of this section (see Watanabe 2001, Cheng 2003a, 2003b and references therein for an overview).

[29] In the intended reading *nani-o* 'what-Acc' takes matrix scope.

would have to reconstruct such overt movement, at which point RM effects may arise. Consequently, one can conclude that the existence of locality restrictions on wh-in situ languages does not posit a challenge for the current approach to RM, understood as a grammaticized form of interference, in contrast to reductionist processing theories of islands. [30]

To sum up the discussion, it has been argued so far that the main properties of RM, as presented in section 1, can be accounted for if RM is a conventionalized property of the grammar that is functionally grounded as a response to memory. Specifically, a number of features of RM can be derived; for example, the fact that RM is similarity-based, its relativization to c-command, the lack of Minimality effects in certain MinDs (cf. Chomsky 1995), the parameterization of RM effects or constraints on covert movement (cf. Huang 1982). Some of these properties are particularly relevant in that a priori they do not seem to straightforwardly follow from the cue-based retrieval parser approach to RM. Nonetheless, it has been shown that such cases follow from either independent properties of the grammar, the parser itself or a combination of both. Consequently, these cases are compatible with the present approach, in contrast to a reductionist processing theory of islands, where the conventions of the grammar, e.g., c-command, play no role.

## 5. A comparison with processing theories of islands

Processing theories of islands are particularly appealing in that the explanation is cost-free: The processing constraints argued to play a role in these theories are independently attested in non-island contexts, in contrast to competence theories. In fact, as Hofmeister & Sag (2010: 402) note 'if grammaticality differences are to explain the processing differences in islands, then some secondary explanation must be invoked to account for the same processing differences in syntactic contexts that do not contain violations of any known grammatical constraint'.

Furthermore, according to Hofmeister & Sag, island constraints are:

> (...) arbitrary in the sense that they bear no relationship to other constraints, emanate from no general principles of language, and have no relevance or parallel outside language. In short, syntactic island constraints mark

---

[30] Under this view which accepts a role of the grammar in explaining the data of wh-insitu languages, unselective binding (Pesetsky 1987) is also an option as suggested by crosslinguistic variation in wh-in situ languages. This mechanism licenses in situ wh-elements which are not subject to islands, e.g., wh-arguments in contrast to wh-adjuncts in Chinese. Allegedly, the choice of unselective binding vs. movement correlates with other properties in these languages, e.g., morphological differences (see Watanabe 2001 and references therein for details).

islands as special within the domain of language, and even more particularly, within the domain of linguistic dependencies. Consequently, island constraints offer little insight into anything about language or cognition, except islands themselves. (Hofmeister & Sag 2010: 406).

However, under the present view which takes RM to be grounded in processing factors (interference and decay), which are not unique to language, these criticisms do not apply or else apply only to a very limited extent. Furthermore, linguistic intervention effects of the RM-kind are not only attested in islands, but also in a diverse number of domains as head movement, (3), NP movement (see the cases of superraising in (4)), agreement (see Holmberg & Hróarsdóttir 2003) and in phonology (see Rizzi 2004 and, in particular, Nevins 2010, a.o.). In these cases we have a Probe Goal system where Goals are targeted on the basis of their features (cf. the content-address component of the cue-based parser and memory in general) and where intervention is determined in terms of feature overlap (cf. cue-overlap). In turn, c-command is relevant for the syntax, though not for the phonology (see Nevins 2010). Moreover, various processing costs postulated by syntactocentric reductionist theories of islands (e.g., processing cost derived from processing referential elements or elements at the edge of CP, etc., as in Kluender 2004) do not play a role, at least, in head movement and in phonology and, therefore, cannot be extended to such a diverse set of data.

## 6. Conclusion

I have claimed that RM (Rizzi 1990 a.o.) entered the grammar due to the cue-based retrieval nature of the parser (e.g., Van Dyke & Lewis 2003 or Lewis & Vasisht 2006, a.o.), whose properties are well attested in other domains of memory and cognition. One of the hallmarks of this kind of parser is that cue-overlap –a factor arguably present in the RM configuration– gives rise to memory interference (see Anderson & Neely 1996, a.o.). The resulting processing burden is amplified by the role of time-based decay of the elements being processed. As a response to these properties of the parser, RM emerges 'for free' in the grammar of natural language. That is to say, RM follows from a so-called 'third factor', namely, the properties of memory, which are language independent. This is a welcome result in that it shows a certain degree of convergence between theoretical linguistics and processing-oriented disciplines. Future work needs to broaden the crosslinguistic data considered and to explore the possibility of applying the present framework to the A-over-A constraint or to locality in other linguistic domains, e.g., in phonology.

## Appendix: On proactive and retroactive interference

Research on learning and forgetting –another area where cue-based retrieval has been applied successfully – has shown that there are two kinds of interference: proactive and retroactive interference. Proactive interference refers to previously learned material hurting our memory for more recently learned items. Retroactive interference refers to impaired memory performance on target items caused by learning new material between the initial encoding of those target items and the final test. It seems that the RM configuration can be equated with retroactive interference. The following question then suggests itself: if interference indeed became grammaticized, why was it that retroactive interference entered the grammar as opposed to proactive interference? This state of affairs actually follows from the very nature of interference effects: it has been shown that retroactive interference yields stronger intervention effects than proactive interference (cf. Anderson & Neely 1996: 251). This accords well with what we find in the grammar, which only bars the former kind of interference. This is probably related to the following factor: the RM and the retroactive interference configuration constitute an environment where *both* proactive and retroactive interference will be found and, therefore, it is predicted to be more problematic than proactive interference on its own. In particular, in the RM configuration X...Z....Y...., X causes proactive interference at the time Z is processed. At the time Y is processed, we find retroactive interference, whereas decay effects cause X to be less accessible. If this discussion is correct, it explains why it is that retroactive interference entered the grammar as opposed to proactive interference. Inasmuch as one can provide a rationale for the resulting state of affairs concerning the relationship between the grammar and memory interference, this discussion constitutes still another argument for the view that RM is indeed a grammaticized subcase of interference.

### REFERENCES

Acheson, Daniel J. & Maryellen C. MacDonald. 2005. A Role for Phonological Information in Relative Clause Processing. Poster presented at the 17th Annual CUNY Conference on Human Sentence Processing, Tucson, AZ.

Adams, Jack. A. 1987. Historical review and appraisal of research on the learning, retention and transfer of human motor skills. *Psychological Bulletin* 101. 41–74.

Adani, Flavia, 2008. The role of features in relative clause comprehension: A study of typical and atypical development. Milan, Italy: University Milano-Bicocca dissertation.

Adani, Flavia, Heather K.J. van der Lely, Matteo Forgiarini & Maria Teresa Guasti. 2010. Grammatical feature dissimilarities make relative clauses

easier: A comprehension study with Italian children. *Lingua* 120. 2148–2166.

Alexopoulou, Theodora & Frank Keller. 2002. Resumption and locality: A crosslinguistic experimental study. In Mary Andronis, Erin Debenport, Anne Pycha & Keiko Yoshimura (eds.), *Papers from the 38th meeting of the Chicago Linguistic Society*. 1–14. Chicago: University of Chicago Press.

Anderson, Michael C. & James H. Neely. 1996. Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (eds.), *Memory handbook of perception and cognition*. 237–313. San Diego, CA: Academia Press.

Aoun, Joseph & Yen-hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar: The diversity of wh-constructions*. Cambridge, MA: MIT Press.

Arnon, Inbal A., Neal Snider, Philip Hofmeister, T. Florian Jaeger & Ivan A. Sag. 2006. Cross-linguistic Variation in a Processing Account: the Case of Multiple Wh-questions. To appear in the *Proceedings of BLS 32*.

Baković, Eric. 1998. Optimality and inversion in Spanish. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough?* 35-58. Cambridge, MA: MIT Press.

Berwick, Rorbert & Amy Weinberg. 1984. *The grammatical basis of linguistic performance*. Cambridge, MA: MIT Press.

Bever, Thomas G. & David Poeppel. 2010. Analysis by synthesis: A (re)Emerging program of research for language and vision. *Biolinguistics* 4, 174–200.

Boeckx, Cedric. 2003. *Islands and chains*. Amsterdam: John Benjamins.

Boeckx, Cedric. 2007. *Understanding minimalist syntax: Lessons from locality in long-distance dependencies*. Malden: Blackwell.

Boeckx, Cedric. 2008. Islands. *Language and Linguistics Compass* 2. 151-167.

Boeckx, Cedric & Kleanthes K. Grohmann. 2007. Putting phases in perspective. *Syntax* 10, 204–222.

Bosque, Ignacio & Javier Gutiérrez-Rexach. 2009. *Fundamentos de sintaxis formal*. Madrid: Akal.

Boston, Marisa Ferrara. 2010. The role of memory in superiority violation gradience. *Proceedings of the 2010 Workshop on Cognitive Modeling and Computational Linguistics, ACL 2010*. 36–44. Uppsala, Sweden: Association for Computational Linguistics.

Bouton, Mark E. 1993. Context, time and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin* 114. 80–99.

Campbell, Jamie I. D. 1991. Conditions of error priming in number-fact retrieval. *Memory and Cognition* 19. 197–209.

Chandler, Christopher C. 1991. How memory for an event is influenced by related events: Interference in modified recognition tests. Journal *of Experimental Psychology: Learning, Memory, and Cognition* 17. 115–125.

Cheng, Lisa L.-S. 2003a. Wh-in-situ. *Glot International* 7 (4). 103–109.

Cheng, Lisa L.-S. 2003b. Wh-in-situ, Part II. *Glot International* 7 (5). 129–137.

Chomsky, Noam. 1964. *Current issues in linguistic theory*. The Hague, The Netherlands: Mouton.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*. 232–286. New York, NY: Holt, Rinehalt, and Winston.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist Inquiries. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. 89–156. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael J. Kenstowicz (ed.), *Ken Hale: A life in language*. 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36. 1–22.

Chomsky, Noam & Howard Lasnik. 1995. The theory of principles and parameters. In Noam Chomsky, *The minimalist program*. 15-127. Cambridge, MA: MIT Press.

Christiansen Morten H. & Nick Chater 2008. Language as shaped by the brain. *Behavioural and Brain Sciences* 31. 489–558.

Dempster, Frank N. 1992. The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review* 12 (1). 45–75.

Eguren, Luis & Olga Fernández Soriano. 2004. *Introducción a una sintaxis minimista*. Madrid: Gredos.

Engdahl, E.1985. Parasitic gaps, resumptive pronouns, and subject extractions. *Linguistics* 23. 3–44.

Fitzpatrick, Justin M. 2002. On minimalist approaches to the locality of movement. *Linguistic Inquiry* 33 (3). 443–463.

Franck, Julie, Glenda Lassi, Ulrich H. Frauenfelder & Luigi Rizzi. 2006. Agreement and movement: a syntactic analysis of attraction. *Cognition* 101. 173–216.

Friedmann, Naama, Adriana Belletti & Luigi Rizzi. 2009. Relativized relatives. Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119. 67–88.

Gibson, Edward. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition* 68. 1–76.

Gordon, Peter C., Randall Hendrick & Marcus Johnson. 2001. Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition* 27. 1411–1423.

Gordon, Peter C., Randall Hendrick & Marcus Johnson. 2004. Effects of noun phrase type on sentence complexity. *Journal of Memory and Language* 51. 97–114.

Gordon, Peter C., Randall Hendrick, Marcus Johnson & Yoonhyoung Lee. 2006. Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory and Cognition* 32 (6). 1304–1321.

Grillo, Antonio. 2008. Generalized minimality: Syntactic underspecification in Broca's aphasia. Utrecht: LOT.

Grodzinsky, Yosef. 2005. Syntactic dependencies as memorized sequences in the brain. *Canadian Journal of Linguistics* 50. 241–266.

Grohmann, Kleanthes. 2000. Proliphic peripheries: a radical view from the left. College Park, MD: University of Maryland dissertation.

Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75. 244–285.

Holmberg, Anders & Thorbjörg Hróarsdóttir. 2003. Agreement and movement in Icelandic raising constructions. *Lingua* 113. 997–1019.

Hofmeister, Philip & Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86. 366–345.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and universal grammar*. Cambridge: Cambridge University Press.

Hornstein, Norbert, Kleanthes Grohmann & Jairo Nunes. 2005. *Understanding minimalism*. Cambridge: Cambridge University Press.

Huang, James C.-T. 1982. Logical relations in Chinese and the theory of grammar. Cambridge, Ma: MIT dissertation.

Huber, David. E. & Randall C. O'Reilly. 2003. Persistence and accommodation in short-term priming and other perceptual paradigms: temporal segregation through synaptic depression. *Cognitive Science* 27. 403–430.

Jaeggli, Osvaldo. 1982. *Topics in Romance syntax*. Dordrecht: Foris.

Jaeggli, Osvaldo. 1985. On Certain ECP Effects in Spanish. Ms., USC.

Jones, Matt & Thad A. Polk. 2002. An attractor network model of serial recall. *Cognitive Systems Research* 3. 45–55.

Jonides, John & Derek E. Nee. 2006. Brain mechanisms of proactive interference in working memory. *Neuroscience* 139. 181–193.

Kail, Robert. 2002. Developmental change in proactive interference. *Child Development* 73 (6). 1703–1714.

Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. In Peter Culicover & Louise McNally (eds.), *Syntax and semantics* 29, 241–279. San Diego, CA: Academic Press.

Kluender, Robert. 2004. Are subject islands subject to a processing account? In Vineeta Chand (ed.), *Proceedings of 23rd West Coast Conference on Formal Linguistics*. 475–499. Somerville, MA: Cascadilla Press.

Koizumi, Masatoshi. 1995. Phrase Structure in Minimalist Syntax. Cambridge, MA: MIT dissertation.

Koopman, Hilda J. & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85. 211–258.

Kuhl, Brice & Anthony D. Wagner. 2009. Forgetting and retrieval. In Gary. G. Berntson & John T. Cacioppo (eds.), *Handbook of neurosciences for the behavioral sciences*, vol. 1. 586–606. Hoboken, NJ: John Wiley and Sons.

Lasnik, Howard. 1999. On the locality of movement: Formalist syntax position paper. In Michael Darnell, Edith Moravscik, Michael Noonan, Frederick J. Newmeyer & Kathleen M. Wheatley (eds.) *Functionalism and formalism in linguistics*, vol. 1: General papers (SLCS 41). 33–54. Amsterdam: John Benjamins.

Lasnik, Howard & Mamoru Saito. 1992. *Move α: Conditions on its applications and outputs*. Cambridge, MA: MIT Press.

Lasnik, Howard, & Juan Uriagereka with Cedric Boeckx. 2005. *A course in minimalist syntax: Foundations and prospects*. Malden, MA: Blackwell.

Lau, Ellen Frances. 2009. The predictive nature of language comprehension. College Park, MD: University of Maryland dissertation.

Lee, Jiyeon & Cynthia K. Thompson. 2010. Real-time production of arguments and adjuncts in normal and agrammatic speakers. To appear in *Language and Cognitive Processes*.

Lewis, Richard. 1996. Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic research* 25. 93–115.

Lewis, Richard. 1998. Interference in working memory: Retroactive and proactive interference in parsing. Ms., Ohio State University.

Lewis, Richard L. & Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29. 1–45.

Lohndal, Terje. 2008. Anti-identity effects. Ms., University of Oslo.

McElree, Brian, Stephani Foraker & Lisbeth Dyer. 2003. Memory structures that subserve sentence comprehension. *Journal of Memory and Language* 48. 67–91.

Müller, Gereon. 2004. Phrase impenetrability and wh-intervention. In Arthur Stepanov, Gisbert Fanselow & Ralf Vogel (eds.), *Minimality effects in syntax*. 289–326. Berlin: Mouton de Gruyter.

Nevins, Andrew. 2010. *Locality and vowel harmony*. Cambridge, MA, MIT Press.

Novick, Jared M., John C. Trueswell & Sharon L. Thompson-Schill. 2005. Cognitive control and parsing: reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, and Behavioral Neuroscience* 5. 263–281.

Omaki, Akira. 2010. Commitment and flexibility in the developing parser. College Park, MD: University of Maryland dissertation.

Ortega-Santos, Iván. 2007a. On neural accommodation and relativized minimality effects. Talk at Biolinguistic Investigations, FUNGLODE, Dominican Republic.

Ortega-Santos, Iván. 2007b. On the emergence of relativized minimality. Talk at the First Meeting of the Arizona Linguistics Circle (ALC 1), University of Arizona. https://umdrive.memphis.edu/rtgsntos/Minimality2007.pdf (1 November, 2011).

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric J. Reuland & Alice G. B. ter Meulen (eds.), *The representation of (in)definiteness*. 98–129. Cambridge, MA, MIT Press.

Phillips, Colin. 1996. Order and structure. Cambridge, MA: MIT dissertation.

Poeppel, David & David Embick. 2005. The relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-first century psycholinguistics: Four cornerstones*. 1–16. Mahwah, NJ: Lawrence Erlbaum.

Pritchett, Bradley. 1991. Subjacency in a principle-based parser. In Robert Berwick, Steve Abney & Carol Tenny (eds.), *Principle-based parsing: computation and psycholinguistics*. 301–345. Dordrecht, The Netherlands: Kluwer.

Reinhart, Tanya. 2006. *Interface strategies: Optimal and costly computations*. Cambridge, MA: MIT Press.

Richards, Norvin W. 1997. What moves when where in which language. Cambridge, MA: MIT dissertation.

Rizzi, Luigi. 1978. Violations of the wh-island Constraint in Italian and the subjacency condition. *Montreal Working Papers in Linguistics* 11. 155–190.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2001. Relativized minimality effects. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*. 89–110. Oxford: Blackwell.

Rizzi, Luigi. 2004. Locality and left periphery. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures*, vol. 3. 223–251. New York: Oxford University Press.

Ross, John R. 1967. Constraints on variables in syntax. Cambridge, MA: MIT dissertation. [Published 1986 as *Infinite syntax!* Norwood, NJ: Ablex.]

Schnur, Tatiana T., Myrna F. Schwartz, Adelyn Brecher & Catherine Hodgson. 2006. Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language* 54. 199–227.

Sprouse, Jon, Matt Wagers & Colin Phillips. Submitted. A test of the relation between working memory capacity and syntactic island effects.

Sprouse, Jon. 2007. A program for experimental syntax. College Park, MD: University of Maryland dissertation.

Starke, Michal. 2001. Move dissolves into merge: a theory of locality. Geneva, Italy: University of Geneva dissertation.

Surprenant, Aimee M. & Ian Neath. 2009. *Principles of memory*. New York: Psychology Press.

Tehan, Gerald & Michael S. Humphreys. 1996. Cuing effects in short-term recall. *Memory and Cognition* 24. 719–732.

Thompson-Schill, Sharon L., John Jonides, Christy Marshuetz, Edward. E. Smith, Mark D'Esposito, Irene P. Kan, Robert T. Knight & Diane Swick. 2002. Effects of frontal lobe damage on interference effects in working memory. *Cognitive, Affective, and Behavioral Neuroscience* 2 (2). 109–120.

Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15. 103–129.

Townsend, David J. & Thomas G. Bever. 2001. *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Ura, Hiroyuki. 2007. Returning to superraising. *Departmental Bulleting Paper* 56. 59–74.

Uriagereka, Juan. Forthcoming. *Multiple spell-out and the minimalist program*. New York: Oxford University Press.

Van Dyke, Julie A. & Richard L. Lewis. 2003. Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of

recovery from misanalyzed ambiguities. *Journal of Memory and Language* 49. 285–316.

Van Dyke, Julie A. &  Brian McElree. 2006. Retrieval interference in sentence comprehension. *Journal of Memory and Language* 55. 157–166.

Vicente, Luis. 2007. *The syntax of heads and phrases: A study of verb (phrase) fronting*. Utrecht: LOT.

Wagers, Matthew Webb. 2008. The structure of memory meets memory for structure in linguistic cognition. College Park, MD: University of Maryland dissertation.

Watanabe, Akira. 2001. Wh-in-situ Languages. Mark Reuben Baltin, Chris Collins (eds.), *The handbook of contemporary syntactic theory*. 203–225. Malden, MA: Blackwell.

Weinberg, Amy. 1988. *Locality principles in syntax and in parsing*. Cambridge, MA: MIT Press.

Xiang, Ming, Brian W. Dillon & Colin Phillips. 2007. Interference effects across dependency types: ERP evidence. Ms., University of Maryland.

**Iván Ortega-Santos**
rtgsntos@memphis.edu
*University of Memphis*