SEMANTIC SIMILARITY EFFECTS ON
WEAK ISLANDS ACCEPTABILITY

Sandra Villata
Université de Genève

Julie Franck
Université de Genève

Abstract: *Two acceptability judgment experiments were conducted to investigate the role of similarity in weak islands environments. We tested predictions stemming from a narrow approach to similarity (Featural Relativized Minimality) and from a broad approach to similarity (Cue-based memory model) in generating intervention effects. According to Featural Relativized Minimality, only syntactic features triggering movement modulate similarity amongst elements, while according to the Cue-based memory model both syntactic and semantic features modulate similarity. Results show significant effects of both animacy and reversibility of thematic roles, two features that do not trigger movement. Therefore, our results support the broad approach to similarity.*

Keywords: *Relativized Minimality, Cue-based memory model, similarity, syntactic features, semantic features, interference, animacy, reversibility of thematic roles*

1. Introduction

It is well known that human language allows for long-distance dependencies, as shown in (1):

(1) Which problem$_i$ do you believe that the linguist solved __$_i$?

The internal object of the verb *solved* has been separated from the verb that selects it and moved to the front of the sentence leaving a gap in its original position (Chomsky 1981, 1986). One of the most interesting properties of long-distance dependencies is their potential unboundedness: in (2) for example, grammar imposes no limit on the length of the interpolated materials between the moved element and its gap:

(2) Which problem$_i$ do you believe that the linguist [and the psycholinguist working together on long-distance dependencies] solved __$_i$?

However, although unbounded, long-distance dependencies are not unconstrained. A first type of constraint arises from the finiteness of the memory system: the moved element has to be

retrieved from memory and integrated at some later point in the sentence in order to derive the proper interpretation. Hence, the structure of the memory system necessarily imposes limits to our ability to keep track of moved elements. A second type of constraint appears to arise from grammar itself. Structures with a short distance between the moved element and its retrieval site impose only a small burden on the memory system, but may nevertheless result ungrammatical.

Two sets of theories were developed to account for constraints on long-distance dependencies, reflecting these two types of constraints. At the level of grammar, the span over which a syntactic rule can apply is regulated by a *locality constraint*. One of the most prominent syntactic theories defining locality constraints is *Featural Relativized Minimality* (henceforth fRM; e.g., Friedmann et al. 2009, Rizzi 1990, 2001, 2004, Starke 2001). According to fRM, a local relation between two elements is disrupted when a third element structurally intervenes between them. Importantly, the disruption is function of the overlap between the syntactic features of the intervener and the distant element: ill-formedness arises when their syntactic features fully match, milder degradation arises when the features partially match, and no degradation arises when there is no match at all.

At the level of memory retrieval mechanisms, constraints on long-distance dependencies are posited by the *Content Addressable* or *Cue-based Memory Model* (henceforth, MM; e.g., McElree 2000, McElree et al. 2003, Lewis et al. 2006, Van Dyke and McElree 2006), according to which the processing of long-distance dependencies involves a limited focus of attention and a content-addressable, cue-based retrieval mechanism sensitive to similarity-based interference.

The two theories were developed along independent lines and typically deal with different types of long-distance dependencies: fRM deals with constraints on grammaticality like weak islands (wh-islands and possibly factive and negative islands), head-movement, A-movement (as superraising), and A'-movement (see Rizzi 1990); MM accounts for the difficulty in the comprehension of a vast range of grammatical structures like object relatives, cleft sentences, and center-embedding structures (e.g., Gordon et al. 2001, 2004, McElree 2000, McElree et al. 2003, Lewis and Vasishth 2005, Lewis et al. 2006). Nevertheless, the two theories bear a striking resemblance in two respects.

First, they grant a key role to the influence of elements perturbing the establishment of the long-distance dependency. But while according to fRM only elements c-commanding the gap are assumed to have the potential to generate *intervention* effects undermining the grammaticality of the sentence, according to MM any element in the sentence has the potential to trigger *interference* in retrieval, independently of its structural or linear position.

Secondly, both theories anchor the phenomenon of intervention/interference[1] in the concept of *similarity*: the long-distance dependency is perturbed as a function of the similarity between the moved element and the intervening/interfering element(s) in the sentence, higher similarity resulting in higher perturbation. Although both accounts hypothesize that similarity modulates the influence from the disturbing element, fRM and MM disagree on the properties that would induce similarity effects. Featural Relativized Minimality suggests that only a subset of syntac-

---

[1] It is worth noticing that whereas the term *intervention* used in fRM refers to a particular configuration of feature overlap in the hierarchical structure, the term *interference* used in MM refers to its consequences at the processing level. We will thus use the term *interference* to refer to the consequence, at the processing level, of structural intervention.

tic features is *seen* in the computation of similarity, namely morphosyntactic features triggering movement (*narrow similarity*). Following MM, on the other hand, any syntactic or semantic feature comes into play in the similarity metric (*broad similarity*).

This paper deals with features inducing similarity effects. Two acceptability judgment experiments in French were undertaken to evaluate the nature of features modulating similarity effects in sentences containing the extraction of a wh-element out of a wh-island. In section 2, we introduce the two theoretical accounts of similarity: the narrow similarity approach of fRM and the predictions stemming from it in section 2.1, and the broad similarity approach and its predictions in section 2.2. In section 3, we report two acceptability studies.

The first tested the role of a syntactic and semantic feature that does not trigger movement, *animacy* (section 3.1), and the second tested the role of a pure semantic feature, *thematic roles reversibility* (section 3.2). Both experiments also manipulated the lexical restriction of the wh-elements, a syntactic feature that triggers movement, previously shown to affect wh-islands acceptability (Atkinson et al. 2015, Villata et al. 2016). Whereas MM predicts that all three variables will affect acceptability judgments, fRM predicts that only lexical restriction will affect them. In section 4, we discuss the implications of our results against the background of narrow and broad approaches to similarity.

2. Theoretical accounts of similarity

2.1 Narrow Similarity: Featural Relativized Minimality

According to Featural Relativized Minimality,[2] a local relation between two non-adjacent elements X and Y is disrupted by an intervening element Z when:

    a.   Z structurally intervenes between X and Y, and
    b.   Z fully matches the featural morphosyntactic specification of X

The first criterion defines the structural position that an element must occupy in order to count as an *intervener*: Z structurally intervenes between X and Y and it c-commands Y. The second criterion specifies that, in order to trigger an *intervention effect* blocking the dependency, the c-commanding intervener Z must also be endowed with the same featural specification as X. Notice that both criteria must be met for an intervention effect to arise, as illustrated in (3) where the intervening wh-subject *who* structurally intervenes between the extracted wh-object *what* and its trace (i.e., *who* c-commands the trace of *what*) and also fully matches its

---

[2] Featural Relativized Minimality is the latest version of Relativized Minimality. In its first formulation, Relativized Minimality (Rizzi 1990) defined similarity in terms of positions of movement (i.e., Head-positions, A-positions, and A'-positions), the core idea being that an element targeting, for instance, an A'-position cannot jump across an intervening A'-specifier. However, this first approach turned out to be too coarse-grained, as not all A'-interveners block A'-chains. This led to a refinement of the theory towards a notion of similarity in terms of featural specification, labelled *Featural Relativized Minimality* (e.g., Friedmann et al. 2009, Rizzi 2004, 2013, Starke 2001).

featural specification (both bearing the [+Q] feature defining questions and attracting movement):

<pre>
      X   Z   Y
</pre>
(3) *What$_i$ do you wonder who wrote __ $_i$?
<pre>
    +Q   +Q
</pre>

Now consider the sentence in (4): despite having the same featural specification as X, Z (i.e., *whom*) does not structurally intervene between X and Y, and the well-formedness of the sentence is thus preserved:

(4) What$_i$ do you believe that the professor of whom wrote __ $_i$?
<pre>
    +Q   +Q
</pre>

Similarly, the well-formedness of a sentence is not undermined when Z c-commands the trace of X without having features in common with it, as in (5): although *the professor* structurally intervenes between *what* and its trace, it does not match the featural specification of *what*, and the well-formedness of the sentence is thus preserved.[3]

(5) What$_i$ do you believe that the professor wrote __ $_i$?
<pre>
    +Q   +N
</pre>

Importantly, the featural approach hypothesizes that different degrees of similarity may give rise to different degrees of degradation, as summarized in (6), where +A and +B stand for the morphosyntactic features attached to X, Z, and Y. Under this approach, the grammar assigns at least three degrees of grammaticality: ill-formedness is predicted for sentences with a full match of features, in which the two featural sets are in a relation of *identity* (6a); moderate degradation is predicted for sentences with a partial match of features, in which the featural set of the intervener is a subset of that of the extractee, therefore in a relation of *inclusion*, (6b); no degradation is predicted for sentences in which the two featural sets are in a relation of *disjunction* (6c).[4]

---

[3] We assume that wh-elements as *what* and *who* bear a [+Q] feature that defines questions and attracts wh-movement. We also assume that definite descriptions as *the professor* bear a [+N] feature, under the reasonable postulation that if [+N] defines the lexically restricted character of complex wh-elements as *which book* (as assumed by Friedmann et al. 2009), it also defines definite descriptions themselves.

[4] Other set-theoretic configurations can be considered like a case of *Inverse Inclusion*, with the specification of the intervener properly including the specification of the extractee (whose role is explored in Villata et al. in press), *Intersection*, in which the extractee and the intervener share some but not all the relevant features (see Belletti et al. 2012), and *Complex Identity*, in which the identity of the featural specification between the extractee and the intervener involves more than one feature, as we will see later in this paper (see also Villata et al. 2016).

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

Experimental evidence in favor of the set-theoretic approach came from Friedmann et al. (2009). In a series of sentence-picture matching experiments, the authors showed that the comprehension of object relatives in Hebrew-speaking children is enhanced as a function of the featural dissimilarity between the extracted object and the intervening subject. In particular, children show above chance performance for object relatives involving a disjunction relation (e.g., free relatives and headed object relatives with an impersonal arbitrary pro subject).

Along the same lines, Adani (2008) and Adani et al. (2010) showed that dissimilarity in the number and gender features of the object and the subject facilitated the comprehension of object relatives by both English and Italian children. These findings can be accounted for by the fact that sentences involving a greater dissimilarity between the intervener and the moved element (i.e., gender and/or number mismatch) are less perturbed than those involving a greater similarity of the two (i.e. gender and/or number match).

In summary, according to fRM a long-distance dependency is perturbed when the featural specification of an intervening c-commanding element matches, either fully or partially, the featural specification of the extracted element. What are the features entering in the calculation of the similarity according to the set-theoretic approach? The theory is very restrictive in this respect: only morphosyntactic features triggering movement (as +Q, +R(elative), +Topic, +Focus, or features belonging to the set of Phi features expressed on the morphology of the verb, as is the case of number and gender in some languages) are relevant in the calculation of the similarity metric (e.g., Belletti et al. 2012, Rizzi 2004, Starke 2001).

This assumption is anchored in a longstanding view in generative grammar according to which movement is always motivated by the need to satisfy some morphosyntactic requirement. In particular, movement is motivated by feature checking (Chomsky 1995 and much subsequent work). It naturally follows that fRM, being a constraint on movement, is restrictively constrained by features triggering movement.

Evidence in favour of this position comes from Belletti et al. (2012) who showed that the comprehension of object relative clauses by Hebrew children is enhanced when the extracted object and the intervening subject mismatch in gender, a feature that triggers movement in that language as it belongs to the Phi features of the verb. Conversely, no improvement was observed for children speaking Italian, in which gender plays no role in movement. We will refer to this approach as the *narrow approach to similarity*.

2.2 Broad Similarity: The Content-Addressable Memory Model

Psycholinguistic research has accumulated evidence showing that the distant argument of a verb is reactivated when the parser reaches the verb (see amongst others, Stowe 1986, Bever & McElree 1988, Osterhout & Swinney 1993). Given the limited focus of attention of one or at

the most two items (McElree 2006), the distant element needs to be retrieved from memory.

According to MM, the retrieval mechanism is *content-addressable* in the sense that the probe (i.e., the verb) triggers a search for its argument on the basis of cues. These cues enable a direct access to the target item, similarly to when an item is retrieved from long-term memory, but in contrast to the word-by-word serial search mechanism operative when order information has to be retrieved (e.g., McElree 2001, McElree & Dosher, 1993, McElree 2000, McElree et al. 2003, Martin & McElree 2008, Van Dyke & McElree 2011). Given its content-addressable nature, the retrieval operation is sensitive to *similarity-based interference*: the more similar to the target a non-target element in the sentence is, the more it interferes with target retrieval (e.g., Gordon et al. 2001, Lewis et al. 2006, McElree 2006). In this framework, any type of information cueing the retrieval process may thus play a role, whether it is syntactic or semantic.

Similarity-based interference manifests itself in terms of slower and/or less accurate sentence comprehension. For example, Van Dyke (2007) showed that, in sentences in which the subject and the verb are separated by a relative clause, a noun phrase inside the relative clause engenders significantly more interference if it occupies a subject position similar to the target subject to retrieve, than if it occupies an object position in a prepositional phrase (see also Van Dyke and Lewis 2003). Interestingly, Van Dyke and McElree (2011) conducted Speed-Accuracy Trade off (SAT) and eye-tracking studies on similar structures and showed that semantic similarity-based interference arises only when the syntactic features of the interfering element match those of the target.

Their result suggests that even though both syntactic and semantic cues play a role in sentence comprehension, syntactic cues are likely to be weighted more strongly in the retrieval process than semantic cues, and possibly play a gating function so that only candidates in memory with matching syntactic properties are considered. Importantly, similarity in the colour of the written target and probe of retrieval does not affect comprehension of a long-distance dependency (Hofmeister and Vasishth 2014), neither does their phonological similarity (Obata et al. 2011). This suggests that surface features are irrelevant to the retrieval process: only those features that are relevant to the syntactic/semantic dependency linking the target and the probe are *seen* by the processor.

Along these lines, MM accounts for the results of Belletti et al. (2012) (see section 2.1) showing that gender mismatch facilitates object relative clause comprehension in Hebrew children but not in Italian children by the fact that gender plays a role in the verbal morphology of Hebrew and therefore is a retrieval cue, while it plays no role in Italian where verbs do not agree in gender with their subjects.

The MM also accounts for the findings of Adani (2008) and Adani et al. (2010) that object relatives involving a number mismatch are easier to process by English and Italian children than those involving a number match by the fact that number is a retrieval cue in these two languages. Since both semantic and syntactic features may enter into the similarity metric, to the condition that they provide cues that the verb may use to retrieve its argument, we will refer to this approach as a *broad approach to similarity*.

So far, we discussed the role of syntactic and semantic features in modulating the ease with which *grammatical* sentences are comprehended. The question remains as to whether these

274

features also have the potential to affect the acceptability of structures that are considered *ungrammatical*. Hofmeister et al. (2013) explored that possibility in sentences containing a superiority violation (e.g., *What who read?*) (for similar results on wh-islands see Atkinson et al. 2015, Villata et al. 2016). They showed that lexically restricted moved wh-objects (e.g., *which book*) are retrieved faster than bare ones (e.g., *what*), even when the intervening subject is also lexically restricted, and concluded that syntactic and semantic richness contributes to reduce interference in ungrammatical sentences by increasing the distinctiveness of the elements in memory.

The present study aims to further explore the role of syntactic and semantic similarity in structures typically considered as ungrammatical: sentences containing an extraction out of a wh-island. Three features were manipulated: lexical restriction, animacy and the reversibility of thematic roles. Lexical restriction carries both syntactic and semantic features, and can be assumed to trigger movement.

Animacy is also both syntactic and semantic in nature. Although there are languages for which one may argue that animacy triggers movement (this could be the case for differential object marking in Spanish, see for instance Torrego 1998, or Rumanian, see Farkas 1978), we will assume that it fails to do so in French where, although it is morphologically realized on the interrogative pronoun (*qu'est-ce que* for inanimate referents and *qui* for animate referents)[5], it is not realized on the verb (see Belletti et al. 2012, Bentea & Durlemann 2014, Bentea et al. 2015, although see Bianchi 2006 for a different standpoint).

Thematic role reversibility is a purely semantic feature. If only narrow similarity matters in the acceptability of wh-islands, as assumed by fRM, only lexical restriction, triggering movement, is expected to modulate judgments. If, in contrast, any type of feature cueing the retrieval of the moved wh-element has the potential to modulate judgments, as assumed under the broad similarity hypothesis of MM, all three variables are expected to influence performance.

3. Experimental investigation of the role of similarity in wh-islands

3.1 Experiment 1

3.1.1 Participants

Forty-two French-speaking students of the University of Geneva took part in the experiment for course credits. They were between 18 and 26 years of age and were all native French speakers.

3.1.2 Materials and Procedure

We manipulated two variables: i) the lexical restriction of the wh-elements (both bare vs.

---

[5] In the *qu'est-ce que* interrogative form, the moved *que* (*what*) is accompanied by the interrogative form *est-ce que* (pronounced /esk/) which is a kind of overt Q marker preventing subject-verb inversion.

both lexically restricted), and ii) the match in animacy between the extracted wh-element and the intervening wh-element (animacy match, where both are animate vs. animacy mismatch, where the extracted wh-element is inanimate while the intervening wh-element is animate). All verbs required animate subjects. The two variables were part of a fully-crossed design involving 4 experimental conditions.

(7)   Qui te demandes-tu qui a apprécié?   (Bare, Animacy match)
    'Who do you wonder who appreciated?'
    +Q,+A   +Q,+A

(8)   Qu'est-ce que tu te demandes qui a apprécié?   (Bare, Animacy mismatch)
    'What do you wonder who appreciated?'
    +Q,-A   +Q,+A

(9)   Quel professeur te demandes-tu quel étudiant a apprécié?   (Restricted, Animacy match)
    'Which professor do you wonder which student appreciated?'
    +Q,+N,+A   +Q,+N,+A

(10)   Quel cours te demandes-tu quel étudiant a apprécié?   (Restricted, Animacy mismatch)
    'Which class do you wonder which student appreciated?'
    +Q,+N,-A   +Q,+N,+A

Thirty-two experimental items were generated. The 128 experimental sentences were split into two between-subjects lists containing 64 items each. The experimental sentences of each list were intermixed with 136 fillers. All fillers were constituted by wh-islands, including the experimental items of Experiment 2. Half of the experimental sentences contained *se demander* (*to wonder*) as a main verb, whereas the other half contained *savoir* (*to know*).

All sentences containing the verb *se demander* were affirmative, while only half of the sentences containing the verb *savoir* were affirmative, with the aim to introduce some variability in the experimental material. Sentences were displayed on a computer screen one at a time. We asked participants to carefully read each sentence and then judge its acceptability on a 7-point Likert scale (1 corresponding to a wholly unacceptable sentence and 7 to a fully acceptable one) by pressing one of the seven numbered buttons on the keyboard. We explicitly informed them that there was no time constraint on responses. Participants were first presented with 10 training items in order to be familiarized to the Likert scale. Three short pauses were administered during the task. The whole session lasted about 20 minutes.

3.1.3 Predictions

According to fRM, the 4 experimental conditions are all cases of identity of morphosyntactic features triggering movement: both the bare and the restricted conditions contain an intervening wh-subject that has the same set of relevant features, namely [+Q] in the bare condition and [+Q,+N] in the restricted condition (see fn. 3). Moreover, animacy does not enter in the definition of the featural sets of the wh-elements as it does not trigger movement. Hence, fRM

predicts all sentences to be highly degraded. In contrast, MM predicts effects of both variables, since both modulate the syntactic and semantic distinctiveness of the elements in memory.

3.1.4 Results

Figure 1 shows the mean acceptability scores for the 4 conditions of Experiment 1. A linear mixed-effects model was fitted to the data using the lmerTest package (http://www.cran.r-project.org/web/packages/lmerTest// lmerTest.pdf) in the R software environment (R Development Core Team, 2011), with random intercepts and slopes for both subjects and items, and lexical restriction and animacy as predictors (fixed factors). Analyses were run on raw scores. A summary of the fixed effects is reported in Table 1.
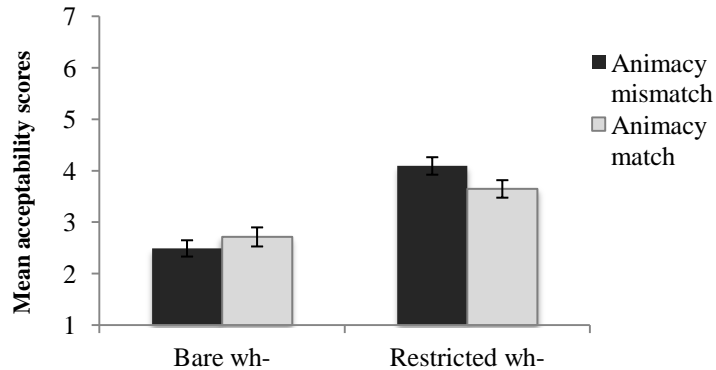

Figure 1. Mean acceptability scores for the 4 conditions of Experiment 1.

Analyses revealed a main effect of Lexical Restriction, attesting to significantly higher rates for sentences with restricted wh-elements ($M = 3.89$) than for sentences with bare wh-elements ($M = 2.60$) ($p < .001$). No main effect of Animacy was observed ($p = 0.236$), but Animacy entered into a significant interaction with Lexical Restriction ($p = 0.003$). Separate mixed effects models were conducted to explore the effect of Animacy for the bare and restricted conditions. An effect of Animacy was found in the restricted condition with higher rates for sentences with animacy mismatch ($M = 4.09$) than for sentences with animacy match ($M = 3.65$) ($\beta = 0.223$, $t = 3.089$, $p = 0.004$). Animacy had no effect in the bare condition ($M = 2.49$ vs. $M = 2.71$) ($\beta = -0.111$, $t = -1.564$, $p = 0.125$).

Variable	Estimate	Std. Error	t	p
Intercept	3.236	0.146	22.205	<.001
Lexical Restriction	0.633	0.083	7.699	<.001
Animacy	0.056	0.046	1.200	=.236
Lexical Restriction*Animacy	0.167	0.055	3.060	=.003

Table 1. Summary of the fixed effects in Experiment 1.

3.1.5 Discussion

Results from Experiment 1 showed that conditions with two restricted wh-elements are rated higher than conditions with two bare wh-elements. This result replicates previous findings on wh-islands (Atkinson et al. 2015, Villata et al. 2016) and superiority violations (Hofmeister et al. 2013). Results also attested to a significant effect of animacy in modulating acceptability judgments when the two wh-elements were lexically restricted, showing that sentences in which these two elements mismatched in animacy are more acceptable than those in which they match.

The results find a natural explanation under the broad similarity hypothesis according to which sentences containing semantically more distinctive elements are easier to parse. For example, two restricted wh-elements as 'which class' and 'which student' in *Which class do you wonder which student appreciated?* are more distinct from one another than two bare wh-elements as 'what' and 'who' in *What do you wonder who appreciated?*. Indeed, students are the kind of entities that can appreciate things, and classes are the kind of entities that can be appreciated, therefore providing a good semantic match to the retrieval process triggered by the verb. Hence, the smaller overlap between the semantic features of restricted wh-elements presumably contributes to reduce similarity-based interference and, as a result, increase sentence acceptability (see Hofmeister et al. 2007, 2013). Along the same lines, an animacy mismatch also increases distinctiveness between wh-elements, thus reducing similarity-based interference and increasing sentence acceptability.

As discussed in detail in Villata et al. (2016), the effect of lexical restriction may also be accounted for by fRM, under additional assumptions (Rizzi 2011). Sentences with two restricted wh-elements may indeed be treated as a configuration of inclusion if one assumes that a restricted wh-element has the option to be attracted both by a complex head [+Q,+N] and by a simple head [+Q], as it meets the attraction conditions of both a complex head and a simple head (while the same does not hold for bare wh-elements that cannot be attracted by a complex head being uniquely specified as [+Q]). Therefore, the restricted condition permits a representation in which the extracted wh-element (e.g., *which class*) is moved to Spec of [+Q,+N], whereas the intervening wh-element (e.g., *which student*) is moved to Spec of [+Q], thus reducing to a configuration of feature inclusion, which is expected to be rated higher than the bare condition that involves a configuration of identity. However, animacy does not belong to the Phi feature set expressed on the verbal inflection, which suggests that animacy does not act as a trigger for movement (Belletti et al. 2012).

Under this perspective, the effect of animacy cannot be accounted for by the narrow similarity hypothesis (but see Bentea et al. 2015 for an opening concerning the possibility that also

278

features morphologically unrealized on the clausal inflectional head can trigger movement, in which case the effect of animacy in the calculation of intervention would be expected).

Why did animacy fail to show an effect in sentences with bare wh-elements? We see three possible explanations. One possibility is that while *qui* (*who*) is endowed with a [+A] animacy feature, *qu'est-ce que* (*what*) is undetermined with respect to animacy, since *what* often, but not always, refers to inanimate entities. For instance, in sentences like *Qu'est-ce qu'il se passe? (What happened?)*, or *Qu'est-ce que tu fais?* (*What are you doing?*)*, what* refers to events that do not seem specified with respect to animacy. A second possibility is that animacy is a property of nouns, not of question operators that are functional particles. In this view, neither *who* nor *what* is endowed with the animacy feature. A third possibility would consist in maintaining that although both *who* and *what* bear the animacy feature ([+A] and [-A] respectively), sentences with two bare wh-elements are too degraded to leave room for the animacy effect to show up.

The novelty of Experiment 1 lies in the finding that the animacy of lexically restricted elements in wh-islands plays a role in their acceptability. However, animacy has both a syntactic and a semantic component; thus, it is unclear whether semantics influences the acceptability of wh-islands since the effect may as well be syntactic. If the animacy effect is a syntactic effect, the results may potentially be accounted for by fRM, under a more relaxed version of the theory, which would assume that syntactic features in the broad sense, including those not triggering movement, have a role to play, although possibly to a lesser degree than syntactic features triggering movement. In Experiment 2, we tested the role of a purely semantic feature with no syntactic correlate: the reversibility of thematic roles, as determined by knowledge of the world.

3.2 Experiment 2

3.2.1 Participants

The same participants recruited for Experiment 1 took part in Experiment 2.

3.2.2. Materials and Procedure

Thirty-two experimental items were created. We manipulated two variables: i) the lexical restriction of the wh-elements (both bare vs. both lexically restricted), and ii) the reversibility of thematic roles (reversible vs. non-reversible). The two variables were part of a fully-crossed design with 4 experimental conditions. All wh-elements were animate.

(11) Qui te demandes-tu qui a licencié?   (Bare, Non-reversible)
      'Who do you wonder who fired?'

(12) Qui te demandes-tu qui a vu?   (Bare, Reversible)
      'Who do you wonder who saw?'

(13) Quel employé te demandes-tu quel chef a licencié?   (Restricted, Non-reversible)

279

'Which employee do you wonder which boss fired?'

(14) Quel employé te demandes-tu quel chef a vu?   (Restricted, Reversible)
       'Which employee do you wonder which boss saw?'

The experimental sentences were split into two between-subjects lists containing 64 items each and intermixed with 136 fillers all constituted by wh-islands, amongst which the experimental items of Experiment 1. The choice of the main verb and its polarity was the same as in Experiment 1. The procedure was the same as in Experiment 1.

3.2.3 Predictions

According to fRM, reversibility should not affect acceptability judgments since this feature does not trigger movement and is not even a syntactic feature. Reversible sentences contained arguments that could be the agent or the patient interchangeably (e.g., the fact that a boss *sees* an employee is equally plausible as an employee seeing a boss), while non-reversible sentences contained arguments whose thematic roles are not interchangeable (e.g., a boss may *fire* an employee but an employee may not fire a boss). The finding of an effect of reversibility would provide strong evidence for a role of semantics in modulating the acceptability of wh-islands.

3.2.4 Results

The same data analyses as in Experiment 1 were conducted. Figure 2 shows the mean acceptability scores for the 4 conditions of Experiment 2, and a summary of the fixed effects is reported in Table 2.
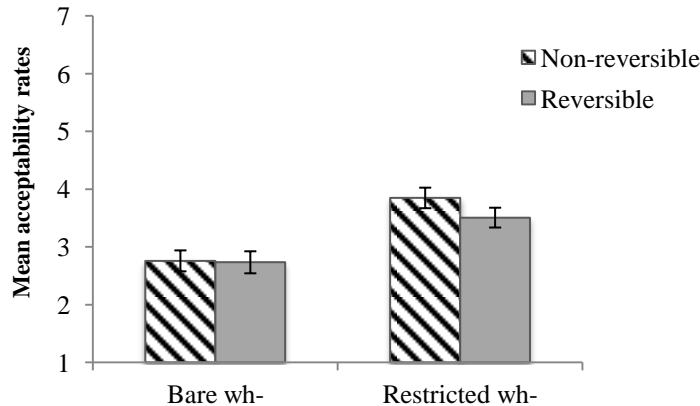

Figure 2. Mean acceptability scores for the 4 conditions of Experiment 1.

Results attested to a significant main effect of Lexical Restriction ($p < 0.001$), with higher scores when both wh-elements are restricted ($M = 3.67$) than when they are bare ($M = 2.74$).

Moreover, a main effect of Reversibility was observed (p = 0.004), attesting to higher scores when the sentence is non-reversible (*M* = 3.3) than when it is reversible (*M* = 3.12). Finally, a significant interaction between Lexical Restriction and Reversibility was also observed (p = 0.038). Separate models were conducted to assess the role of Reversibility for bare and restricted conditions. In the restricted condition, we found higher rates for the non-reversible condition (*M* = 3.85) than for the reversible one (*M* = 3.50) (*β* = 0.170, *t* = 3.047, *p* = 0.004), whereas no effect was observed in the bare condition (*M* = 2.75 vs. *M* = 2.74) (*β* = 0.012, *t* = 0.346, *p* = 0.73).

Variable	Estimate	Std. Error	t	p
Intercept	3.210	0.156	20.468	<.001
Lexical Restriction	0.465	0.077	6.034	<.001
Reversibility	0.091	0.030	3.043	=.004
Lexical Restriction*Reversibility	0.078	0.036	2.164	=.038

Table 2. Summary of the fixed effects in Experiment 2.

3.2.5 Discussion

Results showed a significant effect of reversibility, with higher acceptability rates for non-reversible sentences than for reversible ones. This finding suggests that a pure semantic feature, lying in general knowledge of the world, can modulate acceptability, in line with a broad approach to similarity. This factor has an effect on restricted elements only, suggesting that what really matters is the semantics of the verb-argument relationship determining reversibility and absent in the bare condition, rather than the specificities of the verbs themselves (*fire* vs. *see*). Finally, results from this second experiment also replicate the finding of Experiment 1 that sentences with two restricted wh-elements are rated higher than sentences with two bare wh-elements.

4. General discussion

In two acceptability judgments in French, we tested the possibility that features that do not trigger movement modulate the acceptability of wh-islands, in order to discriminate between two approaches to similarity endorsed by the two major frameworks accounting for long-distance dependencies: fRM and MM. Both approaches grant a key role to feature similarity between the extracted, long-distant wh-element and the intervening wh-element: the more alike they are, the more perturbed is the system. However, the two frameworks differ in the definition of the features relevant to calculate similarity. According to the narrow approach to similarity put forward by fRM, only morphosyntactic features triggering movement enter in this calcula-

tion, whereas the broad approach to similarity defended by MM assumes that similarity is determined also by syntactic features that do not trigger movement and by semantic feature.

Results from both experiments converge in supporting the broad approach. Experiment 1 reported a significant effect of animacy, a morphosyntactic feature that does not trigger movement in French. Experiment 2 reported an effect of the reversibility of the arguments, a strictly semantic feature. Although the role of animacy and reversibility in the comprehension of complex structures is already well-attested (e.g., Caramazza & Zurif 1976 for the role of reversibility in typical and atypical adults, Mak et al. 2002, 2006 and Traxler et al. 2002 for the role of animacy, and Garraffa and Grillo 2008 for both animacy and reversibility in an aphasic patient), this is the first study showing their influence on the comprehension of structures traditionally considered as ungrammatical.

Although both effects find a natural explanation under the broad approach to similarity of MM, they are difficult to account by the narrow approach advocated by fRM. The effect of animacy could potentially be captured by the narrow approach, either by assuming that features that are not morphologically realized on the verb may nevertheless come into play in the similarity metric, or by arguing that animacy triggers movement (along the lines of Bianchi 2006). However, the effect of reversibility, a pure semantic feature, is hardly accountable under a narrow approach to similarity.[6]

Lexical restriction, a feature triggering movement, was also found to play a role in both experiments, replicating previous findings with wh-islands (Atkinson et al. 2015, Villata et al. 2016) and superiority violations (Hofmeister et al. 2013). Here again, the ameliorative effect in sentences with lexically restricted wh-elements finds a natural explanation under MM: lexically restricted elements being endowed with rich semantic features and therefore more distinctive from one another are less susceptible to interference.

Since these sentences appear to involve identical syntactic feature sets, the effect cannot easily be accounted for in terms of fRM. In Villata et al. (2016), the authors argued that the theory may nevertheless find a way to account for this effect by reducing the restricted condition to a configuration of inclusion, which is thus correctly predicted to be rated higher than the bare condition (see Discussion of Experiment 1). However, in three acceptability judgment experiments involving various types of wh-island extractions, the authors showed that the restricted condition is not only rated higher than the bare condition, it is also higher than the inclusion condition.

Hence, without denying the possibility that syntactic features of lexically restricted elements play a role in defining the similarity metric, it seems reasonable to assume that their semantic properties are also involved. In line with that possibility, we note that lexical restriction appears to play a stronger role in modulating wh-islands acceptability than animacy and reversibility, which both showed rather weak effects. The stronger effect of lexical restriction may be explained by the fact that both their syntactic endowment and their multiple semantic features influenced the comprehension of wh-islands.

---

[6] As an anonymous reviewer pointed out, we cannot exclude that participants mostly relied on extragrammatical knowledge to perform the task, rather than parsing the sentence. In this perspective, our results would provide little information about the theoretical question. However, this issue is common to all experimental studies on language comprehension.

It is important to note that judgements remained particularly low (below 4 on the 7-points scale) even in conditions where the moved wh-element and the intervener were maximally distinct, syntactically and semantically. The moderate effect of these featural variations contrasts with the major effect of obtained when the wh-intervener is replaced by a noun phrase, distinct from the extracted wh-element with respect to their syntactic class (as in example (5)), which gives rise to fully grammatical sentences (Villata et al. 2015). Further research is needed to develop an exhaustive map of the various properties (features and syntactic classes) modulating wh-island acceptability, joining the fine characterization of syntactic representations provided by syntactic theory and cognitive models of sentence processing.

References

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

Adani, Flavia, Heather K. J. van der Lely, Matteo Forgiarini and Maria Teresa Guasti. 2010. "Grammatical feature dissimilarities make relative clauses easier: a comprehension study with Italian children". *Lingua* 120(9): 2148-2166.

Alboiu, Gabriela. 2002. *The Features of Movement in Romanian*. Bucharest University Press: Bucharest.

Ambar, Manuela. 1988. *Para uma sintaxe da inversão sujeito-verbo em portuguê*s. Lisbon: Estudos Lingüìsticos.

Atkinson, Emily, Apple Aaron, Rawlins Kyle and Omaki Akira. 2015. "Similarity of wh-phrases and acceptability variation in wh-islands". *Frontiers in Psychology* 6:2048. doi: 10.3389/fpsyg.2015.02048.

Belletti, Adriana, Naama Friedmann, Brunato Dominique and Luigi Rizzi. 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.

Bentea, Anamaria, Stephanie Durrleman and Luigi Rizzi. 2015. "Refining intervention: The acquisition of featural reations in object A-bar dependencies". *Lingua* 169: 21-41.

Bever, Thomas G. and Brian McElree. 1988. "Empty categories access their antecedents during comprehension". *Linguistic Inquiry* 19: 35–43.

Bianchi, Valentina. 2006. "On the syntax of personal arguments". *Lingua* 116: 2023–2067. doi.org/10.1016/j.lingua.2005.05.002.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht : Foris.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York : Praeger.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge. MA: MIT Press.

Farkas, Donka F. 1978. "Direct and indirect object reduplication in Romanian". *Proceedings of the Chicago Linguistic Society* 14: 88 –97

Friedmann, Naama, Adriana Belletti and Luigi Rizzi. 2009. "Relativized relatives: Types of intervention in the acquisition of A-bar dependencies". *Lingua 119*(1): 67-88.

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

Gordon, Peter C., Randall Hendrick and Marcus Johnson. 2004. "Effects of noun phrase type on sentence complexity". *Journal of Memory and Language* 51(1): 97–114.

Hofmeister, Philip, T. Florian Jaeger, Ivan A. Sag, Inbal Arnon and Neal Snider. 2007. "Locality and accessibility in wh-questions". In *Roots: Linguistics in Search of its Evidential Base*, ed. by Sam Featherston and Wolfgang Sternefeld, 185-206. Berlin: Mouton de Gruyter.

Hofmeister, Philip, Florian Jaeger, Ivan A. Sag, Inbal Arnon and Neal Snider. 2013. "The source ambiguity problem: distinguishing effects of grammar and processing on acceptability judgments". *Language and Cognitive Processes* 28(1): 48-87.

Hofmeister, Philip and Shravan Vasishth. 2014. "Distinctiveness and encoding effects in online sentence comprehension". *Frontiers in Language Sciences* 5: 1237.

Lewis, R. Lewis and Shravan Vasishth. 2005. "An activation-based model of sentence processing as skilled memory retrieval". *Cogn. Sci.* (29): 375– 419.

Lewis, R. Lewis, Shravan Vasishth and Julie Van Dyke. 2006. "Computational principles of working memory in sentence comprehension". *Trends in Cognitive Science* 10(10): 447–454.

Martin, Andrea E. and Brian McElree. 2008. "A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis". *Journal of Memory and Language* 58: 879-906.

McElree, Brian. 2000. "Sentence comprehension is mediated by content-addressable memory structures". *Journal of Psycholinguistic Research* 29: 111-123.

McElree, Brian. 2001. "Working memory and focal attention". *Journal of Experimental Psychology: Learning, Memory & Cognition* 27: 817-835.

McElree, Brian. 2006. "Accessing recent events". In *The psychology of learning and motivation*, vol. 46, ed. by B. H. Ross, 155-200. San Diego: Academic Press.

McElree, Brian and Barbara A. Dosher. 1993. "Serial retrieval processes in the recovery of order information". *Journal of Experimental Psychology: General* 122: 291-315.

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

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

Obata, Miki, Richard Lewis, Samuel Epstein, Brian Bartek and Julie Boland. 2010. "Featural analysis and short-term memory retrieval in on-line parsing: Evidence for syntactic, but not phonological, similarity-based interference". In Proceedings of NELS 41, vol.2, ed. by Yelena Fainleib, Nicholas La Cara and Yangsook Park, 69-82. Amhrest, MA: GLSA

Osterhout, Lee and David A. Swinney. 1993. "On the temporal course of gap-filling during comprehension of verbal passives". *Journal of Psycholinguistic Research*, 22: 273–286.

R Development Core Team. 2011. R: A Language and Environment for Statistical Computing. Vienna, Austria : the R Foundation for Statistical Computing. Available online at http://www.R-project.org/

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

Rizzi, Luigi. 2001. "On the position "Int(errogative)" in the left periphery of the clause". In *Current studies in Italian syntax, Essays offered to Lorenzo Renzi*, ed. by Guglielmo Cinque and Giampaolo Salvi, 287- 296. Amsterdam: Elsevier.

Rizzi, Luigi. 2004. "Locality and the left periphery". In *Structures and Beyond: The cartography of syntactic structures, vol. 3*, ed. by Adrian Belletti, 223-251. Oxford: Oxford University Press.

Rizzi, Luigi. 2011. "Minimality", in *The Oxford Handbook of Linguistic Minimalism*, ed. by Cedric Boeckx, 220-238. Oxford-New York: Oxford University Press.

Rizzi, Luigi. 2013. "Locality". *Lingua* 130:169-186.

Soare, Gabriela. 2009. "The Syntax-Information Structure Interface: a Comparative View from Romanian". PhD disseration, University of Geneva.

Starke, Michal. 2001. "Move Dissolves into Merge: A Theory of Locality". PhD dissertation, University of Geneva.

Stowe, Laurie. A. 1986. "Parsing WH-constructions: Evidence for on-line gap location". *Language and Cognitive Processes* 1:227-245.

Torrego, Esther. 1998. *The dependencies of objects.* Linguistic Inquiry Monographs 34. Cambridge, Mass.: MIT Press.

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

Van Dyke, Julie and Richard Lewis. 2003. "Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities". J*ournal of Memory and Language* 49: 285–316.

Van Dyke, Julie. 2007. "Interference effects from grammatically unavailable constituents during sentence processing". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33: 407–430.

Van Dyke, Julie and Brian McElree. 2011. "Cue-dependent interference in comprehension". *Journal of Memory and Language* 65: 247-263.

Villata, Sandra, Luigi Rizzi and Julie Franck. 2015. "Intervention effects in wh-islands and non-islands: experimental evidence". Poster presentation at IGG41. Perugia (Italy), February 26-28.

Villata, Sandra, Luigi Rizzi and Julie Franck. (2016). "Intervention effects and Relativized Minimality: New experimental evidence from graded judgments". *Lingua*.