

A Construction-based Analysis of Verbless Relative Adjuncts in French and Romanian

Gabriela Bîlbîie and Frédéric Laurens

LLF, Université Paris Diderot Paris 7

Proceedings of the HPSG09 Conference

Georg-August Universität Göttingen, Germany

Stefan Müller (Editor)

2009

CSLI Publications

<http://csli-publications.stanford.edu/>

Abstract

French and Romanian verbless relative adjuncts are incidental adjuncts which have been described as elliptical relative clauses. We show that this analysis is not empirically adequate and propose an alternative non-elliptical analysis. We analyze verbless relative adjuncts as sentential fragments whose head can be a cluster of phrases. They are marked by a functor phrase which displays selection properties with respect to the head phrase and makes an essential contribution to the semantics of the adjunct. The analysis relies on the interaction of grammatical constraints introduced by various linguistic objects, as well as on a constructional analysis of verbless relative adjuncts distinguishing several subtypes.

1 Introduction

Both French and Romanian have verbless adjuncts whose form is reminiscent of relative clauses. These verbless adjuncts are exemplified in (1) for French and in (2) for Romanian. They are characterized by the presence of a fronted constituent (shown in *italics*) which can either be a prepositional phrase containing a WH form (1a, 2a, 2b) or the form *dont* in French (1b). We label those constructions Verbless Relative Adjuncts (henceforth VRAs).

- (1) a. Trois personnes, [*parmi lesquelles* Jean], sont venues.
three people(FEM), [among which.FEM John], AUX come
‘Three people, among which John, have come.’
- b. Trois personnes sont venues, [*dont* une hier].
three people(FEM) AUX come, [DONT one.FEM yesterday]
‘Three people have come, one of them yesterday.’
- (2) a. Au venit trei persoane, [{ *printre* | *între* } *care* (și) Ion].
AUX come three people, [{among | among} which.ACC (also) John]
‘Three people have come, among which (also) John.’
- b. Au venit trei persoane, [*dintre care* una ieri].
AUX come three people(FEM), [among which.ACC one.FEM yesterday]
‘Three people have come, one of them yesterday.’

French and Romanian VRAs have been described as elliptical relative clauses (Grevisse 1993, Gheorghe 2004 and Gheorghe 2005) on the basis of semantic and syntactic similarities with non-restrictive relative clauses (3).

[†]Part of this work has been presented in Bucharest at the 11th Conference of the English Department. Many thanks to A. Abeillé, D. Arnold, O. Bonami, D. Godard, J-M. Marandin, the audience of the HPSG09 Conference and three anonymous reviewers for helpful discussions or useful suggestions.

- (3) a. Il a écrit de nombreux romans, [*dont* deux ont été publiés le mois dernier].
 ‘*He has written many novels, two of which have been published last month.*’
- b. El a scris multe romane, [*dintre care* două au fost publicate luna trecută].
 ‘*He has written many novels, two of which have been published last month.*’

Part of the data used in this work is based on a corpus study. The French data comes from the French Treebank of Paris 7 (Abeillé et al. 2003). Lacking a similar corpus for Romanian, we collected examples from press texts.

The paper is structured as follows. We first focus on the constituency of VRAs. We show that the elliptical analysis of VRAs is not empirically adequate and propose an alternative non-elliptical analysis. Then, we discuss the distributional and functional properties of VRAs before turning to their semantic properties. The analysis is then couched in a constructional version of HPSG (i.e. SBCG, Sag 2007). The analysis relies on the interaction of grammatical constraints introduced by various linguistic objects, as well as on a constructional analysis of VRAs distinguishing several subtypes.

For reasons of space, properties of VRAs which are the same in both languages are only illustrated for French. Romanian examples are only used when the property is absent or less clear in French.

2 Constituency

French and Romanian VRAs are constituted of two parts. The first part (labeled the *fronted constituent*) is reminiscent of the extracted phrase or of the complementizer of a relative clause. The second part (labeled the *body*) is composed of either a single phrase or of a sequence of phrases which are not related by grammatical functions. We show that an analysis of VRAs as elliptical relative clauses is not empirically supported and propose an alternative analysis.

2.1 The fronted constituent

The fronted constituent of VRAs contains lexical items which are found in relative clauses. Those items include forms of the WH lexeme *lequel* in French and *care* in Romanian, and the form *dont* in French. While French *lequel* and Romanian *care* are found in interrogative clauses too, French *dont* is only found in relative clauses apart from its use in VRAs.

The WH forms always appear in prepositional phrases in VRAs. The prepositions heading the fronted constituent all have an abstract partitive meaning similar to that of the preposition *among* in English. This includes *parmi* in French and *dintre*, *între*, and *printre* in Romanian.

In both languages, more complex expressions are also found, such as *au {premier | second} rang desquels* ‘in the {first | second} row of which’, *au sommet desquels* ‘at the top of which’, *au sein desquels* ‘in the middle of which’, *au centre desquels* ‘in the center of which’ or *au nombre desquels* ‘in the count of which’ in French, and *în rândul căroră* ‘in the row of which’, *în mijlocul căroră* ‘in the middle of which’ or *în fruntea căroră* ‘at the top of which’ in Romanian. Although these expressions almost always compositionally denote spatial relations, they are only found with an abstract partitive meaning in VRAs. Note that some of these expressions additionally suggest a ranking among the subparts.

The WH form appearing in the fronted phrase is coreferential with a noun phrase preceding the VRA in the host clause, which we label the *licenser* (see section 3.1). This is signalled by morphological agreement on the WH form in French: *lesquels* (MASC) vs. *lesquelles* (FEM). Agreement cannot be observed in Romanian because both the accusative form *care* and the genitive plural form *căroră* do not vary in gender.

In French, the fronted phrase can also contain the form *dont*. Apart from its use in VRAs, the form *dont* appears only in relative clauses where it has been analyzed as a complementizer (Godard 1988).

Like the complementizers *que* and *qui* which are also found in relative clauses, the complementizer *dont* only combines with finite clauses (4a). WH relative clauses on the other hand are not always finite. Unlike prepositional WH forms like *duquel* ‘of which’, the complementizer *dont* cannot be embedded within a filler phrase (4b), and does not show morphological agreement with an antecedent. Finally, complementizer *dont* marks relative clauses containing a missing constituent which would be marked by the form *de* (4c) or any proform coreferential with the antecedent of the relative clause, as long as the proform is embedded under a propositional attitude predicate, such as *penser* ‘to think’ or *être vrai* ‘to be true’ (4d).

- (4) a. * Un livre [*dont* parler]
 a book [DONT talk.INF]
 ‘A book which we should talk about’
- b. Un homme, [*le frère* { * *dont* | *duquel* } est venu hier]
 a man, [the brother { DONT | of.which } AUX come yesterday]
 ‘A man, the brother of which has come yesterday’
- c. Un livre [*dont* on parle beaucoup]
 ‘A book which one talks about’
- d. Un livre [*dont* il est vrai qu’il coûte cher]
 ‘A book of which it is true that it is expensive’

VRAs’ *dont* shares some of its properties with the complementizer *dont*. For instance, it cannot be embedded within the fronted constituent as well. However, it

is unclear whether other selection properties of the complementizer are shared by VRAs' *dont*. Since an elliptical analysis of VRAs is not empirically adequate (see section 2.3), we argue that it is not the case. Moreover, while complementizer *dont* is assumed to have no semantic contribution, such a statement is hard to make for VRAs' *dont* since it forces a partitive semantics (5), although the partitive semantics may also be assumed to be constructionally introduced in VRAs.

- (5) a. Au total, dix livres ont été commandés, [(**dont*) tous pour toi].
'In total, ten books have been ordered, all of them for you.'
- b. Au total, dix livres ont été commandés, [(*dont*) deux pour toi].
'In total, ten books have been ordered, two of them for you.'

The left edge of the fronted phrase must coincide with the left edge of the VRA. It cannot be linearized in or after the body (6a) or be preceded by adverbials (6b).

- (6) a. *Plusieurs personnes, [Jean *parmi lesquelles*], sont venues.
'Several people, among which John, have come.'
- b. *Plusieurs personnes, [notamment *parmi lesquelles* Jean], sont venues.
'Several people, among which notably John, have come.'

2.2 The body

The body of VRAs is constituted of either a single phrase or a sequence of phrases. When the body is a single phrase, it can be either a noun phrase in the citation form (i.e. no prepositional marking in French and Romanian and nominative case in Romanian) or a phrase of any category whose form parallels the corresponding phrase in the host clause.

Not every VRA type allows for the two options. In French, WH VRAs do not allow for marked single phrases and *dont* VRAs disprefer it. *Dont* VRAs with a single marked constituent are not as bad as their WH VRA counterparts and can be improved with adverbs, such as *notamment* 'notably' (7).

- (7) Un jeune homme annonce à diverses personnes sa mort prochaine, [{**parmi lesquelles* | % *dont*] notamment à un psychiatre qui décide de l'aider].
'A young man announces his imminent death to several people, {among which | DONT} notably to a psychiatrist who decides to help him.'

In Romanian, *dintre* is incompatible with preposition or case marking (8a)¹, while no such restriction is found with *printre* and *între* as long as the adverb *și* 'also' precedes them (8b).

- (8) a. Ion lucrează cu șapte doctori, [*dintre care* (**cu*) doi ruși].
 John works with seven doctors, [among which with two Russian]
'John works with seven doctors, two of them are Russian.'

¹ Adverbs such as *mai ales* 'especially' can improve the acceptability of (8a).

- b. Ion a oferit flori mai multor fete, [*printre care* și
 John AUX offered flowers ADV many.DAT girls, [among which also
 {Maria | Mariei}].
 {Mary.NOM | Mary.DAT}]
 ‘John has offered flowers to many girls, Mary among them.’

The body of a VRA can also be constituted of a sequence of phrases (i.e. a cluster). Three types need to be distinguished. Clusters of type I mimic the syntax of the host clause. In those clusters, each of the phrases has to be marked like the corresponding phrase in the host clause (9a). Clusters of type II contain a noun phrase in the citation form and a predicative phrase expressing a property of that noun phrase (9b). Clusters of type III are only found in VRAs whose licenser is a past participle used to express a functional property assumed by some individuals within a event. For example, *un blessé* ‘an injured person’ is the patient of an event in which someone gets hurt. In clusters of that type, the event relation of the cluster is contributed by the past participle. As a result, adverbs, rather than adjectives, are used to modify that relation (9c).

- (9) a. Certains ont parlé à mes amis, [*dont* Marie *(à) Marc].
 ‘Some have spoken with my friends, Mary with Mark.’
 b. Je vends dix jeux, [*dont* la plupart encore dans leur boîte].
 ‘I sell ten games, the majority of them still in their original box.’
 c. L’accident a fait douze blessés, [*dont* cinq grièvement].
 ‘The accident left twelve injured, five of them critically.’

Fronted phrases show selection properties regarding the type of the cluster they combine with. For instance, in French VRAs with *parmi*, at least one phrase in the cluster must be a noun phrase. These properties of fronted phrases combined with properties of clusters of type I can result in ungrammaticality (10).

- (10) J’ai parlé à plusieurs personnes hier, [{*dont* | **parmi lesquelles*} à Marie de linguistique].
 ‘I spoke with several people yesterday, of which with Mary about linguistics.’

2.3 VRAs are not elliptical relative clauses

VRAs are usually referred to as elliptical relative clauses. However, an elliptical analysis of VRAs faces two kinds of problems. Under an elliptical approach, VRAs are assumed to be relative clauses which have the additional property of having some of their syntactic or phonological material removed. An elliptical analysis is of interest if, and only if, (I) one can reconstruct a relative clause from any VRA in a regular fashion and (II) the semantic properties of VRAs are the same as that of relative clauses. We argue that none of these conditions are verified.

2.3.1 Arguments against syntactic reconstruction

There are at least three options for the reconstruction of a verbal form in VRAs. For an example like (11), the first option is to reconstruct a verbal form of the same lexeme as the verb of the host clause of the VRA (11b). This is the obvious option when the VRA contains a cluster of type I. The second option is the reconstruction of an existential verb (11c). The third option is the reconstruction of a quotation verb (11d). In some cases, none of these options will work.

- (11) a. Plusieurs ont eu un cadeau, [*dont* Marie un livre].
'Several have had a present, of which Mary a book.'
- b. *Plusieurs ont eu un cadeau, [*dont* Marie a eu un livre].
'Several have had a present, of which Mary has had a book.'
- c. *Plusieurs ont eu un cadeau, [*dont* est Marie un livre].
'Several have had a present, of which is Mary a book.'
- d. *Plusieurs ont eu un cadeau, [*dont* on cite Marie un livre].
'Several have had a present, of which one mentions Mary a book.'

Within a syntactic reconstruction approach, the choice of a verbal form is dependent on lexical constraints, such as subcategorization properties, which are not correlated with semantic properties. For instance, it is possible to reconstruct the verb *figurer* 'to appear' within a *parmi* VRA but not in a *dont* VRA, because *figurer* can subcategorize a PP marked by *parmi* but not a PP marked by *de*. The reverse is true for an expression such as *faire partie de* 'to belong to'.

2.3.2 VRAs do not have the same semantic properties as relative clauses

Non-restrictive relative clauses behave semantically like independent clauses that contain a proform (Arnold 2004). As a result, their semantic contribution is largely independent from that of their host clause. This is not the case for VRAs, as shown by the contrast in (12). While the sequence of utterances in (12a), which contains a relative clause, is coherent, the sequence in (12b) is contradictory because whales are said both to have and not to have apparent ears. This is so because VRAs are sentential fragments (see section 4.3). Thus, only a syntactic reconstruction which reproduces the content of the host clause is compatible with the semantics of VRAs. However, this kind of reconstruction is most often impossible on syntactic grounds.

- (12) a. Non, tu te trompes! Bien que beaucoup de mammifères, [*dont* les baleines sont un exemple], aient des oreilles apparentes, les baleines n'en ont pas.
'No, you're wrong! While many mammals, of which whales are a example, do have apparent ears, whales do not have any.'

- b. Non, tu te trompes! Bien que beaucoup de mammifères, [*dont* les baleines], aient des oreilles apparentes, les baleines n'en ont pas.
'No, you're wrong! While many mammals, whales among others, do have apparent ears, whales do not have any.'

Another problem faced by the elliptical account is that it predicts that some VRAS should be well-formed, while they are ill-formed for semantic reasons. This is so because the elliptical account assumes that the partitive semantics of VRAS comes from the elided verbal predicate rather than from the fronted phrase. In Romanian, the preposition *dintre* cannot cooccur with a body containing a referential noun phrase such as a definite one. However, a verb form can be reconstructed without difficulty yielding a well-formed relative clause (13).

- (13) Au venit mai multe persoane, [*dintre care* {**Maria* | *o amintim pe Maria*}].
'Many people have come, among which {Mary | one mentions Mary}.'

2.4 Non-elliptical alternative

Non-elliptical analyses differ from elliptical ones in that they do not link form constraints on clusters (such as those exhibited by clusters of type I) to the presence of a syntactic head in the structure. As a result, they make no prediction on the distribution of clusters. We assume that the body of VRAS has exactly the structure it seems to have at first sight: it has a flat structure and has no syntactic head. As for VRAS as a whole, we argue in favor of an analysis in which the body is the head and the fronted phrase is a functor phrase.

The selection properties of VRAS are best attributed to the body. This is so because most of the phrases functioning as the body in VRAS can also function as incidental adjuncts alone with a similar semantics. This is especially the case of those which contain adverbials such as *notamment* 'notably' (14a) or are coordinated structures (14b).

- (14) a. De nombreuses espèces, [(*dont*) notamment les oursins], ont souffert de la pollution.
'Many species, (among which) notably urchins, have suffered from pollution.'
- b. Plusieurs personnes, [une hier et deux ce matin], se sont plaintes de l'organisation.
'Several people, one yesterday and two this morning, complained about the organization.'

The selection properties of VRAS are distinct from those of the fronted phrase. Apart from the French form *dont* whose category is unclear, the fronted constituent is always a prepositional phrase. A preposition like *parmi* 'among' in French introduces a semantic relation between two arguments, one of which is typically

realized as a complement (e.g. *lesquelles* in (15)). The other argument (the external argument) is usually not realized within the preposition phrase itself. Rather it is selected for by the prepositional phrase. In VRAs, the external argument of the fronted phrase is realized within the body (e.g. *Jean* in (15)). Thus, it is reasonable to assume that the selection properties of the fronted phrase and those of the VRA are distinct and therefore not to analyze the fronted phrase as the head.

- (15) plusieurs personnes, [[*parmi lesquelles*] Jean]
 several people.FEM.PL [among which.FEM.PL John(MASC.SG)]
 ‘several people, among which John’

Following Van Eynde 2003, we use the function functor, which replaces the functions specifier, marker and (prehead) adjunct. Functors select a head and are able to mark it, hence modifying its distribution. The fronted phrase is best analyzed as a functor because (I) it must appear before the body (unlike other adjuncts, such as *notamment* ‘notably’), (II) it displays selection properties, and (III) (at least in certain cases) it modifies the distribution of the phrase it combines with.

3 Function in the host phrase

VRAs are incidental adjuncts. They are linearized within a host phrase which must contain a noun phrase introducing a sum-denoting entity. That noun phrase is labeled the *licenser*. The syntactic relation between the VRA and its licenser is submitted to both linear order and locality constraints.

3.1 Properties of the licenser

The licenser must denote a plural entity whose subparts are accessible (i.e. an entity which can be described as a sum of entities) (Lasersohn 1995). This does not entail that the licenser of a VRA will always have plural morphology (16). In most corpus examples however, the licenser is a plural indefinite.

- (16) **Un important volume de gaz** s’est échappé du cratère, [*dont* environ 25% de dioxyde de soufre].
 ‘A great volume of gaz has been released from the crater, including about 25% of sulfur dioxide.’

The licenser can be a dependent of a head of any category (17a) and assume any function including adjunct (17b).

- (17) a. Des représentants de **plusieurs pays**, [*dont* le Brésil], y assistent.
 ‘Representatives of several countries, among which Brasil, attend it.’
 b. Je l’ai attendu **plusieurs heures**, [*dont* une sous la pluie].
 ‘I have been waiting for him for several hours, one of which in the rain.’

In some cases, it might be tempting to describe a VRA as having two licensors or more (18). This analysis is only possible when the fronted phrase does not make it explicit which of the noun phrase is truly the licensor of the VRA. We will come back to the analysis of these examples later in section (4.3).

- (18) J’ai offert **des cadeaux à plusieurs personnes**, [*dont un livre à Marie*].
‘I have offered presents to many people, among which a book to Mary.’

3.2 Ordering and locality constraints

The constraints on the linearization of VRAs need to make reference to the relative linearization of their licensor, as well as to the syntactic structure of their host.

VRAs must follow their licensor (19a). If the licensor is a dependent of a clause’s head, the VRA can be linearized anywhere in the clause after the licensor (19b). However, if the licensor is not a dependent of a clause’s head, the VRA needs to follow its licensor directly (19c).

- (19) a. ([**Dont Marie*]), plusieurs personnes sont venues, ([*dont Marie*]).
‘Several people have come, among which Mary.’
 b. J’ai demandé à plusieurs personnes hier, [*dont Marie*], de m’aider un peu.
‘I asked several people yesterday, among which Mary, to help me a bit.’
 c. Des représentants de plusieurs pays, ([*dont le Brésil*]) se sont réunis, ([**dont le Brésil*]).
‘Representatives from several countries, among which Brasil, have met.’

The relation between a VRA and its licensor obeys the right roof constraint which generally applies on rightwards non-local relations such as extraposition or right dislocation.

- (20) *Que deux personnes viennent ne m’a pas étonné, [*dont Marie*].
‘The fact that two people come, among which Mary, has not surprised me.’

Note that if a VRA contains a cluster of type I, the VRA must follow every phrase of the host clause which is paralleled in the cluster.

- (21) Plusieurs, ([**dont Marie un livre*]), m’ont offert des cadeaux, ([*dont Marie un livre*]).
‘Several have offered me presents, of which Mary a book.’

VRAs are also at least compatible with ‘comma intonation’. Thus, they are a kind of incidental adjuncts, incidentality being defined as a syntactic property which correlates with both phonological and linearization properties (Bonami and Godard 2003).

4 Semantic Properties

VRAs can be associated with two distinct semantics. Although VRAs always introduce a subpart of the sum individual denoted by their licenser, the referential properties of the introduced subpart are not always the same. VRAs are non-restrictive modifiers which, unlike most other non-restrictive modifiers, must scope under propositional attitude verbs. This is explained by the fact that VRAs are sentential fragments.

4.1 VRAs have a partitive semantics

VRAs always introduce an entity which must be interpreted as a subpart of the sum individual denoted by their licenser. Sum individuals are not always composed of atomic parts. As a result, a noun phrase containing a mass noun can function as the licenser of a VRA (16). Quantified noun phrases which do not denote sum individuals are not suitable licensers for VRAs (22).

- (22) *Tout étudiant doit venir, [*dont* Marie].
'Every student must come, among which Mary.'

VRAs cannot introduce any other semantic relation between their licenser and the phrase introducing the subpart. Meronymy and possession, for instance, are ruled out (23).

- (23) *Plusieurs personnes sont venues, [*dont* leur chien].
'Several people have come, among them their dog.'

VRAs can have two distinct semantics depending on the referential properties of the introduced subpart. Exemplifying VRAs introduce a subpart which is referential (i.e. It can be identified independently of the fact that it is a subpart of a sum individual) (24a). This is the case of the noun phrases *Marie* and *Jean* in (24a). Partitioning VRAs, on the other hand, introduce a subpart which is not referential but can be defined within the sum individual as having some properties which are not shared by other subparts. The property in question can be a property of the entity which is the subpart or a property of the subevent in which the subpart takes part (24b). One of the most striking differences between exemplifying and partitioning VRAs is that only the latter can introduce a list of subparts which are coextensive with the sum individual denoted by the licenser (24b).

- (24) a. { *Deux | trois } personnes sont venues, [*dont* Marie et Jean].
'{Two | three} people have come, among which Mary and John.'
- b. Trois personnes sont venues, [*dont* une lundi et deux mardi].
'Three people have come, one on Monday and two on Tuesday.'

Within a noun phrase coordination, it is sufficient that one of the noun phrases be non referential in order for the partitioning semantics to be available (25).

- (25) Prends deux objets, [*dont* cette bouteille et {un | *ce} couteau].
'Take two objects, that bottle and {a | that} knife.'

The semantics of the head of the fronted phrase also plays a role in the semantics of the VRA. In French, *parmi* is always exemplifying, while *dont* can be exemplifying or partitioning. In Romanian, *dintre* is always partitioning, while *printre* and *între* can be both exemplifying or partitioning. These lexical properties can be observed in other uses of the prepositions as well (26).

- (26) a. Avem {majoritatea | spionii} *printre* noi.
'We have {the majority | spies} among us.' (partitioning or exemplifying)
b. {majoritatea | *spionii} *dintre* copii
'the {majority | spies} of the children' (partitioning only)

4.2 VRAs have a non-restrictive semantics

Restrictive modifiers have an intersective interpretation and therefore introduce an implicit 'contrast set', which can be accessed by anaphors like *the others* (Arnold 2004). Non-restrictive modifiers, on the other hand, are not intersective modifiers and introduce no such 'contrast set'. VRAs are non-restrictive modifiers of their licenser as well as of their host. The VRA in (27) does not restrict the set of Mary's friends to the one including John. Therefore, there is no possible antecedent for *the other friends of Mary*. Neither does the VRA in (27) restrict the event denoted by the host clause to the event such that *Some friends of Mary came and John came* as opposed to another event such that *Some friends of Mary came but John didn't*.

- (27) Certains amis de Marie, [*parmi lesquels* Jean], sont venus. #Les autres amis de Marie viendront demain.
'Some friends of Mary's have come, among which John. The others friends of Mary will come tomorrow.'

Unlike most non-restrictive modifiers, the content of VRAs is part of the asserted content of the utterance containing them. As a result, VRAs which are linearized within a clause which is itself in the scope of a propositional attitude verb must be interpreted in the scope of that verb too. Thus, (28) entails that *Peter believes that vervain can heal ulcers* but not that it is actually true or even that the speaker believes it.

- (28) Pierre croit que certaines plantes, [*dont* la verveine], soignent les ulcères.
'Peter believes that some plants, vervain among them, can heal ulcers.'

4.3 VRAs are sentential fragments

VRAs are sentential fragments and describe a subevent of the event denoted by their host clause. This is especially clear in partitioning examples where the subpart is not defined by a property of the subpart itself but by a property of the subevent (24b). However, this is also true of exemplifying examples as can be seen by the ungrammaticality of (29) in which a VRA is embedded within a non-event for which no subevent can be defined.

- (29) *Aucun représentant de ces quatre pays, [*dont* le Brésil], n’a participé à la conférence.
‘No representative of those four countries, among which Brasil, has attended the conference’

VRAs can appear in declarative, imperative (25) or interrogative clauses. In interrogative clauses, they share their abstracted parameter with the host clause and thus can be used as a hint. In (30), if one can answer the question about *France*, then one can answer the question about *countries of the OECD* too. The hint is that knowledge about France is sufficient to answer the question.

- (30) En quelle année, plusieurs pays de l’OCDE, [*dont* la France], ont-ils signé ce traité?
‘In which year, several countries of the OECD, among which France, did sign this treaty?’

The fact that VRAs are sentential fragments allows one to account for cases where VRAs seem to have several licensors as cases where VRAs have only one licensor but also introduce a sum-subpart relation because the event they introduce is a subevent of the one introduced by the host. In that kind of implied sub-subpart relation, the corresponding element in the host does not need to denote a sum-individual at all (31).

- (31) Paul a offert un cadeau à plusieurs personnes, [*dont* un livre *(à Marie)].
‘Paul has offered a present to several people, among which a book (to Mary).’

5 HPSG Grammar Fragment

The properties of VRAs are best analyzed as resulting from the interaction of various syntactic and semantic constraints applying on distinct linguistic objects. An HPSG grammar fragment accounting for the properties of VRAs minimally requires (I) a theory of fragments, (II) a theory of clusters, and (III) a theory of locality of selection. Several aspects of the grammar introduced here are constructional in nature. The analysis is couched in a constructional version of HPSG, namely SBCG (Sag 2007).

5.1 A theory of fragments

Fragments are expressions which convey a semantic content which is not given by their form alone. Rather, the semantic content conveyed by a fragment is a function of (I) the type of the fragment, (II) the semantic content of the constituent(s) in the fragment, and (III) contextual information which can be linguistic in nature or not (Fernández et al. 2007).

A sentential fragment such as the short question *when* in (32a) is interpreted as having the same semantic content as the clause *when she will come* in (32b). This semantic content comes in part from the type of the fragment (short questions have the same content type as interrogative clauses, i.e. a propositional abstract), the semantic content of the constituent(s) in the fragment (*when* provides the parameter for the propositional abstract), and contextual information (*Mary will come* functions as the antecedent of the fragment providing the proposition used to build the propositional abstract).

- (32) a. Mary will come but nobody knows [when].
b. Mary will come but nobody knows [when she will come].

More generally, fragments are reminiscent of description anaphora, as opposed to instance anaphora. In instance anaphora, what is shared between an anaphoric expression and its antecedent(s) are indices. In description anaphora, what is shared is some aspects of the description of the antecedent(s) which apply to a new entity with a different index. This is the case with *one* anaphora in English (33) (Arnold and Borsley 2008).

- (33) Here is a small red mugs with flowers and here is a bigger one.

In all of these cases, computing the semantics of the fragment can be achieved by expressing constraints between four semantic representations: two complete ones (the meaning of the antecedent and the meaning of the fragment) and two partial ones (the content which is anaphorically shared between the antecedent and the fragment, and the content which is given by the constituents in the fragment).

We use MRS (Copestake et al. (2005)) to represent incomplete semantic representations as the underspecification of a complete semantic representation. In MRS, meaning is represented as bags of elementary predications. Connection between these elementary predications is achieved through index-sharing and label-sharing. MRS representations are suitable to express partial meaning.

For an example such as (34), we want to obtain the four bags of elementary predications in (35). The bags are related by two meta-constraints (bag unification and bag intersection) shown at the bottom of figure (35).

- (34) Plusieurs personnes sont venues, [*dont Marie hier*].
'Several people have come, of which Mary yesterday.'

$$(35) \left[\begin{array}{l} \text{ANTECEDENT } [A] \\ \text{VARIABLE } [B] \\ \text{FRAGMENT } [C] \\ \text{CONTENT } [D] \end{array} \left[\begin{array}{l} \text{HOOK } [\text{IND } 1] \\ \text{RELS } [A'] \left\langle \left[\begin{array}{l} \text{message-rel} \\ \text{IND } 1 \end{array} \right], \left[\begin{array}{l} \text{come-rel} \\ \text{IND } 1 \\ \text{ARG1 } 2 \end{array} \right], \left[\begin{array}{l} \text{several-rel} \\ \text{IND } 2 \end{array} \right], \left[\begin{array}{l} \text{person-rel} \\ \text{IND } 2 \end{array} \right] \right\rangle \\ \text{HOOK } [\text{IND } 5] \\ \text{RELS } [B'] \left\langle \left[\begin{array}{l} \text{message-rel} \\ \text{IND } 5 \end{array} \right], \left[\begin{array}{l} \text{come-rel} \\ \text{IND } 5 \\ \text{ARG1 ind} \end{array} \right] \right\rangle \\ \text{HOOK } [\text{IND } 3] \\ \text{RELS } [C'] \left\langle \left[\begin{array}{l} \text{message-rel} \\ \text{IND } 3 \end{array} \right], \left[\begin{array}{l} \text{name-rel} \\ \text{IND } 4 \end{array} \right], \left[\begin{array}{l} \text{yesterday-rel} \\ \text{IND } 3 \end{array} \right] \right\rangle \\ \text{HOOK } [\text{IND } 3] \\ \text{RELS } [D'] \left\langle \left[\begin{array}{l} \text{message-rel} \\ \text{IND } 3 \end{array} \right], \left[\begin{array}{l} \text{come-rel} \\ \text{IND } 3 \\ \text{ARG1 } 4 \end{array} \right], \left[\begin{array}{l} \text{name-rel} \\ \text{IND } 4 \end{array} \right], \left[\begin{array}{l} \text{yesterday-rel} \\ \text{IND } 3 \end{array} \right] \right\rangle \end{array} \right] \right]$$

- $[A'] \cap [D'] = [B']$
- $[B'] \cup [C'] = [D']$

Bag unification: Let A, B and C be bags of elementary predications. A is the unification of B and C iff each element R in A, Q in B and S in C appears in either: (I) a pair $\langle R, Q \rangle$ where R and Q have the same description, (II) a pair $\langle R, S \rangle$ where R and S have the same description or (III) a triple $\langle R, Q, S \rangle$ where R is the unification of Q and S. No element appears in more than one tuple.

Bag intersection: Let A, B and C be bags of elementary predications. A is the intersection of B and C iff (I) B is a possible result of the unification of A and B (i.e. using only triples), (II) C is a possible result of the unification of A and C and (III) there is bag A' such that B is a possible result of the unification of A' and B, C is a possible result of the unification of A' and C and A' has more element than A.

The use of condition III in bag intersection is motivated by the fact that VRAS have a greedy interpretation (i.e. they share everything with their antecedent except what is literally introduced by the fragment). Note that in the two meta-constraints, elementary predications can only be unified if they have the exact same type.

The account is integrated within an HPSG grammar using a feature FRAGMENT. The value of the feature FRAGMENT has two features: ANTECEDENT and VARIABLE which are of type *sem-obj* (36).

$$(36) \left[\begin{array}{l} \text{MTR} \\ \text{DTRS} \end{array} \left[\begin{array}{l} \text{SEM} \left[\begin{array}{l} \text{RELS } [1] \\ \text{RELS } [2] \end{array} \right] \\ \text{C-SEM} \left[\begin{array}{l} \text{RELS } [3] \\ \text{RELS } [4] \end{array} \right] \\ \text{FRAGMENT} \left[\begin{array}{l} \text{ANTECEDENT} \left[\text{RELS } [5] \right] \\ \text{VARIABLE} \left[\text{RELS } [6] \right] \end{array} \right] \end{array} \right] \right]$$

- $\boxed{3} \cap \boxed{1} = \boxed{4}$
- $\boxed{4} \cup \text{union}(\boxed{5}, \boxed{6}, \boxed{2}) = \boxed{1}$

Additional constraints on the semantics of fragments come from the syntax-semantics interface. Constraints can be imposed on by clusters or by the VRA-construction. Incomplete meaning for the fragment (i.e. bags of elementary predications that cannot be mapped into a connex graph) should be excluded as well.

5.2 A theory of clusters

Clusters are sequences of phrases which are not related by functions but nevertheless display constituent properties. Clusters are fragments but the reverse is not always true. The fragment *when* in (32a), for instance, is composed of a single phrase.

Clusters do not always have the same distribution as a headed phrase with a similar semantic content. For French, it has been noticed that some items or constructions are compatible with clusters which are sentential fragments but not with clauses. This is the case of the conjunction *ainsi que* ‘as well as’, which may combine with a cluster but not with a finite clause (Abeillé and Godard 1996, Mouret 2006).

- (37) Paul offrira un livre à Marie, ainsi qu’(*il offrira) un CD à Paul.
‘Paul will offer a book to Mary, as well as (he will offer) a CD to Paul.’

Some cluster types are submitted to form constraints which instantiate lexical knowledge about subcategorization properties of lexical items which are not realized within the cluster.

That property leads Ginzburg and Sag 2000 to analyze sentential fragments as single daughters of a phrase with full clausal properties including the syntactic category VERB. This is problematic for VRAs because sentential fragments and clauses do not have the same distribution. We use the definition given by Mouret 2006 (38). The feature CLUSTER allows one to select or subcategorize for a sign which is a cluster.

$$(38) \text{ cluster-cxt} \Rightarrow \text{phrasal-cxt} \ \& \ \left[\begin{array}{c} \text{MTR} \left[\begin{array}{c} \text{SYN} \left[\begin{array}{c} \text{CAT} \left[\begin{array}{c} \text{cluster} \\ \text{CLUSTER } \boxed{1} \end{array} \right] \end{array} \right] \end{array} \right] \\ \text{DTRS } \boxed{1} \text{ list}(\text{sign}) \end{array} \right]$$

Several constraints can be expressed on clusters, such as the fact that they must contain a noun phrase in the citation form or that their form instantiates knowledge about the subcategorization of a word which is not present in the structure (39).

(39) NP-cluster-cxt \Rightarrow cluster-cxt &

$$\left[\begin{array}{c} \text{MTR} \\ \left[\begin{array}{c} \text{SYN} \\ \left[\begin{array}{c} \text{CAT} \\ \left[\begin{array}{c} \text{CLUSTER} \end{array} \right] \text{ contains} \left(\left[\begin{array}{c} \text{SYN} \\ \left[\begin{array}{c} \text{CAT} \\ \left[\begin{array}{c} \text{noun} \\ \text{CASE} \quad \text{nominative} \end{array} \right] \\ \text{MRKG} \quad \text{det} \end{array} \right] \end{array} \right] \end{array} \right) \end{array} \right] \end{array} \right] \end{array} \right]$$

We will assume that VRAs with only one phrase in the body have a unary cluster body because it allows one to express generalizations in a simplified way.

5.3 A theory of locality of selection

The selection properties of VRAs and the selection properties of their fronted phrase are submitted to the same kind of locality constraints. When a VRA modifies a clause, the licenser of the VRA must be a direct dependent of the clause's head. When a fronted phrase modifies a cluster, the phrase introducing a subpart of the plural entity denoted by the antecedent must be a direct daughter of the cluster.

In order to state these locality constraints on selection, we introduce a set-valued feature *ANCHORS* which contains indices of semantic entities which are accessible to adjunct selection. This feature was originally proposed by Kiss 2005 to analyze extraposed relative clauses. The two following constraints on the propagation of anchors are introduced. Constraint (40) ensures that only direct dependents of a clause's head are accessible via the anchor set on the level of the clause. In (40) and (41), *c* stands for the predicate *contains* which is a relation expressing that a list contains some element. In both constraints, the element contained in the antecedent of the constraint is universally quantified.

$$(40) \left[\begin{array}{c} \text{word} \\ \text{DEPS } c \left(\left[\begin{array}{c} \text{SEM} \\ \text{HOOK} \left[\begin{array}{c} \text{anchor} \\ \text{LABEL} \quad \boxed{1} \\ \text{IND} \quad \boxed{2} \end{array} \right] \end{array} \right] \right) \end{array} \right] \Rightarrow \left[\begin{array}{c} \text{MTR} \\ \text{SEM} \left[\begin{array}{c} \text{ANCHORS } c \left(\left[\begin{array}{c} \text{anchor} \\ \text{LABEL} \quad \boxed{1} \\ \text{IND} \quad \boxed{2} \end{array} \right] \right) \end{array} \right] \end{array} \right]$$

Constraint (41) ensures that only direct daughters of a cluster are accessible via the anchor set on the level of the clause. Another virtue of the anchor constraints is that they restrict semantic selection to material which is literally introduced in the cluster. Thus, reconstructed semantic relations are not available for semantic selection.

$$(41) \left[\begin{array}{c} \text{cluster-cxt} \\ \text{DTRS } c \left(\left[\begin{array}{c} \text{SEM} \\ \text{HOOK} \left[\begin{array}{c} \text{anchor} \\ \text{LABEL} \quad \boxed{1} \\ \text{IND} \quad \boxed{2} \end{array} \right] \end{array} \right] \right) \end{array} \right] \Rightarrow \left[\begin{array}{c} \text{MTR} \\ \text{SEM} \left[\begin{array}{c} \text{ANCHORS } c \left(\left[\begin{array}{c} \text{anchor} \\ \text{LABEL} \quad \boxed{1} \\ \text{IND} \quad \boxed{2} \end{array} \right] \right) \end{array} \right] \end{array} \right]$$

5.4 A theory of VRAs

We analyze VRA constructs as subtypes of the head-functor construct as defined in Van Eynde 2003 and Sag 2007. Functors have two main properties. They select the head sign and contribute a mark to the construct (42).

$$(42) \text{ hd-func-cxt} \Rightarrow \text{hd-cxt} \& \left[\begin{array}{l} \text{MTR} \left[\begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{VAL} \quad [1] \\ \text{MRKG} \quad [2] \end{array} \right] \\ \text{DTRS} \left\langle \begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{CAT} \quad [\text{SELECT} \quad [3]] \\ \text{MRKG} \quad [2] \end{array} \right] \\ \text{HD-DTR} \quad [3] \left[\begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{VAL} \quad [1] \end{array} \right] \end{array} \right] \end{array} \right\rangle \end{array} \right]$$

The VRA construct itself is defined in (43). The head of the construct is a sentential fragment. It is selected for by the fronted phrase which contributes a sum-subpart relation which is characteristic for the construction. The sum-subpart relation is assumed to have a partitioning subtype and an exemplifying subtype. The construction itself contributes a second subpart relation which links the event denoted by the host clause to that introduced by the sentential fragment. The construction also selects a nominal licenser. Notice the use of the set ANCHORS to express the locality of selection of the fronted phrase and of the VRA construct.

$$(43) \text{ VRA-cxt} \Rightarrow \text{phrasal-cxt} \& \text{hd-fun-cxt} \& \left[\begin{array}{l} \text{MTR} \left[\begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{CAT} \quad [1] \\ \text{VAL} \quad [3] \end{array} \right] \left[\begin{array}{l} \text{SELECT} \left[\begin{array}{l} \text{SEM} \left[\text{ANCHORS} \quad \text{contains}([\text{IND} \quad [2]]) \end{array} \right] \end{array} \right] \end{array} \right] \\ \text{SEM} \left[\begin{array}{l} \text{HOOK} \quad [\text{IND} \quad [7]] \end{array} \right] \\ \text{C-SEM} \left[\begin{array}{l} \text{RELS} \left\langle \begin{array}{l} \text{sum-subpart-rel} \\ \text{SUBPART} \quad [7 \text{ event}] \\ \text{SUM} \quad [6 \text{ event}] \end{array} \right\rangle \end{array} \right] \end{array} \right] \\ \text{HD-DTR} \quad [4] \left[\begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{CAT} \quad [1] \\ \text{VAL} \quad [3] \\ \text{MRKG} \quad \text{none} \end{array} \right] \\ \text{SEM} \left[\begin{array}{l} \text{HOOK} \quad [\text{IND} \quad [7]] \\ \text{ANCHORS} \quad \text{contains}([\text{IND} \quad [8]]) \end{array} \right] \\ \text{FRAGMENT} \left[\begin{array}{l} \text{ANTECEDENT} \quad [A] \left[\text{HOOK} \quad [\text{IND} \quad [6]] \right] \\ \text{VARIABLE} \quad [C] \end{array} \right] \\ \text{GAP} \quad \{ \} \end{array} \right] \\ \text{DTRS} \left\langle \begin{array}{l} \text{SEM} \left[\begin{array}{l} \text{HOOK} \quad [\text{LTOP} \quad [5]] \\ \text{RELS} \left\langle \begin{array}{l} \text{sum-subpart-rel} \\ \text{LABEL} \quad [5] \\ \text{SUBPART} \quad [8] \\ \text{SUM} \quad [2] \end{array} \right\rangle \end{array} \right] \end{array} \right\rangle \end{array} \right]$$

WH-VRA-construct is a subtype of VRA-construct. The fronted phrase of a VRA-construct contains a WH form which is coreferential with the nominal antecedent of the VRA-construct (44).

$$(44) \text{ WH-VRA-ctx} \Rightarrow \text{VRA-ctx} \& \left[\begin{array}{l} \text{MTR} \left[\begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{SELECT} \left[\begin{array}{l} \text{SEM} \left[\begin{array}{l} \text{ANCHORS} \text{ contains}([\text{IND } \boxed{1}]) \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \\ \text{DTRS} \left\langle \left[\begin{array}{l} \text{WH} \left\{ \left[\begin{array}{l} \text{HOOK} [\text{IND } \boxed{1}] \end{array} \right] \right\}, \text{sign} \end{array} \right] \right\rangle \end{array} \right]$$

The prepositions functioning as the head of the fronted phrase of a WH-VRA-construct have the following lexical properties (45). They have an argument structure containing two elements, one of which is an internal argument realized as the complement of the preposition. The other argument is the external argument of the preposition and is not realized as a dependent of the preposition. Rather the preposition selects for a phrase which contains an anchor coindexed with its external argument. Finally the preposition has to introduce a sum-subpart relation between its two arguments: the internal argument denoting a sum and the external a subpart of that sum. Prepositions may differ regarding the exact type of sum-subpart relation. Some introduce an exemplifying relation (Fr. *parmi*), some a partitioning one (Ro. *dintre*), some an underspecified one (Ro. *printre* and *între*).

$$(45) \text{ PARMI-word} \Rightarrow \text{word} \& \left[\begin{array}{l} \text{ARG-ST} \left\langle \boxed{1} \left[\begin{array}{l} \text{syn} \left[\begin{array}{l} \text{cat} \text{ noun} \\ \text{VAL} \langle \rangle \\ \text{MRKG} \text{ det} \end{array} \right] \\ \text{SEM} \left[\begin{array}{l} \text{HOOK} [\text{IND } \boxed{3}] \end{array} \right] \end{array} \right] , \boxed{2} \left[\begin{array}{l} \text{SYN} \left[\begin{array}{l} \text{CAT} \text{ noun} \\ \text{VAL} \langle \rangle \\ \text{MRKG} \text{ det} \end{array} \right] \\ \text{SEM} \left[\begin{array}{l} \text{HOOK} [\text{IND } \boxed{4}] \end{array} \right] \end{array} \right] \right\rangle \\ \text{SYN} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{preposition} \\ \text{XARG} \boxed{1} \\ \text{SELECT} \left[\begin{array}{l} \text{SEM} \left[\begin{array}{l} \text{ANCHORS} \text{ contains}([\text{IND } \boxed{3}]) \end{array} \right] \end{array} \right] \end{array} \right] \\ \text{VAL} \langle \boxed{2} \rangle \end{array} \right] \\ \text{SEM} \left[\begin{array}{l} \text{HOOK} [\text{LTOP } \boxed{5}] \\ \text{RELS} \left[\begin{array}{l} \text{exemplifying-sum-subpart-rel} \\ \text{LABEL} \boxed{5} \\ \text{SUBPART} \boxed{3} \\ \text{SUM} \boxed{4} \end{array} \right] \end{array} \right] \end{array} \right]$$

We assume that *dont* is a marker which has no argument structure but introduces an underspecified sum-subpart relation, although it could also be introduced constructionally. *Dont* selects for a phrase which (I) contains an anchor for the subpart argument of its sum-subpart relation and (II) selects for a phrase containing an anchor for the sum argument of its sum-subpart relation.

$$(46) \text{ DONT-word} \Rightarrow \text{word \&}$$

$$\left[\left[\left[\begin{array}{c} \text{SYN} \\ \text{CAT} \\ \text{SELECT} \end{array} \right] \left[\begin{array}{c} \text{SYN} \mid \text{CAT} \mid \text{SELECT} \mid \text{SEM} \mid \text{ANCHORS contains} \left\{ \left[\text{IND} \quad \boxed{4} \right] \right\} \\ \text{SEM} \mid \text{ANCHORS contains} \left\{ \left[\text{IND} \quad \boxed{3} \right] \right\} \end{array} \right] \right] \left[\begin{array}{c} \text{VAL} \quad \langle \rangle \\ \text{MRKG} \quad \text{dont} \end{array} \right] \left[\begin{array}{c} \text{HOOK} \quad \left[\text{LTOP} \quad \boxed{5} \right] \\ \text{RELS} \quad \left\{ \left[\begin{array}{c} \text{sum-subpart-rel} \\ \text{LABEL} \quad \boxed{5} \\ \text{SUBPART} \quad \boxed{3} \\ \text{SUM} \quad \boxed{4} \end{array} \right] \right\} \end{array} \right] \right] \right]$$

(47) DONT-VRA-cxt \Rightarrow VRA-cxt & $\left[\text{MTR} \left[\text{SYN} \left[\text{MRKG} \quad \text{dont} \right] \right] \right]$

6 Conclusion

References

- Abeillé, A., Clément, L. and Toussanel, F. 2003. Building a treebank for French. In A. Abeillé (ed.), *Treebanks*, Dordrecht: Kluwer.

- Abeillé, A. and Godard, D. 1996. La complémentation des auxiliaires en français. *Langages* 122, 32–61.
- Arnold, D. 2004. Non-Restrictive Relative Clauses in Construction Based HPSG. In St. Müller (ed.), *Proceedings of the HPSG04 Conference*, pages 27–47, Stanford: CSLI Publications.
- Arnold, D. and Borsley, R. D. 2008. Non restrictive relative clauses, ellipsis and anaphora. In St. Müller (ed.), *Proceedings of the HPSG08 Conference*, pages 5–25, Stanford: CSLI Publications.
- Bonami, O. and Godard, D. 2003. Incidental adjuncts: an overlooked type of adjunction. Presentation given at *the 10th International Conference on HPSG*. East Lansing: Michigan State University.
- Copestake, A., Flickinger, D., Pollard, C. and Sag, I. A. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation* 3(4), 281–332.
- Fernández, R., Ginzburg, J. and Lappin, Sh. 2007. Classifying Non-Sentential Utterances in Dialogue: A Machine Learning Approach. *Computational Linguistics* 33(3), 397–427.
- Gheorghe, M. 2004. *Propoziția relativă*. Bucharest: Editura Paralela 45.
- Gheorghe, M. 2005. Construcții cu propoziții relative. In *Gramatica limbii române*, volume II, Bucharest: Editura Academiei Române.
- Ginzburg, J. and Sag, I. A. 2000. *Interrogative Investigations: the form, meaning and use of English interrogatives*. Stanford: CSLI Publications.
- Godard, D. 1988. *Syntaxe des relatives en français*. Paris: CNRS Editions.
- Grevisse, M. 1993. *Le Bon Usage*. Boeck Duculot, 13. edition.
- Kiss, T. 2005. Semantic constraints on relative clause extraposition. *Natural Language and Linguistic Theory* 23(2), 281–334.
- Lasnik, P. 1995. *Plurality, Conjunction and Events*. Dordrecht: Kluwer.
- Mouret, F. 2006. A Phrase Structure Approach to Argument Cluster Coordination. In St. Müller (ed.), *Proceedings of the HPSG06 Conference*, pages 247–267, Stanford: CSLI Publications.
- Sag, I. A. 2007. Sign-Based Construction Grammar: An informal synopsis. Manuscript. Stanford University.
- Van Eynde, F. 2003. On the Notion ‘Determiner’. In St. Müller (ed.), *Proceedings of the HPSG03 Conference*, pages 391–396, Stanford: CSLI Publications.