Chapter 11

Complex predicates

Danièle Godard
Université Paris Diderot

Pollet Samvelian
Université Sorbonne Nouvelle

Complex predicates are constructions in which the head attracts arguments from its predicate complement. They characterize auxiliaries, copulas and predicative verbs, certain control or raising and causative verbs, and light verbs. This phenomenon has been studied in HPSG in different languages, Romance and Germanic languages, Korean and Persian, which illustrate different aspects. Romance languages show that argument inheritance is distinct from structure: it is compatible with different structures. German and Dutch and Korean show that argument inheritance induces different word order properties, and Persian that it can be preserved by a derivation rule (nominalization from a verb), and, most importantly in that language which has relatively few simplex verbs, that light verb constructions are used to turn a noun into a verb.

1 What are complex predicates?

Stefan: To make the volume as such more uniform, please start with an introduction explaining what you are doing. A section should never start with a subsection right away.

1.1 Definition

The term *complex predicate* does not have a universally accepted definition. In the HPSG tradition, a complex predicate is composed of two or more words, which

Danièle Godard & Pollet Samvelian. 2020. Complex predicates. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 381–445. Berlin: Language Science Press. DOI: ??

are predicates. By predicate, we mean either a verb or a word of a different category (noun, adjective, preposition) which is associated with an argument structure. A complex predicate is a construction in which the head attracts the arguments of the other predicate, which is its complement: the arguments selected by the complement predicate "become" the arguments of the head (Hinrichs & Nakazawa 1989; 1998). The phenomenon is called *argument attraction*, *composition*, *inheritance* or *sharing*.

To take an example, tense auxiliaries and the participle in Romance languages are two different words, since they can be separated by adverbs, as in (1), but the two verbs belong to the same clause, and, more precisely, the syntactic arguments belong to one argument structure. We admit that the property of monoclausality can manifest itself differently in different languages (Butt 2010). In the case of Romance auxiliary constructions, the first verb (the auxiliary) hosts the clitics which pronominalize the arguments of the participle: corresponding to the NP complement in (1a), the pronominal clitic *l(e)* is hosted by the auxiliary *a* in (1b), (1c). This contrasts with the construction of a control verb such as *vouloir* 'to want' in French, where the clitic corresponding to the argument of the infinitive is hosted by the infinitive (2) (from Abeillé & Godard 2002: 406):

(1) a. Paul a rapidement lu son livre. (French)
Paul has quickly read his book
'Paul has quickly read his book.'

b. Paul l'a rapidement lu.
Paul it-has quickly read
'Paul has quickly read it.'

c. * Paul a rapidement le-lu.
Paul has quickly it-read
Intended: 'Paul has quickly read it.'

(2) a. Paul veut lire son livre. (French)
Paul wants read his book
'Paul wants to read his book.'

b. Paul veut le lire.
Paul wants it-read
'Paul wants to read it.'

c. * Paul le veut lire.[1]
Paul it-wants read
Intended: 'Paul wants to read it.'

This approach to complex predicates goes back to Relational Grammar (Aissen & Perlmutter 1983): although formalized in a different way, their analysis of causative constructions in Romance languages relies on such argument attraction, under the name of *clause union*. Similarly, in Lexical Functional Grammar, Andrews & Manning (1999) speak of complex predicates as building a domain of grammatical relations sharing. It is also present in Categorial Grammar (Geach 1970), with complex categories whose definition takes into account the nature of the argument they combine with, and the operation of function attraction. In particular, Kraak (1998: 301) accommodates complex predicates by introducing a specific mode of combination called *clause union mode*, where two verbs (two lexical heads) are combined. But, in this account, there is no argument attraction in general, the mechanism being specifically defined in order to account for clitic climbing.

There are other definitions of complex predicates. The term has been used to describe the complex content of a word, when it can be decomposed. For instance, the verb *dance* has been analyzed as incorporating the noun *dance* and considered a "complex predicate" (Hale & Keyser 1997). In the sense adopted here, complex predicates involve at least two words, and are syntactic constructions. Closer to what we consider here complex predicates is the case of Japanese passive or causative verbs, illustrated in (3).

(3) tabe-rare-sasete-i-ta. (Japanese)
eat-PASS-CAUS-PROGR-PAST
'(someone) was causing (something) to be eaten.'

The causative morpheme adds a causer argument, and behaves as if it took the verb stem as its complement (more precisely the verb stem with the passive morpheme in this case), whose expected subject appears as the object of the causative verb. This operation is like argument attraction. However, it happens in the lexicon rather than in syntax: the elements in (3) are bound morphemes, and they form a word (Manning et al. 1999; Gunji & Hasida 2012). Thus, we do not consider causative verbs in Japanese to constitute complex predicates.

Complex predicates are sometimes given a semantic definition: the two elements together describe one situation (Butt 1995). Such a semantic definition

[1]Possible in an earlier stage of French.

does not coincide with the syntactic one. It is true that the head verb of a complex predicate tends to add tense, aspectual or modal information while the other element describes a situation type. Thus, in (1), the two verbs jointly describe one situation, the auxiliary adding tense and aspect information. But the semantics of a complex predicate is not always different from that of ordinary verbal complements. Thus, there is no evident semantic distinction depending on whether the Italian restructuring verb *volere* 'to want' is the head of a complex predicate (4a) or not (4b), and the two verbs do not seem to describe just one situation (Monachesi 1998: 314).

(4) a. Anna lo vuole comprare. (Italian)
Anna it wants buy
'Anna wants to buy it.'

b. Anna vuole comprarlo.
Anna wants buy.it
'Anna wants to buy it.'

The same point is made for Hindi in Poornima & Koenig (2009). They show that there exist two structures combining an aspectual verb and a main verb; in one of them the aspectual verb is the head of a complex predicate while, in the other one, it is a modifier of the main verb. In more general terms, complex predicates show that syntax and semantics are not always isomorphic in a language. Thus, although the semantic definition of complex predicates may be useful for some purposes, we will ignore it here.

The distinction between complex predicates and *Serial verb constructions* (SVC), illustrated in (5) (from Haspelmath 2016: 294), is not evident (e.g. Andrews & Manning 1999; Haspelmath 2016). The main reason is that the constructions which have been dubbed SVC are different in different languages: following Andrews & Manning (1999), they do not share a grammatical mechanism, but more superficial tendencies, such as their resemblance to paratactic constructions, due to the absence of marking of complementation or coordination, and involve more semantic relations than are usually associated with complementation and coordination.

(5) Òzó sàán rrá ógbà. (Edo)
Ozo jump cross fence
'Ozo jumped over the fence.'

Accordingly, SVC are not within the purview of complex predicates, and will not be studied in this chapter.

1.2 Constructions involving complex predicates

Complex predicates enter into a number of constructions across languages. They differ from ordinary constructions by different properties, depending on the construction, such as the position of pronominal clitics in Romance languages ("clitic climbing"), word order, or special semantic combinations.

The following have been particularly studied in HPSG:

- Tense auxiliaries, the copulas and other verbs taking predicative complements, restructuring verbs, headed by certain subject raising or control verbs, and certain causative and perception verbs in Romance languages (Abeillé & Godard 1994; 2000; 2001a,b; 2002; 2010; Abeillé et al. 1995; Abeillé, Godard, Miller & Sag 1998; Abeillé, Godard & Sag 1998; Monachesi 1998);
- Certain constructions in German and Dutch, called coherent constructions, headed by tense auxiliaries, certain raising and control verbs, certain verbs with predicative complements, as well as the copula and particle verbs (Hinrichs & Nakazawa 1989; 1994; Rentier 1994; Kiss 1994; 1995; Bouma & van Noord 1998; Hinrichs & Nakazawa 1998; Kathol 1998; 2000; Meurers 2000; 2001; De Kuthy & Meurers 2001; Müller 2002; 2003; 2018);
- Causatives in different languages (among which German, Italian, Turkish), including both analytical causatives (complex predicates in the sense adopted her) and synthetic causatives (Webelhuth 1998).
- Korean auxiliaries, control verbs and *ha* causative verb (Chung 1998; Sells 1991; Yoo 2003; Kim 2016);
- Hindi aspectual predicates (Poornima & Koenig 2009).
- Light verb constructions (combination of a semantically light verb with a predicate belonging to diverse categories) in Persian (Bonami & Samvelian 2010; Müller 2010; Samvelian 2012; Bonami & Samvelian 2015), and Korean (Ryu 1993; Lee 2001; Choi & Wechsler 2002; Kim 2016).

In this chapter, we examine some of these constructions which illustrate the different ways in which complex predicates differ from ordinary verbs with their complements.

2 The basic mechanism in HPSG: Argument attraction

In HPSG, complex predicates are analyzed in the following way: one of the predicates is the head of the construction, and it attracts the syntactic arguments of the other predicate, that is, its complements and, possibly, its subject. The phenomenon is called *argument attraction*, *composition*, *inheritance*, *raising*, or *sharing*. We illustrate it with tense auxiliaries in French (Abeillé & Godard 1994; 2002).

In French, auxiliary constructions consist of a tense auxiliary (*avoir* 'to have' or *être* 'to be' followed by a past participle and its complements as illustrated in (1). The auxiliary is the head: it bears inflectional affixes (for tense and person), like any other verb; in (1), it has the form of a present indicative, 3rd person; it is in the indicative, as expected in a declarative sentence. It hosts pronominal clitics, like verbal heads in general (1b), (1c). Moreover, it can be gapped alone (6a), while the participle can only be gapped with the auxiliary (6b), (6c)[2]; this is expected if the auxiliary is the head, since it behaves like *pense* 'think' (6d) while the participle behaves like the infinitive in (6e), (6f).

(6) a. Lola a acheté des pommes, et Alice (a) cueilli des pêches.
Lola has bought some apples and Alice has picked some peaches
'Lola has bought apples, and Alice (has) picked peaches.'

b. Lola a acheté des pommes, et Alice (a acheté) des pêches.
Lola has bought some apples and Alice has bought some peaches
'Lola has bought apples, and Alice (has bought) peaches.'

c. # Lola a acheté des pommes, et Alice a des pêches.
Lola has bought some apples and Alice has some peaches
'Lola has bought apples, and Alice has peaches.'

d. Lola pense acheter des pommes, et Alice (pense) cueillir des pêches.
Lola thinks buy some apples and Alice thinks pick some peaches
'Lola is thinking of buying apples, and Alice (is thinking of) picking peaches.'

e. Lola pense acheter des pommes, et Alice (pense acheter) des
Lola thinks buy some apples and Alice thinks buy some

[2]Note that (6c) is acceptable with the possession verb *avoir*.

pêches.
peaches
'Lola is thinking of buying apples, and Alice (is thinking of picking) peaches.'

f. * Lola pense cueillir des pommes et Alice pense des pêches.
Lola thinks pick some apples and Alice thinks some peaches
Intended: 'Lola is thinking of picking apples and Alice is thinking of (picking) peaches.'

The auxiliary construction in French is a complex predicate: The clitic corresponding to a complement of the participle is hosted by the auxiliary (it is said to "climb") as in (1b); moreover, it occurs in bounded dependencies such as the infinitival complement of adjectives like *facile* 'easy', whose nominal complement is unexpressed (7a); this unexpressed complement can be that of a participle (7c) but not that of an infinitive complement (7b). This follows if this unexpressed complement is in fact treated as the complement of the auxiliary.

(7) a. Cette technique est impossible à maîtriser en un jour. (French)
this technique is impossible to master in one day
'This technique is impossible to master in one day.'

b. * Cette technique est impossible à réussir à maîtriser en un jour.
this technique is impossible to manage to master in one day
Intended: 'This technique is impossible to manage to master in one day.'

c. Cette technique est impossible à avoir maîtrisé en un jour.
this technique is impossible to have mastered in one day
'This technique is impossible to have mastered in one day.'

These two properties follow if the complements of the participle become those of *avoir* 'to have': the auxiliary "attracts" the complements of the participle. In addition, the tense auxiliary *avoir* 'to have' is a subject raising verb (see Abeillé (2020), Chapter 12 of this volume): the subject is selected by the participle and shared by the auxiliary. For instance, *Paul* is an agent in (1a) (*Paul a lu son livre*, 'Paul has read his book' because *lire* 'to read' requires an agent subject, and it is the impersonal subject *il* in *Il a fait froid* (lit. It has made cold, 'It [the weather] was cold'), because the subject of *faire froid* is impersonal. Thus, the auxiliary *avoir* (like tense auxiliary *être* 'to be') is, in fact, a generalized raising verb: its whole argument structure is identified with that of the participle. A simplified

description of subject raising verbs and tense auxiliaries is given in (8) (for the feature [LIGHT ±], see Section 3).

(8) a. Subject raising verb:

$$\left[\text{ARG-ST } \boxed{1} \oplus \left\langle \left[\text{SUBJ } \boxed{1}\right] \right\rangle \oplus \boxed{2}\right]$$

b. Tense auxiliary:

$$\left[\text{ARG-ST } \boxed{1} \oplus \left\langle \begin{bmatrix}\text{ARG-ST } \boxed{1} \oplus \boxed{2} \\ \text{LIGHT } +\end{bmatrix} \right\rangle \oplus \boxed{2}\right]$$

The subject raising verb takes a complement saturated complement, which is described as the second element of the argument structure, expecting a subject $\boxed{1}$ identified with the subject of the raising verb. The notation $\boxed{1}$ without $\langle\ \rangle$ indicates that this element may be absent: it is meant to accommodate subjectless verbs. In addition, the raising verb may have its own complements, noted here $\boxed{2}$. On the other hand, the auxiliary is not only a subject raising verb, but takes as a complement a participle which has not combined with any complements and only has attracted complements.

The arguments of a word are made up of subject and complements. The relation between (expected) arguments and realized subject and complements, is as in (9) (see Ginzburg & Sag 2000: 171; Bouma et al. 2001). The arguments include the subject, the complements (and the specifier), but also a list of non-canonical elements (possibly empty) (see below).

(9) Argument Realization Principle

$$\textit{word} \Rightarrow \left[\text{SYNSEM} \mid \text{LOC} \left[\begin{array}{l}\text{CAT} \begin{bmatrix}\text{SUBJ } \boxed{1} \\ \text{COMPS } \boxed{2} \\ \text{SPR } \boxed{3}\end{bmatrix} \\ \text{ARG-ST } \boxed{1} \oplus \boxed{2} \oplus \boxed{3} \bigcirc \text{list}(\textit{non-canon})\end{array}\right]\right]$$

In (10a), the participle *lu* 'read' selects the argument *son livre* 'her book', which is attracted by the auxiliary *a* 'has'. Accordingly, it is realized as the complement of the auxiliary *a*. The structure of the VP in (10a) is given in Figure 1.

(10) a. Marie a lu son livre.
Marie has read her book
'Marie has read her book.'

b. Marie l'a lu.
Marie it.has read
'Mary has read it.'

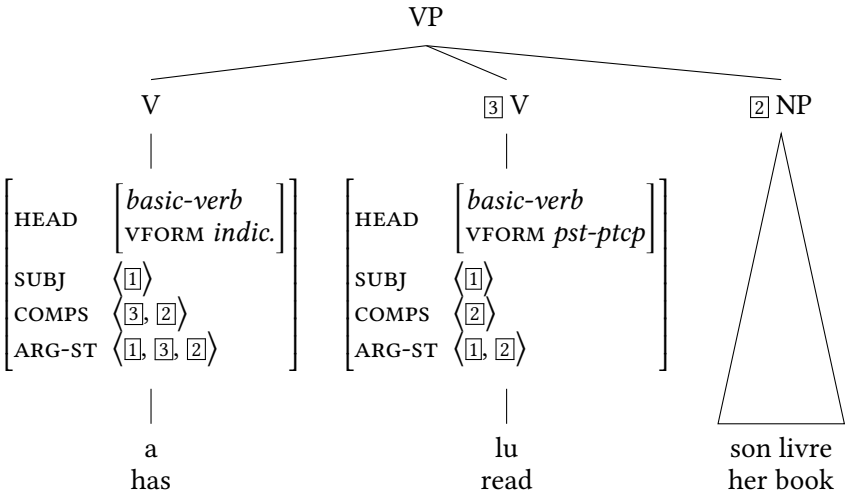


Figure 1: VP structure in French

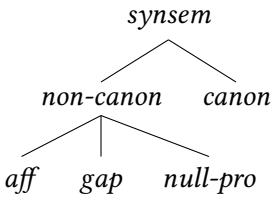


Figure 2: Subtypes of *synsem*

Let us turn to pronominal clitics. The arguments are *synsems*, which can have different subtypes (Figure 2). Usually, they are not specified on the lexeme description, but they are on words.

Romance clitics, illustrated by *l(e)* in (10b), are analyzed as affixes (*aff*) on verbs, which correspond to arguments of the verb (Miller & Sag 1997). They belong to the argument structure of the participle, and are attracted by the auxiliary, although they are not realized as complements. In (10b) and Figure 3, the arguments of the auxiliary are the subject 1, the participle 3, and 2; 2 is typed as an affix, 3rd person, masculine singular. It belongs to the argument structure, but not to the complement list of the auxiliary (see (9)).

We distinguish between *basic verbs* and *reduced verbs*, following Abeillé, Godard & Sag (1998). With basic verbs, the argument list is simply the concatenation of the subject and complements, while reduced verbs have at least one affix ar-

gument which belongs to the argument list, but not to the complement list. Such verbs are subject to a morphological rule which realizes this affixal argument as an affix, the so-called clitic pronoun *l(e)*. Thus, in Figure 1, both the auxiliary *a* 'has' and the participle *lu* 'read' are basic verbs: the arguments tagged [3] and [2] are also complements. On the other hand, in Figure 3, the participle is a basic verb – argument [2] is typed as an affix, but is also a complement – while the auxiliary is a reduced verb: argument [2] is not a complement of the auxiliary, and the verb hosts the affix *l(e)*.

In French, past participles never host clitics (1c), which we assume to be a morphological property. But, in Italian, past participles may host clitics, although never when they combine with the auxiliary. The specification that the participle complement of the auxiliary is a basic verb accounts for this property, because basic verbs are not the target of the morphological rule realizing the affixal argument as an affix. Although both verbs in Figure 3 have an affixal argument, one is a basic verb (the participle), the affixal argument being also an expected complement, and the other is a reduced verb (the auxiliary), this affixal argument not being an expected complement.[3]

3 Different structures for complex predicates: Restructuring verbs and the copula in Romance languages

In addition to tense auxiliaries, Romance languages have other cases of complex predicates: they are headed by restructuring verbs, by the copula and other verbs taking predicative complements, and by certain causative and perception verbs.

[3]It is worth noting that tense auxiliaries can take as complement a coordination of participles:

(i) Jean a acheté et lu ce livre.
Jean has bought and read this book
'Jean bought and read this book.'

(ii) Jean l'a acheté et lu.
Jean it-has bought and read
'Jean bought and read it'

This may be seen as raising a difficulty for the analysis of their complement based on argument structure sharing, since argument structure characterizes words rather than phrases. However, coordinations of words are a special kind of phrases, since the conjuncts must share their argument structure. It is plausible that such coordinations inherit an argument structure from the conjuncts.

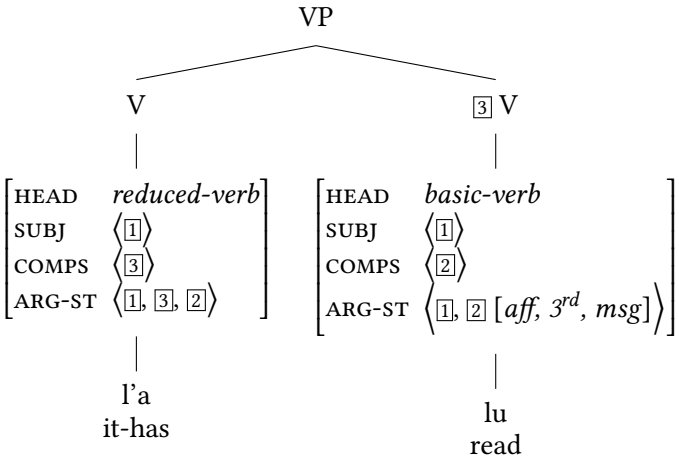


Figure 3: Clitic climbing in French

We focus here on restructuring verbs and the copula. An analysis of causative and perception verbs is proposed in (Abeillé et al. 1995; Abeillé, Godard, Miller & Sag 1998; Abeillé & Godard 2010).

A comparison of the properties of constructions headed by restructuring verbs in different Romance languages illustrates an important aspect of the phenomenon: argument attraction is compatible with different syntactic structures. Restructuring verbs enter either a flat structure, or a verbal complex (Monachesi 1998; Abeillé & Godard 2001a; 2010). As for the copula, it differs from tense auxiliaries and restructuring verbs in two respects: its complement always behaves like a phrase, although it can be fully saturated for its complements, partially saturated, or not saturated at all (Abeillé & Godard 2001b; 2002); and it has a uniform behavior and analysis across the Romance languages.

3.1 Romance restructuring verbs as head of complex predicates

Certain verbs in Romance languages, called *restructuring verbs*, exhibit two behaviors: either as ordinary verbs taking a VP complement or as heads of complex predicates (Rizzi 1982; Aissen & Perlmutter 1983). Restructuring verbs are modal, aspectual, or movement verbs (such as *venire* 'to come', *andare* 'to go', *correre* 'to run', *tornare* 'to come back' in Italian). However, it must be kept in mind that this behavior is lexical: verbs which are close semantically may not be heads of complex predicates.

Several properties show that they may head complex predicates (Monachesi 1998: 323-328). First, clitic climbing, which is optional (while it is obligatory with tense auxiliaries). The examples in (11) all mean 'John wants to eat them' (examples from Abeillé & Godard 2010: 113). For each language, the first example illustrates the complex predicate, and the second one the VP complement construction, with the clitic downstairs.

(11) a. Giovanni *le* vuole mangiare. (Italian)
Giovanni them wants eat
'Giovanni wants to eat them.'

b. Giovanni vuole mangiar*le*.
Giovanni wants eat.them
'Giovanni wants to eat them.'

c. Juan *las* quiere comer. (Spanish)
Juan them wants eat
'Juan wants to eat them.'

d. Juan quiere comer*las*.
Juan wants eat.them
'Juan wants to eat them.'

e. O João quere-*as* comer. (Portuguese)
o João wants-them eat
'João wants to eat them.'

f. En Joan *les* vol menjar. (Catalan)
en Joan them wants eat
'Joan wants to eat them.'

g. En Joan vol menjar-*les*.
en Joan wants eat-them
'Joan wants to eat them.'

Second, the medio-passive or middle *si* construction, where the verb hosts the reflexive clitic *si* or *se* (12b) (depending on the language), and the subject corresponds to the object of the active construction (12a), with an interpretation close to that of passive. The construction is possible with restructuring verbs such as *potere* 'to be able to' (12c), (12d) (see Monachesi 1998: 333-336), but not with verbs only taking an infinitival VP complement such as *parere* 'to appear' (12e) (examples (12d) and (12e) from Abeillé & Godard 2010: 122).

(12) a. Giovanni stira queste camicie facilmente. (Italian)
Giovanni irons these shirts easily
'Giovanni irons these shirts easily.'

b. Queste camicie si stirano facilmente.
these shirts SI iron easily
'These shirts iron easily.'

c. Giovanni può stirare queste camicie facilmente.
Giovanni can iron these shirts easily
'Giovanni can iron these shirts easily.'

d. Queste camicie si possono stirare facilmente.
these shirts SI can iron easily
'These shirts can be ironed easily.'

e. * Queste camicie si paiono stirare facilmente.
these shirts SI appear iron easily
Intended: 'These shirts appear to be ironed easily.'

The medio-passive verb alternates with a transitive verb: it is the result of a Lexical Rule (13), which takes a transitive verb like *stirare* (12a) to give a verb whose subject corresponds to the expected object of the transitive verb, and acquires a reflexive clitic noted as *a-aff* (realized *si* or *se*) as in (12b) (Abeillé, Godard & Sag 1998: 31; Monachesi 1998).

(13) Medio-Passive Lexical Rule

$$\left[\text{ARG-ST}\left\langle \text{NP}, \text{NP}\ [acc]_j \right\rangle \oplus \boxed{1}\right] \mapsto \left[\text{ARG-ST}\left\langle \text{NP}_j, [a\text{-}aff,\ acc]_j \right\rangle \oplus \boxed{1}\right]$$

What is crucial here is that the input is a verb taking an accusative NP complement. Hence, a verb taking a VP complement like Italian *potere* 'to be able to' or *parere* 'to appear' cannot be the input, since it lacks an NP complement. On the other hand, the corresponding restructuring verb *potere* can since it inherits such a complement from the infinitive: the verb *potere* in (12c) inherits *queste camicie* 'these shirts' from *stirare* 'to iron', and is the input to Rule (13), giving the verb which occurs in (12d). On the other hand, the verb *parere* which is not a restructuring verb does not have an NP object and cannot be the input to this Rule (13).

The third relevant property is their acceptability in bounded dependencies, as illustrated in (7) for tense auxiliaries, and (14) for restructuring verbs. (14b) (from Monachesi 1998: 341) relies on *cominciare* 'to begin' being a restructuring verb, while *promettere* 'to promise' is not (14c).

(14) a. Questa canzone è facile da apprendere. (Italian)
this song is easy to learn
'This song is easy to learn.'

b. Questa canzone è facile da cominciare a apprendere.
this song is easy to begin to learn
'This song is easy to begin to learn.'

c. * Questa canzone è facile da promettere di apprendere.
this song is easy to promise to learn
Intended: 'This song is easy to promise to learn.'

The complement of adjectives such as 'easy' in Romance languages is a bounded dependency: they take an infinitival complement whose own expected complement (we analyze it as a null pronoun, see Figure 2) is coindexed with its subject (Abeillé, Godard & Sag 1998; Monachesi 1998).[4]

(15)
$$\left[\begin{array}{ll} \textsc{head} & \textit{adjective} \\ \textsc{arg-st} & \left\langle \textsc{xp}_j,\ \textsc{vp} \left[\begin{array}{ll} \textsc{vform} & \textit{infinitive} \\ \textsc{marking} & \textit{da} \\ \textsc{comps} & \left\langle \textit{null-pro}\ \left[\textit{acc}\right]_j \right\rangle \oplus \boxed{2} \end{array}\right] \right\rangle \end{array}\right]$$

Complex predicates can occur in this construction because their head attracts the complement of their complement. Thus, in (14b), *cominciare* 'to begin' is expecting the same object as *apprendere* 'to learn', which is coindexed with the subject of the copular construction, in the same way as *apprendere* is expecting an object in (14a).

Finally, the possibility of preposing the verbal complement of a verb which can take a VP complement or be the head of a complex predicate disappears when there is evidence of a complex predicate. For the sake of simplification, we now concentrate on Italian and Spanish. The data in (16), with a preposed VP, contrast with those in (17) (both examples from Abeillé & Godard 2010: 132), where the head verb bears a clitic corresponding to the expected complement of the infinitive. Preposing of the verbal complement is associated with pronominalization (*lo*) in Italian (16a) not in Spanish (16b), where it is more natural in contrastive contexts.

[4] Forms such as *a*, *da*, *di*, which introduce infinitival complements in (14) are not analyzed as heads, but as markers, a part of speech which has the feature MARKING, whose value is specific to the form. Markers select the head with which they combine (for instance, *da* selects an infinitival VP in (14a)), and the feature is shared by the whole VP. Hence, the adjective *facile* 'easy' in Italian takes as a complement an infinitival VP [MARKING *da*].

(16) [Context] Does he want to talk to Mary?

a. Parlare a Maria, certamente lo vuole. (Italian)
talk to Maria certainly it wants
'Talk to Maria, certainly he wants to.'

b. Hablarle a María, seguramente quiere (pero no a su madre).
talk.to.her to María certainly wants but not to her mother
(Spanish)
'Talk to Maria, certainly he wants to (but not to her mother).'

(17) a. * Parlare, certamente glielo vuole. (Italian)
talk certainly to.him/it wants
Intended: 'Talk to him, he certainly wants to.'

b. * Hablar, le quiere (pero no mucho tiempo). (Spanish)
talk to.him/her wants but not a.long time
Intended: 'Talk to him/her he wants to (but not for a long time).'

We assume that restructuring verbs have two possible descriptions: as ordinary verbs taking an infinitival VP complement, or as heads of complex predicates. They are related by the Argument Attraction Lexical Rules (18) (adapted from Monachesi 1998: 331).[5]

(18) Argument Attraction Lexical Rules for Romance restructuring verbs

a. Subject control verbs

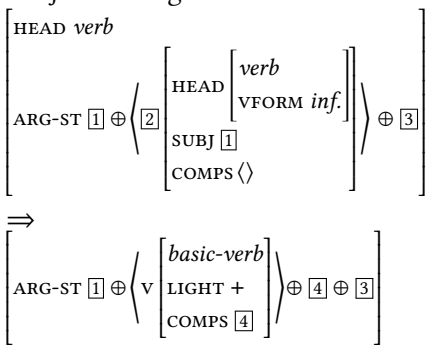
$$\begin{bmatrix} \textsc{head}\ \textit{verb} \\ \textsc{arg-st}\left\langle \textsc{xp}_i, \boxed{2}\begin{bmatrix} \textsc{head}\begin{bmatrix} \textit{verb} \\ \textsc{vform}\ \textit{inf.} \end{bmatrix} \\ \textsc{subj}\left\langle \textsc{xp}_i \right\rangle \\ \textsc{comps}\ \langle\rangle \end{bmatrix} \right\rangle \oplus \boxed{3} \end{bmatrix}$$

$$\Rightarrow$$

$$\begin{bmatrix} \textsc{arg-st}\left\langle \textsc{xp}_i, \textsc{v}\begin{bmatrix} \textit{basic-verb} \\ \textsc{light}\ + \\ \textsc{comps}\ \boxed{4} \end{bmatrix} \right\rangle \oplus \boxed{4} \oplus \boxed{3} \end{bmatrix}$$

[5] We leave aside the object control and object raising verbs (verbs of influence or perception verbs) which can also be the head of a complex predicate, hence be the target of a similar Lexical Rule (Abeillé, Godard, Miller & Sag 1998; Abeillé & Godard 2010), concentrating on the case of so-called *restructuring verbs*.

b. Subject raising verbs



In the input description, the verbal complement is saturated for its complements. The verb may have other complements in addition to the saturated infinitival VP, noted as list 3. We distinguish between subject control verbs and subject raising verbs to accommodate the case where the complement verb is subjectless, but with complements that can be attracted. In (19a), the verb *sembra* 'seems' is a raising verb, and the infinitive *piacere* 'to please' is an impersonal verb with no subject, but with a complement, realized by *gli* on the head verb *sembra* (there is another interpretation where *gli* is the complement of *sembra*, which is irrelevant).[6] Note that there is speakers' variation: *sembrare* 'to seem' is not a restructuring verb for all Italian speakers (hence % on the examples).

The category of the subject of control verbs is not specified: it can be an infinitival VP as well as an NP (or even a sentence); in the first case, the index is that of the situation (19c), in the second, it is the index of the nominal entity (19b). Again, the upstairs clitic *gli* corresponds to the argument of *piacere* 'to please':

(19) a. % Gli sembra piacere molto. (Italian)
to.him seems please a.lot
'It seems that he likes it a lot.'

b. % [Questo regalo] gli sembra piacere.
this gift to.him seems please
'This gift seems to please him.'

[6] Alternatively, in a grammar with null pronouns, impersonal and unaccusative verbs in Romance languages could be analyzed as having a null pronoun subject, a representation which allows a common input for subject control and raising verbs in the Argument Attraction Lexical Rule (as in Monachesi 1998: 331).

c. % [Andare in vacanza] gli sembra piacere
go.away on vacation to.him seems please
'To go away on vacation seems to please him.'

3.2 The different structures of complex predicates with restructuring verbs

The point of this section is to show that argument attraction is compatible with different structures: these are two different phenomena. In Romance languages, restructuring verbs can take a VP complement, or be the head of a complex predicate. In the latter case, there are two possible structures: they enter either a flat structure or a verbal complex. We speak of a flat structure when the complement verb as well as the complements that it subcategorizes for are all sisters of the head. We speak of a verbal complex when the head verb and the complement verb form a constituent by themselves, to the exclusion of their complements (see Figure 4).

We contrast Italian and Spanish.[7] Note that in Spanish, there is variation among speakers: we describe here one usage of Spanish complex predicates.

The impossibility of Preposing illustrated in (17) shows that the sequence of the complement verb and its complements does not form a constituent (a VP) when there is a complex predicate, a point made by Rizzi (1982) for Italian, on the basis of a series of constructions (pied-piping, clefting, Right Node Raising, Complex NP shift). However, the two languages differ with respect to other properties. The presence of a clitic on the head verb indicates that there is a complex predicate.

First, adverbs occur between the restructuring verb and the infinitive in Italian (20a), but not in Spanish in a general way (20b) (a few adverbs, such as *casi* 'nearly', *ya* 'already', *apenas* 'barely' are possible). In Spanish, an adverb may occur after the verb and before the infinitive if the complement is a VP (20c) (examples (20) from Abeillé & Godard 2010: 139).

(20) a. Giovanni *lo* vuole spesso leggere. (Italian)
Giovanni it wants often read
'Giovanni wants to read it often.'

[7] In Portuguese restructuring verb constructions are also a flat structure, but with different ordering constraints from Italian; the other variety of Spanish is similar to Portuguese. Except for the copula (see Section 3.4), complex predicate constructions with head verbs entering only one structure also distribute between these two structures among Romance languages: tense auxiliaries in French, Italian, Portuguese as well as Romanian modal *a putea* 'can' are the head of a flat structure, while tense auxiliaries in the variety of Spanish described here and in Romanian enter a verbal complex (Abeillé & Godard 2010).

b. * Juan *lo* quiere a menudo leer. (Spanish)
Juan it wants often read
Intended: 'Juan wants to read it often.'

c. Juan quiere a menudo leer*lo*.
Juan wants often read.it
'Juan wants to read it often.'

Second, an inverted subject NP can occur between the two verbs of a complex predicate in Italian (21a), but not in Spanish (21b). The subject can occur postverbally in interrogative sentences. In Italian, it can occur between the two verbs with a special prosody, indicated by the small capitals in (21a), and with speaker's variation (Salvi 1980). In Spanish, this is not possible (except for the pronominal subject) (Suñer 1982).

(21) a. % Lo comincia MARIA a capire, il problema, oppure no?
it begins Maria to understand the problem or no
(Italian)
'Maria, she's beginning to understand it, the problem, yes or no.'

b. * ¿Lo comienza Juan a comprender? (Spanish)
it begins Juan to understand
'Is Juan beginning to understand it?'

c. ¿Comienza Juan a comprenderlo?
begins Juan to understand.it
'Is Juan beginning to understand it?'

Finally, Italian heads of complex predicates can have scope over the coordination of infinitives with their complements (22a), while this is not the case in Spanish (22b). Again, the presence of a clitic on the head verb (*lo vuole* lit. it wants, *le volvió* lit. to.him started.again) shows that this is a complex predicate construction (examples from Abeillé & Godard 2010: 136-137).

(22) a. % Giovanni lo vuole comprare subito e dare a Maria.
Giovanni it wants buy immediately and give to Maria
(Italian)
'Giovanni wants to buy it immediately and give it to Maria.'

b. * Le volvió a pedir un autógrafo y a hacer
to.him/her started.again to ask an autograph and to make

proposiciones. (Spanish)
propositions
Intended: 'He started again to ask him for an autograph and to make propositions to him.'

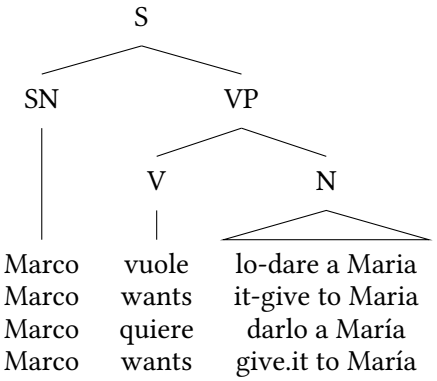
Constituency tests such as preposing (17) show that the verbal complement is not a VP in either language. The verbal complex, in which the two verbs form a constituent without the complements, is well-suited to account for the absence of adverbs (in a general way) and of subject NPs, if such combinations exclude elements other than verbs (adverbs in particular). This constraint can be captured by the feature [LIGHT +][8], which has been used in Romance languages for other phenomena as well (Abeillé & Godard 2000) (see Section 3.3). Hence, complex predicate constructions in Spanish contain a verbal complex, while they form a flat structure in Italian containing the complement verb and its complements.

This is illustrated with examples in Figures 4, which all mean 'Marco wants to give it to Maria'. The verb takes a VP complement in Figure 4a in both languages, it is the head of a flat VP in Italian in Figure 4b, and enters a verbal V-V complex in Spanish in Figure 4c (from Abeillé & Godard 2010: 146).

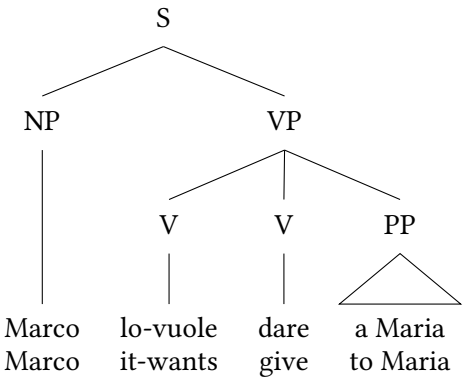
The possibility of the coordination in (22a) has been viewed as an argument in favor of a complement VP even when there is argument attraction (Andrews & Manning 1999). The data go against such an analysis for Spanish, since the coordination is not acceptable. For Italian, although such sequences as (22a) can be analyzed as coordinations of VP, they can also be Non Constituent Coordinations (NCC) (*John gives a book to Maria and discs to her brother*). So, the question becomes: why is (22b) not an acceptable NCC in Spanish? Abeillé & Godard (2010) propose that NCC coordinations are subject to a general constraint in Romance languages: the parallel elements of the coordination must be at the same syntactic level, otherwise the acceptability is degraded. An example is the contrast between (23a) and (23b) in Spanish. The structure of (22b), repeated in (23c), is similar to that of (23b), if it is a verbal complex ((23) from Abeillé & Godard 2010: 137, 144).

(23) a. Juan da [el libro de Proust] [a María] y [el (libro) de Juan gives the book of Proust to María and the book of Camus] [a Pablo]. (Spanish)
Camus to Pablo
'Juan gives the book by Proust to María and the book by Camus to

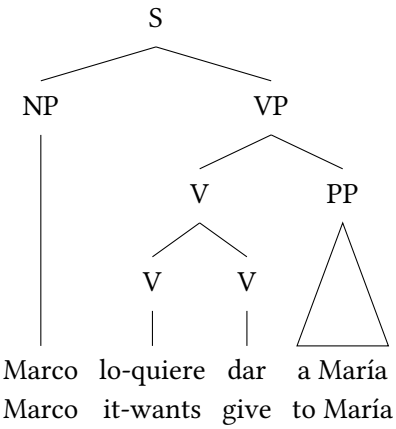
[8]The adverbs admissible in the Spanish verbal complex are light.



(a) VP complement



(b) Flat structure



(c) Verbal complex

Figure 4: Three complementations for Romance restructuring verbs

Pablo.'

b. ?? Juan da [el libro de Proust] [a María] y [de Camus] [a Pablo].
Juan gives the book of Proust to María and of Camus to Pablo
Intended: 'Juan gives the book by Proust to María and the book by Camus to Pablo.'

c. * [Le volvió a pedir] [un autógrafo] y [a hacer] [proposiciones].
to.him/her started.again to ask an autograph and to make propositions
Intended: 'He started again to ask him an autograph and to make propositions to him/her.'

In (23a) the NP *el de Camus* 'the one by Camus' is parallel to and at the same level as *el libro de Proust* 'the book by Proust', the PP *a Pablo* 'to Pablo' is parallel to and at the same level as *a María* 'to María', the NP and the PP are both complements of *da* 'gives'. But, in (23b), *de Camus* 'by Camus' is parallel to *de Proust* 'by Proust', and not at the same level as *el libro de Proust* or as *a Pablo*: *a Pablo* corresponds to the complement of *da* 'gives' while *de Camus* corresponds to the complement of the noun *libro* 'book'. The acceptability is degraded.

If the structure of a complex predicate is that of a verbal complex in Spanish, the structure of (23c) is similar to that of (23b): *a hacer* corresponds to a *a pedir*, which is the complement V of *volvió* in a V-V constituent, and is not at the same level as *proposiciones* which corresponds to *un autógrafo* which is outside the V-V constituent.

3.3 Analysis of Romance restructuring verb constructions in HPSG

It has been shown in Section 3.1 that the different Romance languages all have complex predicate constructions, and, in Section 3.2, that, although they share some properties (such as clitic climbing, occurrence in bounded dependencies), they also show syntactic differences among themselves (separability of the head and the infinitive or participle in Italian, but not in Spanish, possibility of coordination of the complement verb with its complements in Italian, but not in Spanish). The flexibility of HPSG grammar allows us to describe both the commonalities and the differences: the common behavior follows from the fact that they share the mechanism of argument attraction, which characterizes certain

classes of verbs; the differences follow from a different phrase structure: the restructuring verb in Italian enters a flat structure (Figure 4b), while it enters a verbal complex in Spanish (Figure 4c) (Figures 4 from Abeillé & Godard 2010: 146). This analysis contrasts with that of Andrews & Manning (1999) in LFG, who propose that complex predicates in Romance languages arise when two verbs have a common domain of grammatical functions, but correspond to just one phrase structure, all these verbs taking a VP complement. It is not clear how they can account for the differences between the two languages.

Two phrase structure rules combining a head with its complements account for the distinction between the flat structure and the verbal complex: the usual head-complements phrase, and a different one, the head-cluster phrase, also used in German (see Section 4.1.2). The difference between the flat structure and the verbal complex is attributed to the feature [LIGHT ±].

The *head-complements-phrase* is defined as follows:

(24) *head-complements-phrase* (Romance languages) ⇒

$$\begin{bmatrix} \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{COMPS} & \boxed{3} \\ \text{LIGHT} & - \end{bmatrix} \\ \text{HEAD-DTR} \mid \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{COMPS} & \boxed{2} \bigcirc \boxed{3} \\ \text{LIGHT} & + \end{bmatrix} \\ \text{NON-HEAD-DTRS} \quad \boxed{2} \text{ non-empty list} \end{bmatrix}$$

The head-complements-phrase is usually saturated for the expected complements, but not always: list [3] is usually empty, but does not have to be (see the case of the copula in Section 3.4). An example of the flat structure with a restructuring verb is given in Figure 5.

In the flat structure the head verb takes as complements the infinitival verb and the canonical complements expected by the infinitive, and combines with them (Figure 5). The VP, corresponding to the head-complements-phrase, is complement saturated.

The verbal complex corresponds to another kind of head-complements-phrase, called the head-cluster-phrase, given in (25) (see Müller 2002: 6; Müller 2018: 39).[9]

(25) *head-cluster-phrase* (Spanish) ⇒

[9] This rule is also used in Romanian. As in German, we do not specify the category of the complement (which can be a noun in Spanish, for instance).

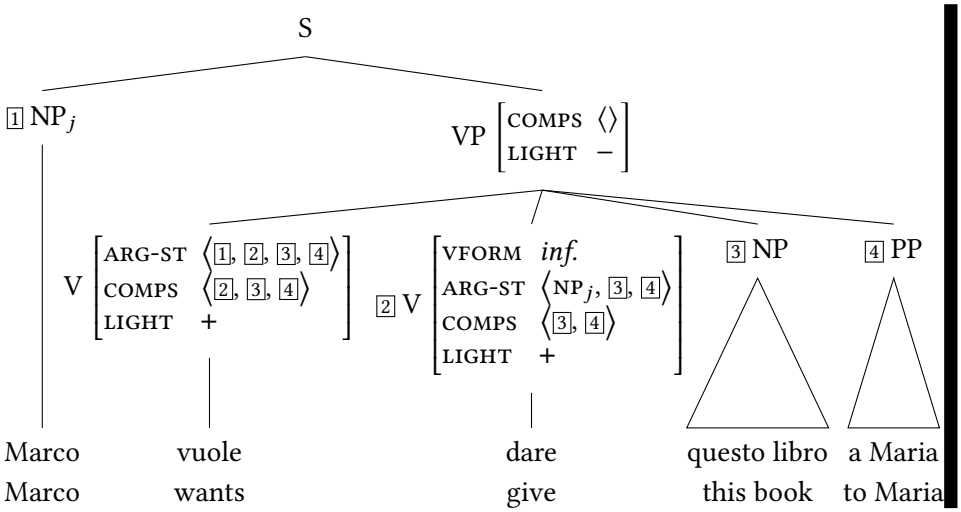


Figure 5: Flat VP structure with an Italian restructuring verb

$$\begin{bmatrix} \textsc{synsem} \mid \textsc{loc} \mid \textsc{cat} & \begin{bmatrix} \textsc{comps} & \boxed{1} \\ \textsc{light} & + \end{bmatrix} \\ \textsc{head-dtr} \mid \textsc{synsem} \mid \textsc{loc} \mid \textsc{cat} & \begin{bmatrix} \textsc{head} & \textit{verb} \\ \textsc{comps} & \boxed{1} \oplus \langle \boxed{2} \rangle \\ \textsc{light} & + \end{bmatrix} \\ \textsc{non-head-dtrs} & \left\langle \left[\textsc{synsem} \boxed{2} \left[\textsc{light} + \right] \right] \right\rangle \end{bmatrix}$$

It differs from the usual head-complements-phrase on two accounts: there is only one non-head daughter, and all the constituents are [LIGHT +]. The LIGHT feature (Bonami & Webelhuth 2012) renames the WEIGHT feature proposed in Abeillé & Godard (2000), as well as the LEX feature used in German (e.g. Hinrichs & Nakazawa 1989; 1994; Kiss 1995; Meurers 2000; Müller 2002; Höhle 2018). The LIGHT feature has ordering as well as structural consequences (Abeillé & Godard 2000; 2010). It is appropriate both for words and phrases. Words can be light or non-light; lexical verbs (finite verbs, participles or infinitives without complements) are light. Most phrases are non-light; in particular, the VP, that is, the phrase which combines with the subject in Romance languages, is non-light.[10] But some phrases can be light if they are composed of light constituents. Such is

[10] Note that the Head-only phrase is non-light. Hence, the VP which dominates a lexical verb only is non-light.

the case of the head-cluster phrase.

The head-cluster-phrase is illustrated in Figure 6: the phrase *quiere dar* corresponds to the head-cluster-phrase, while the whole VP (*quiere dar aquel libro a María* 'wants to give that book to María') corresponds to the usual head-complements-phrase in (24).

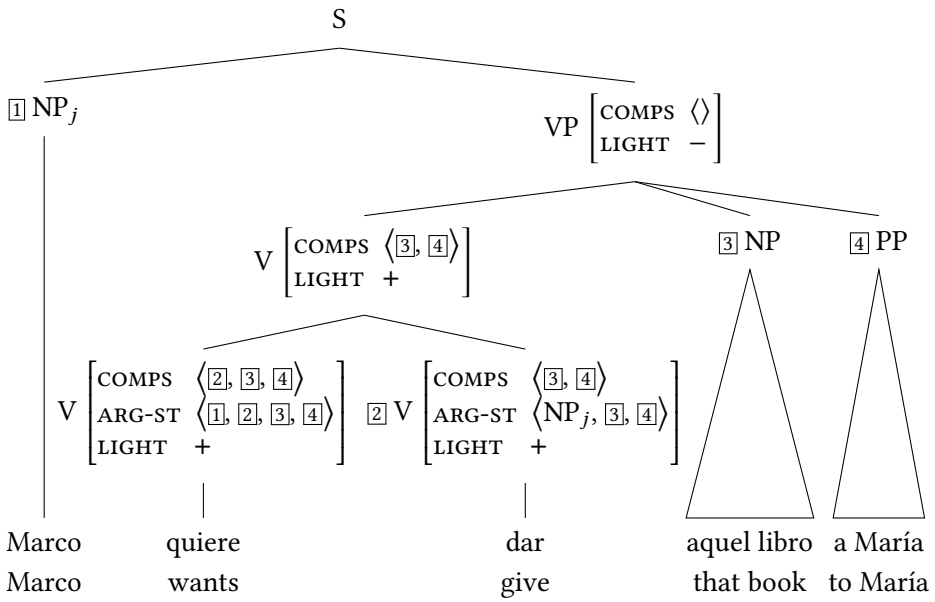
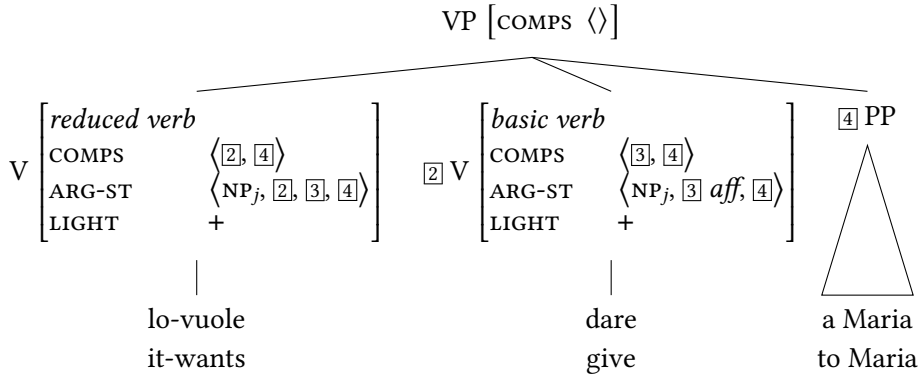


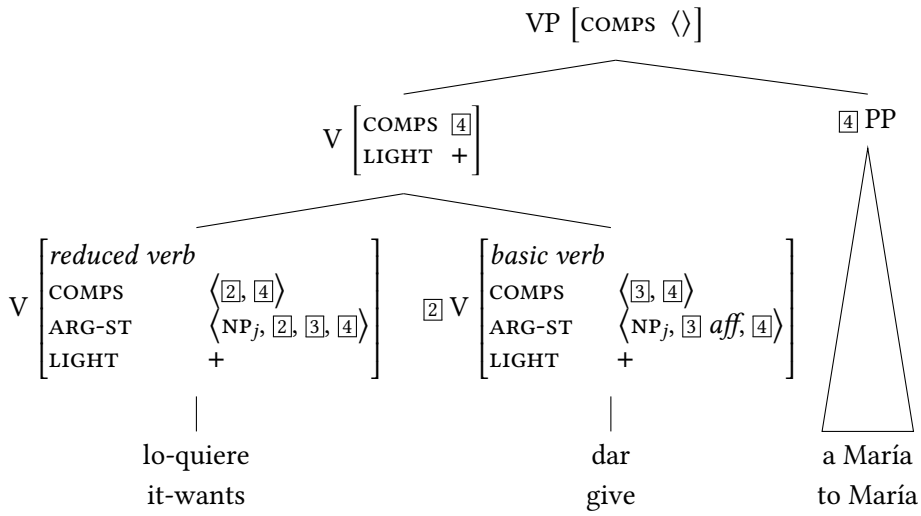
Figure 6: VP with a verbal complex with a Spanish restructuring verb

Regarding the canonical complements in the verbal complex construction, the requirement is passed up by the verbal complex, according to the description in (25) (the list 1 is non empty). The verbal complex itself combines with the canonical complements expected by the infinitive (here, 3 and 4).

More has to be said regarding the clitic *lo* in Italian *Marco lo-vuole dare a Maria* 'wants to give it to Maria' and Spanish *Marco lo-quiere dar a María* 'wants to give it to María') in Figure 4. The infinitive is a basic verb: there is no difference between the complements and the arguments (except for the subject); its complement list contains an affixal element (see Section 2). Following Rule (18a), this element is attracted to the argument list of the head verb, but it is not realized as a complement, as is expected given Principle (9); the head verb is then a reduced verb (see Figures 7), which is the target of a morphological rule of cliticization, hence the clitic *lo* 'it' on the head verb *vuole* or *quiere* 'wants'.



(a) Italian clitic climbing



(b) Spanish clitic climbing

Figure 7: Italian and Spanish clitic climbing with Italian and Spanish restructuring verbs

It remains to ensure that Spanish restructuring verbs are characterized by a verbal complex, and Italian ones by a flat structure. We assume an additional constraint on phrases in Spanish. According to (26), if the phrase is light, it follows that the non-head daughters are also light, and, conversely, if the phrase is non-light, the non-head daughters are non-light.

(26) $\left[\begin{array}{l}\textit{phrase}\\ \textsc{light}\ \boxed{1}\end{array}\right] \Rightarrow \left[\textsc{non-head-dtrs} \mid \textsc{light}\ \boxed{1}\right]$ (Spanish)

The structure of the flat VP does not obey this constraint: the infinitival verb which is a non-head daughter is light, while the other complements are non-light (see Figure 5). When constraint (26) applies, he head of a restructuring verb cannot enter a flat structure.

In order to prevent a verbal complex in Italian (and French), we assume a different additional constraint.

(27) $\left[\begin{array}{l}\textit{phrase}\\ \textsc{light}\ +\end{array}\right] \Rightarrow \left[\textsc{non-head-dtrs}\ \text{list}\ ([\textsc{head}\ \neg\ \textit{verbal}])\right]$ (Italian)

If the structure included a verbal complex in Italian, it would be light, but this is not possible because constraint (27) says that the non-daughter in a light phrase cannot be verbal, which excludes the participle or the infinitive.

Romance languages follow the general constraints on ordering in non-head final languages. According to constraint (28), the verb precedes the complements it subcategorizes for. This is relevant not only for the head of the complex predicate, but also for the participle complement of the tense auxiliary or the infinitive complement of a restructuring verb. Although the latter do not combine with their expected complements, they still subcategorize for them.

(28) $\left[\textsc{v}\left[\textsc{comps}\left\langle \ldots, \boxed{1}, \ldots \right\rangle\right]\right] < \left[\textsc{synsem}\ \boxed{1}\right]$ (Romance languages)

3.4 The complements of the copula in Romance languages

It is an interesting fact that, while Romance restructuring verbs enter two different structures (the flat structure and the verbal complex), the copula has the same complementation across Romance languages (Abeillé & Godard 2001b; 2010)[11]. Moreover, this complementation differs both from the flat structure and the verbal complex: the copula takes a non-light complement, which can be saturated or not.

[11]We concentrate on the predicative use of the copula.

The complement of the copula is underspecified: it is predicative (noted [PRD +]), but can be an adjective, a noun, a preposition or a passive participle (for the passive construction, see Abeillé & Godard 2002). We illustrate clitic climbing with the same example in different Romance languages (examples from Abeillé & Godard 2010: 120).

(29) a. Jean lui était fidèle. (French)
Jean to.him/her was faithful
'Jean was faithful to him/her.'

b. Giovanni le era fedele. (Italian)
Giovanni to.her was faithful
'Giovanni was faithful to her.'

c. Juan le era fiel. (Spanish)
Juan to.him/her was faithful
'Juan was faithful to him/her.'

d. O Joãn era-lhe fiel. (Portuguese)
o Joãn was-to.him/her faithful
'O Joãn was faithful to him/her.'

e. En Joan li era fidel. (Catalan)
en Joan to.him/her was faithful
'En Joan was faithful to him/her.'

f. Ion îi era credincios. (Romanian)
Ion to.him/her was faithful
'Ion was faithful to him/her.'

The properties of the construction differentiate it clearly from tense auxiliaries and restructuring verbs. For the sake of simplification, we restrict the examples to French, Italian and Spanish. The sequence of the head of the complement with its complements is a constituent, since, for instance, it can be dislocated and pronominalized (30) (examples (30)-(31) from Abeillé & Godard 2010: 133-134).

(30) [Context] Is John faithful to his friends?

a. Fidèle à ses amis, il l'est plus qu'à ses convictions
faithful to his friends he it.is more than to his convictions
politiques. (French)
political
'Faithful to his friends, he is, more than to his political ideas.'

b. ? Fedele ai suoi amici, (lo) è più che alle sue idee
faithful to.the his friends it is more than to.the his ideas
politiche. (Italian)
political
'Faithful to his friends, he is, more than to his political ideas.'

c. Fiel a sus amigos, lo es más que a sus convicciones políticas.
faithful to his friends it is more than to his convictions political
(Spanish)
'Faithful to his friends, he is, more than to his political ideas'

Crucially, the construction differs from that of restructuring verbs in that the dislocated constituent can leave behind its complements (31).

(31) a. Fidèle, il l'est plus à ses amis qu'à ses convictions
faithful he it.is more to his friends than.to his convictions
politiques. (French)
political
'As for being faithful, he is to his friends more than to his political convictions.'

b. Fedele, lo è ai sui amici più che alle sue idee politiche.
faithful it is to.the his friends more than to.the his ideas political
(Italian)
'As for being faithful, he is to his friends more than to his political convictions.'

c. Fiel, lo es más a sus amigos que a sus convicciones políticas.
faithful it is more to his friends than to his convictions political
(Spanish)
'As for being faithful, he is to his friends more than to his political convictions.'

Similarly, the predicative complement can be extracted with its complements or leave them behind. In the latter case, it can be cliticized, as shown in (32c) (compare with examples (16) and (17) with restructuring verbs). In (32), the adjective is extracted (it corresponds to the predicative complement of *être* 'to be', as part of a concessive adjunct (examples (32) and (33) from Abeillé & Godard 2010: 146, 148).

(32) [Context] Is he really faithful to his friends?

a. Aussi fidèle à ses amis qu'il soit, il ne perd pas de vue ses
as faithful to his friends as he is he NE lose not of sight his
intérêts. (French)
interests
'As faithful to his friends as he is, he does not lose sight of his interests.'

b. Aussi fidèle qu'il soit à ses amis, il ne perd pas de vue ses
as faithful as he is to his friends he NE lose not of sight his
intérêts.
interests
'As faithful as he is to his friends, he does not lose sight of his interests.'

c. Aussi fidèle qu'il leur soit, il ne perd pas de vue ses
as faithful as he to.them is he NE lose not of sight his
intérêts.
interests
'As faithful to them as he is, he does not lose sight of his interests.'

Moreover, an adverb may intervene between the copula and the adjective, not only in French or Italian, where it is expected (it is possible with tense auxiliaries and restructuring verbs) but also in Spanish, where it is not expected, if the structure is the same as with restructuring verbs. We illustrate this possibility with cliticization, in order to make the contrast with restructuring verbs clearer.

(33) a. Roméo lui sera probablement fidèle. (French)
Roméo to.him/her will.be probably faithful
'Roméo will probably be faithful to him/her.'

b. Romeo le sarà probabilmente fedele. (Italian)
Romeo to.her will.be probably faithful
'Romeo will probably be faithful to her.'

c. Romeo le será probablemente fiel. (Spanish)
Romeo to.him/her will.be probably faithful
'Romeo will probably be faithful to him/her.'

The data show that, contrary to restructuring verbs, the copula, in Romance languages, has only one complementation. Abeillé & Godard (2002; 2010) propose that the copula takes a "phrasal" complement, which can be saturated or

not. This analysis is implemented by saying that the predicative complement is non-light, whether it is saturated or not, and that it is underspecified with respect to complement saturation or attraction.

(34) Description of the copula in Romance languages

$$\left[\text{ARG-ST}\left\langle \boxed{1}, \left[\begin{array}{l}\text{HEAD [PRD +]}\\ \text{SUBJ } \boxed{1}\\ \text{COMPS } \boxed{2}\\ \text{LIGHT } -\end{array}\right]\right\rangle \oplus \boxed{2}\right]$$

Like tense auxiliaries, the copula is a subject raising verb, hence the identical value 1 for its subject and that of the complement, which allows it to be empty. Its complement differs from that of a tense auxiliary (8b) on several accounts: it is predicative, which is not the case for tense auxiliaries, and it is non-light; in addition, it is not specified for its category. Being non-light, it can have combined with its complements or some of them, while the complement of the auxiliary is light, hence all its complements are attracted.[12]

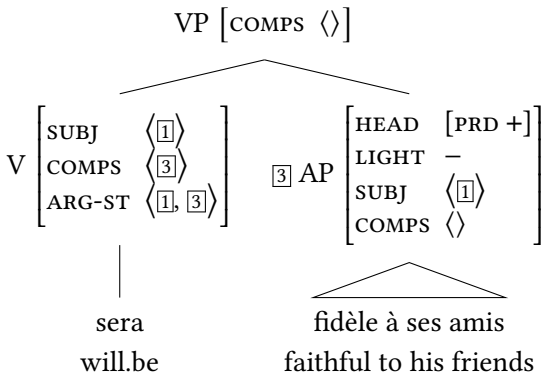


Figure 8: The Romance copula with a saturated complement

Figure 9 illustrates a case where the affix complement of the adjective is attracted to the copula. For cliticization and the notion of reduced verb, see Section 2.

[12] Note that the complements included in a predicative PP are not attracted to the copula in a general way.

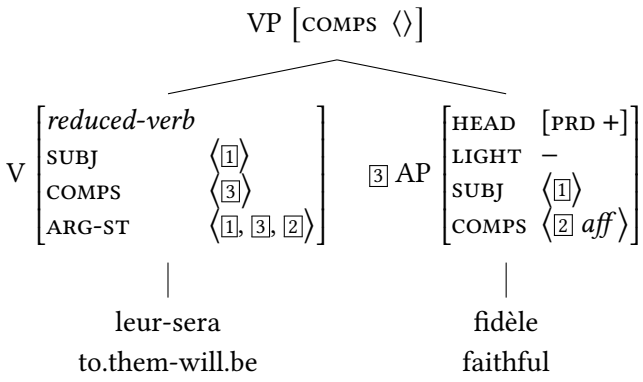


Figure 9: Clitic climbing with the Romance copula

4 Complex predicates and word order

In certain languages, a complex verb construction signals itself essentially by properties of word order. This is the case for instance in German (Hinrichs & Nakazawa 1989; 1994; Kiss 1994; 1995; Hinrichs & Nakazawa 1998; Kathol 1998; Hinrichs & Nakazawa 1999; Kathol 2000; Meurers 2000; 2001; De Kuthy & Meurers 2001; Müller 2002; 2003; 2012) and Dutch (Rentier 1994; Bouma & van Noord 1998), as well as Korean (Sells 1991; Chung 1998; Yoo 2003; Kim 2016). We concentrate on coherent constructions in German, and on Korean auxiliaries.

4.1 Verbal complexes in German

4.1.1 Coherent and incoherent constructions in German

Among verbs with an infinitival complement, German distinguishes between coherent and incoherent constructions (Bech 1955). We speak of constructions rather than verbs, because, although the constructions are triggered by lexical properties of verbs, many verbs can be constructed either way. Verbs entering coherent constructions, obligatorily or optionally, belong to different classes: tense auxiliaries (the verbal complement is an infinitive or a participle), modals, subject and object raising, subject and object control verbs, the copulas and predicative constructions, particle verbs (see Müller 2002, Chapters 2 and 6).

Coherent and incoherent constructions differ with respect to several properties (separability of the head verb and the infinitive, extraposition of the infinitive with its complements, pied-piping in relative clauses, scope of adjuncts). In incoherent constructions, an adverb such as *nicht* 'not' may occur between the two

verbs (35a) (from Müller 2002: 42), the infinitival phrase can be extraposed (35b, 35c), and the infinitive is pied-piped with its relative pronoun complement (35d) (examples from Hinrichs & Nakazawa 1998: 117–118).

(35) a. … dass Karl zu schlafen nicht versucht. (German)
that Karl to sleep not tries
'that Karl does not try to sleep'

b. … dass Peter Maria das Auto zu kaufen überredet.
that Peter Maria the car to buy persuades
'that Peter persuades Maria to buy the car'

c. dass Peter Maria überredet, [das Auto zu kaufen].
that Peter Maria persuades the car to buy
'that Peter persuades Maria to buy the car'

d. Das ist das Auto, das zu kaufen er Peter überreden wird.
that is the car which to buy he Peter persuade will
'That is the car, which he will persuade Peter to buy.'

On the other hand, coherent constructions, of which the combination of the future auxiliary *wird* 'will' (36a) or the raising verb *scheinen* 'to seem') with an infinitival complement (36d) are typical examples, do not allow for a non-verbal element between the two verbs (36b), nor for extraposition of the infinitive with its complements (36c), (36e) (examples (36a), (36c), (36d) et (36e) from Müller 2002: 43), or pied-piping of the infinitive in relative clauses (36f)–(36g) (examples adapted from Hinrichs & Nakazawa 1999: 66).[13]

(36) a. … dass Karl das Buch lesen *wird*. (German)
that Karl the book read will
'that Karl will read the book'

b. * … dass Karl das Buch lesen nicht wird.
that Karl the book read not will
Intended: 'that Karl will not read the book'

c. * … dass Karl wird das Buch lesen.
that Karl will the book read
Intended: 'that Karl will read the book'

[13]The head verb in coherent constructions is italicized.

d. … weil Karl das Buch zu lesen *scheint.*
because Karl the book to read seems
'because Karl seems to read the book'

e. * … weil Karl scheint das Buch zu lesen.
because Karl seems the book to read
Intended: 'because Karl seems to read the book'

f. * Das ist das Buch das lesen Karl wird.
this is the book that read Karl will
Intended: 'This is the book that Karl will read.'

g. * Das ist das Buch das zu lesen Karl scheint.
this is the book that to read Karl seems
Intended: 'This is the book that Karl seems to read.'

Scrambling of the complements of the two verbs, or of the subject of the head verb with the complements of the infinitival is possible in a coherent construction. In (37a) the complements of *sehen* 'see' (*Peter*) and of *kaufen* 'buy' (*das Auto* 'the car') are not interleaved. In (37b), *Peter*, the complement of *sehen*, occurs between *das Auto*, which is the complement of *kaufen*, and *kaufen* (example (37b) from Hinrichs & Nakazawa 1998: 117).

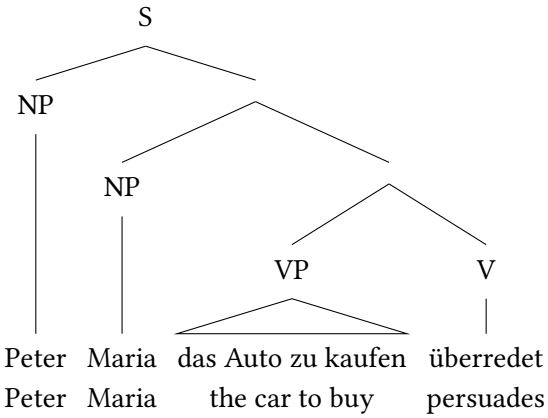
(37) a. … dass er Peter das Auto kaufen *sehen wird.* (German)
that he Peter the car buy see will
'that he will see Peter buy the car'

b. … dass er das Auto Peter kaufen *sehen wird.*
that he the car Peter buy see will
'that he will see Peter buy the car'

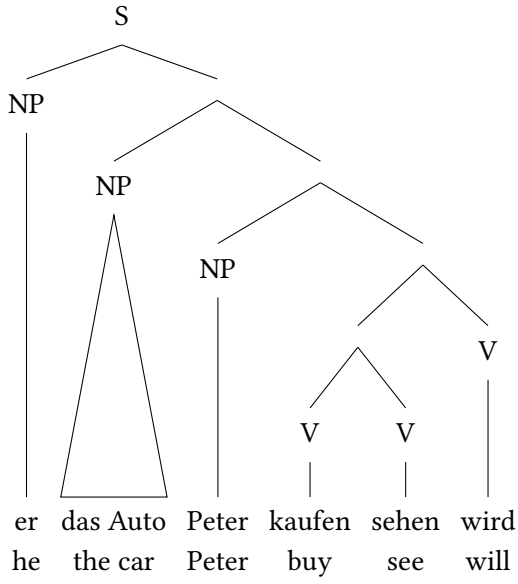
In the complex predicate approach of this chapter, these data point to the following analysis: incoherent constructions involve a saturated VP complement, while coherent constructions do not: they involve a complex predicate, with a verb attracting the complements of its complement. We assume here a verbal complex for the complex predicate. Figure 10a represents example (35b), and Figure 10b represents example (37b).

4.1.2 Coherent constructions in HPSG

One might wonder whether it is possible to analyze the data in terms of word order instead of structure: a verb governing a coherent construction would trig-



(a) Incoherent construction (embedded clause)



(b) Coherent construction (embedded clause)

Figure 10: Incoherent and coherent constructions in German

ger a modification of the ordering domain. More precisely, it would induce domain union of the two ordering domains associated with the two verbal projections (see Müller 2020, Chapter 10 of this volume for a discussion of order domains). Usually, the domain in which constituents are ordered is identical with the phrase or the sentence which dominates them. In the linearization approach (Reape 1994), dominance and ordering can be distinguished. In certain circumstances, the domain for ordering is larger than the domain of constituency, so that the elements belonging to different phrases can be reordered and interleaved, a phenomenon called domain union. Domain union could be responsible for the order in (37b): the structure would be the same as in incoherent constructions (see Figure 10a), but the ordering domain would be the whole sentence.

The existence of the remote (or long) passive goes against such an analysis (Hinrichs & Nakazawa 1994; Kathol 1998; Müller 2002). A complex predicate construction can be passivized in such a way that the subject (in the nominative case) of the passive auxiliary corresponds to the object of the active infinitive complement. An (impersonal) passive construction like (38a) with an infinitival VP containing an accusative object (*den Wagen* 'the car') alternates with a coherent construction such as (38b), with a corresponding nominative (examples (38a)–(38b) from Müller 2002: 137, (38c)–(38d) from Müller 2003: 40).

(38) a. … weil oft versucht wurde, [den Wagen zu reparieren].
because often tried was the car to repair
'because many attempts were made to repair the car'

b. … weil der Wagen oft zu reparieren *versucht wurde*.
because the car often to repair tried was
'because many attempts were made to repair the car'

c. Karl darf nicht versuchen zu schlafen.
Karl is.allowed not try to sleep
'Karl is not allowed to try to sleep.'
'Karl is allowed to not try to sleep.'

d. Karl darf versuchen, nicht zu schlafen.
Karl is.allowed try not to sleep
'Karl is allowed to try not to sleep.'

In (38a), the infinitival VP is extraposed. In (38b), there is no infinitival VP, as shown by the position of the adverb *oft* 'often', which occurs before *zu reparieren* 'to repair', while modifying *versucht* 'tried'. In a coherent construction, an adverb can scope over any of the verbs that belong to it. In (38c), *zu schlafen*

'to sleep' is not part of the coherent construction, because it is extraposed; *nicht* 'not' can have scope over *darf* 'is allowed' or *versuchen* 'to try', not over *schlafen* 'to sleep'. In (38d), *nicht* belongs to the extraposed infinitival; accordingly, it can only scope over it. The fact that *oft* can scope over *versucht* in (38b) shows that they belong to the same coherent construction, which allows for passivization: *versuchen* attracting the complement of *reparieren* can be passivized.

German differs from Romance languages in not distinguishing structurally between the subject and the complements: the subject can be considered as a complement, and introduced by the same rule. The structure of the sentence is usually represented as having binary branching daughters (see Figures 10). The constraint is as follows (Müller 2018: 21).

(39) *head-complement-phrase* (German) ⇒

$$\begin{bmatrix} \textsc{synsem} & \begin{bmatrix} \textsc{loc} \mid \textsc{cat} \mid \textsc{comps} \boxed{1} \oplus \boxed{3} \\ \textsc{light} - \end{bmatrix} \\ \textsc{head-dtr} \mid \textsc{synsem} & \begin{bmatrix} \textsc{loc} \mid \textsc{cat} \mid \textsc{comps} \boxed{1} \oplus \left\langle \boxed{2} \right\rangle \oplus \boxed{3} \end{bmatrix} \\ \textsc{non-head-dtrs} & \left\langle [\textsc{synsem} \boxed{2}] \right\rangle \end{bmatrix}$$

Following constraint (39), the head combines with one complement at a time, noted $\boxed{2}$. The presentation of the list as composed of three parts, with the relevant one in any position, allows for a free order. The lexical verb is [LIGHT +], and the phrase combining the verb with a complement is [LIGHT-].[14] The structure of (40) is exemplified in Figure 11 (Müller 2018: 22).

(40) ... weil das Buch jeder kennt
because the book everybody knows
'because everybody knows the book.'

Turning to complex predicates, they form a verbal complex phrase: they cannot be separated by an adverb or an NP (36b)–(36c). Given the structure of the German sentence with binary branching, illustrated in Figure 10, this verbal complex only shows structurally when there is a series of verbs attracting the complements of their complements, as in (37) (see Figure 10b).

[14] The feature LIGHT is the equivalent of LEX used in German studies, although the properties of light elements may differ depending on the language. It does not belong to LOCAL features in (39), because an extracted constituent may differ from its trace as regards lightness (Müller 2018) (see Borsley & Crysmann (2020), Chapter 13 of this volume for extraction).

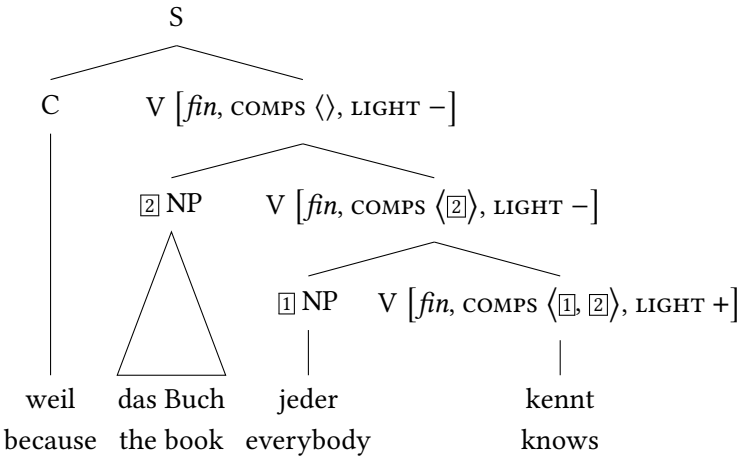


Figure 11: Clause structure in German

The phrase structure constraint allowing complex predicates is as in (41) (Müller 2012; Müller 2018: 39). It is called *head-cluster-phrase*, rather than *verbal-complex-phrase*, because it is not specialized for verbal heads.[15]

(41) *head-cluster-phrase* (German) ⇒

$$\left[\begin{array}{ll} \textsc{synsem} & \left[\begin{array}{l}\textsc{loc}\mid\textsc{cat}\mid\textsc{comps}\ \boxed{1}\\ \textsc{light}\ +\end{array}\right] \\ \textsc{head-dtr}\mid\textsc{synsem} & \left[\begin{array}{l}\textsc{loc}\mid\textsc{cat}\mid\textsc{comps}\ \boxed{1}\oplus\left\langle\boxed{2}\right\rangle\\ \textsc{light}\ +\end{array}\right] \\ \textsc{non-head-dtrs} & \left\langle[\textsc{synsem}\ \boxed{2}\ [\textsc{light}\ +]]\right\rangle \end{array}\right]$$

We illustrate the analysis with sentence (37b) (… *dass er das Auto Peter kaufen sehen wird* 'that he will see Peter buy the car'), elaborating on Figure 10b. The description of *werden* (the future auxiliary), a subject raising verb, is as in (42) (from Müller 2018: 39), and that of *sehen* 'to see', an object raising verb and an obligatorily coherent verb, is as in (43) (simplified from Müller 2002: 102). The

[15] Following Hinrichs & Nakazawa (1994) and De Kuthy & Meurers (2001), but contrary to Müller (2018), we mention the lightness of the mother. Müller's decision is motivated by the fact that infinitive intransitive verbs may be analyzed as argument saturated (and non-light) if their subject is represented as a head feature rather than a complement. However, it leads to formal complications which are best ignored in this presentation. Hence, we assume here, for the sake of simplification, that the subject of the infinitival verb is a complement in German, like the subject of a finite verb.

subject is assumed here (for simplification) to be part of the list of complements of infinitives as well as of finite verbs; hence the raised subject of the infinitive complement of *werden* 'will' is included in list 1, and that of *sehen* is included in list 2. The subject is distinguished from the other elements of this list by its semantic role (it is the first semantic argument of the infinitive).

(42) *werden* (future auxiliary):

$$\begin{bmatrix} \textsc{head}\ \textit{verb} \\ \textsc{comps}\ \boxed{1} \oplus \left\langle \textsc{v}\ [\textit{bse}, \textsc{comps}\ \boxed{1}, \textsc{light}\ +] \right\rangle \end{bmatrix}$$

(43) *sehen* (obligatory coherent verb):

$$\begin{bmatrix} \textsc{head}\ \textit{verb} \\ \textsc{comps}\ \boxed{1} \oplus \boxed{2} \oplus \left\langle \textsc{v}\ [\textit{bse}, \textsc{comps}\ \boxed{2}, \textsc{light}\ +] \right\rangle \end{bmatrix}$$

Sentence (37b) is represented in Figure 12.

Stefan: I changed this figure. It is now CP and VFORM is under head.

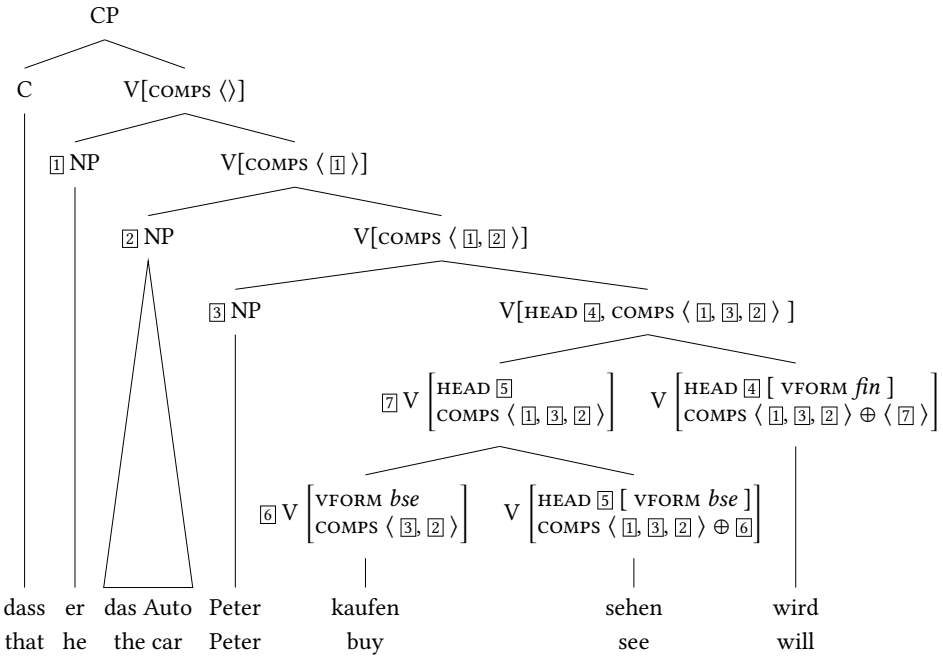


Figure 12: Coherent construction with verbal complexes in German

4.1.3 The German copula

The copula in German, with an adjectival argument, is also the head of a complex predicate.[16] The subject of the copula and the complements of the adjectives can be permuted (examples from Müller 2002: 68) (see (37) for coherent verbs):

(44) a. … dass die Sache dem Minister ganz klar war.
that the matter.NOM the minister.DAT completely clear was
'that the matter was completely clear to the minister.'

b. … dass dem Minister die Sache ganz klar war.
that the minister.DAT the matter.NOM completely clear was
'that the matter was completely clear to the minister.'

Adverbs can have different scopings: in (45) (from Müller 2002: 68), *immer* 'always' can modify the modal or the adjective. This follows if there is just one coherent construction, both the modal and the copula being the head of a complex predicate (see Section 4.1.2, example (38b) for coherent verbs).

(45) … weil der Mann ihr immer treu sein wollte
because the man.NOM her.DAT always faithful be wanted.to
'because the man always wanted to be faithful to her.'
'because the man wanted to be faithful to her forever.'

Müller (2002) also shows that the copula does not take a saturated AP complement. Contrary to a construction with an incoherent verb, this AP cannot be extraposed (46b), or pied piped with a relative pronoun (46d) (from Müller 2002: 70) (compare with (35c), (35d)).

(46) a. Karl ist auf seinem Sohn stolz gewesen.
Karl is on his son proud been
'Karl was proud of his son.'

b. * Karl ist gewesen auf seinem Sohn stolz.
Karl is been on his son proud
Intended: 'Karl was proud of his son.'

c. der Sohn, auf den Karl stolz gewesen ist
the son on whom Karl proud been is
'the son of whom Karl was proud'

[16] As in Romance languages, the German copula accepts nominal and prepositional predicative complements. However, they are complement saturated.

d. * der Sohn, auf den stolz Karl gewesen ist
the son on whom proud Karl been is
Intended: 'the son of whom Karl was proud'

In addition, the German copula, like the Romance copula, is a subject raising verb: the semantic properties of the subject depend on the adjective (a human is proud or faithful, and a matter is clear, cf. Karl's faithfulness, the clarity of the matter); moreover, the sentence can be subjectless (from Müller 2002: 72):

(47) Am Montag ist schulfrei.
at.the Monday is school.free
'There is no school on Monday.'

The description of the German copula, restricted to its predicative use, and to its syntactic part, is as follows:

Stefan: Where is this from? Provide reference. The German literature never used a type *sein*. I would just omit it and say that it is *sein* in the heading.

(48)
$$\begin{bmatrix} \textit{sein} \\ \textsc{head}\ \textit{verb} \\ \textsc{comps}\ \boxed{1} \oplus \left\langle \begin{bmatrix} \textsc{head}\ [\textsc{prd}\ +] \\ \textsc{comps}\ \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix}$$

It differs from the Romance copula in not specifying the lightness of its predicative complement. The COMPS list includes the subject, while subject and complements are distinguished in Romance languages.

4.2 Argument attraction with Korean auxiliaries

4.2.1 Scrambling with auxiliaries in Korean

Korean resembles German in that a complex predicate signals itself notably by its word order properties (see Sells 1991; Chung 1998; Yoo 2003; Kim 2016). We illustrate here the case of auxiliaries.[17]

Korean auxiliaries semantically resemble aspectual or modal verbs rather than tense auxiliaries: they include such verbs as *iss-* 'to be in the process/state of', *chiwu-* 'to do resolutely', *siph-* 'to want', but also the verb of negation *anh-* 'not' (see also Kim 2020, Chapter 18 of this volume). They bear the tense marking

[17] Chung (1998) also considers control verbs to be the head of complex predicates, and Kim's study (2016) includes serial verb and light verb constructions.

for the sentence (49a), impose a certain ending to their verbal complement (*-e* in (49a)), and, when they also have a use as ordinary verbs (49b), they have an argument structure which is absent in their auxiliary use (examples from Kim 2016: 85-86).

(49) a. Mia-ka wul-e pely-ess-ta.
Mia-NOM cry-PART end.up-PST-DECL
'Mia ended up crying.'

b. Mimi-nun congi-lul hyucithong-ey pely-ess-ta.
Mimi-TOP paper-ACC trash.can-LOC throw.away-PST-DECL
'Mimi threw away the paper in the trash can.'

In (49b), the verb has three arguments: agent subject, theme object, and locative complement. This argument structure is absent in (49a).

Consider the sentences (50). There is no evidence of scrambling in (50a): the subject *Maryka* starts the sentence, and the complement of the verb *ilkko* 'read' immediately precedes it. However, in (50b), the subject of the head verb *issta* 'be in the process of' occurs between the complement of *ilkko* and the verb *ilkko* itself.

(50) a. Mary-ka ku chayk-ul ilk-ko iss-ta.
Mary-NOM the book-ACC read-PART be.in.the.process.of-DECL
'Mary is in the process of reading the book.'

b. Ku chayk-ul Mary-ka ilk-ko iss-ta.
the book-ACC Mary-NOM read-PART be.in.the.process.of-DECL
'Mary is in the process of reading the book.'

A priori, these data could be explained in two ways: either the auxiliary always takes a VP complement, and scrambling is due to linearization, in which case the domains of the two verbs are unioned (Reape 1994); or there is a complex predicate: the complement of the embedded verb (*the book*) is attracted by the auxiliary verb.

There are several properties which show that scrambling is due to argument attraction. First, the presence of the auxiliary allows for case alternation: the argument of a verb like *mek-* 'to eat' is assigned accusative case (51a); however, when the verb is the complement of the auxiliary verb *siphessta* 'would like' in (51b), it can be either accusative or nominative (examples (51) from Kim 2016: 87).

(51) a. Mimi-ka sakwa-lul/*ka mek-ess-ta.
Mimi-NOM apple-ACC/NOM eat-PST-DECL
'Mimi ate an apple.'

b. Mimi-ka sakwa-lul/ka mek-ko siph-ess-ta.
Mimi-NOM apple-ACC/NOM eat-PART wish-PST-DECL
'Mimi would like to eat an apple.'

Given that case assignment is a local phenomenon, and a verb does not influence the case of the complement of its complement, this indicates that *sakwa-* 'becomes' the complement of the auxiliary (see also Yoo 2003). Moreover, in Korean, a negative polarity item such as *amwukesto* 'anything' is licensed by a clause-mate negated element (52a) (52c). (52a) and (52b) show that the negative verb *anh-* allows this negative polarity item as the argument of *mekko*, complement of the auxiliary *siphci* 'to wish' (examples from Kim 2016: 91).

(52) a. Mimi-nun amwukes-to mek-ci anh-ass-ta.
Mimi-TOP anything-also eat-PART not-PAST-DECL
'Mimi did not eat anything.'

b. Mimi-nun amwukes-to mek-ko siph-ci anh-ass-ta.
Mimi-TOP anything-also eat-PART wish-PART not-PAST-DECL
'Mimi did not feel like eating anything.'

c. * Mimi-lul amwukes-to mek-tolok seltukha-ci anh-ass-ta.
Mimi-ACC anything-also eat-PART persuade-PART not-PAST-DECL
Intended: '(We) did not persuade Mimi to eat anything.'

Finally, the same argument can be levelled against an analysis which appeals to linearization, as above in German (4.1.2): so-called long passivization is possible with certain auxiliaries like *chiwu* 'to do resolutely', which cannot be accounted for by appeal to linearization and domain union (examples from Chung 1998: 164)[18]. (53a) exemplifies the active sentence, and (53b), the passive one. In (53a), *malssengmanhun solul* 'the troublesome cow' is the complement of the complement verb *phala* 'sell'. In (53b), *malssengmanhun soka* is the subject of the passivized verb *chiwe ciessta*.

(53) a. Ku nongpwu-ka malssengmanhun so-lul phal-a
the farmer-NOM troublesome cow-ACC sell-PART

[18] However, such passives are judged unnatural by native speakers, hence the question mark.

chiw-ess-ta.
do.resultely-PAST-DECL
'The farmer resolutely sold the troublesome cow.'

b. ? Malssengmanhun so-ka (ku nongpwu-eyuyhay) phal-a
Troublesome cow-NOM the farmer-by sell-PART
chiw-e ci-ess-ta.
do.resolutely-PART PASS-PAST-DECL
'The troublesome cow was resolutely sold (by the farmer).'

Since passivization only affects the complement of the verb which is itself passivized, it follows that *malssengmanhun solul* is the complement of the auxiliary in (53a).

The scrambling data with control verbs are very similar to those with auxiliaries (54). There is no scrambling in (54a): the dative complement of the head verb is followed by the other complement, a VP. However, in (54b), the subject of the head verb (*Maryka* 'Mary' + nominative) occurs between the complement of the complement verb (*ku chaykul* 'the book' + accusative) and the dative complement of the head verb (*Johnhanthey* 'John' + dative).

(54) a. Mary-ka John-hanthey [ku chayk-ul ilk-ulako]
Mary-NOM John-DAT the book-ACC read-PART
seltukha-yess-ta.
persuade-PAST-DECL
'Mary persuaded John to read the book.'

b. Ku chayk-ul Mary-ka John-hanthey ilk-ulako
the book-ACC Mary-NOM John-DAT read-PART
seltukha-yess-ta.
persuade-PAST-DECL
'Mary persuaded John to read the book.'

However, we do not observe case alternation, and control verbs fail to allow the negative polarity item *amwukesto* 'anything' as the complement of the verb complement (Kim 2016: 91).

(55) a. Mimi-lul amwukes-to an mek-tolok selkhuta-yess-ta.
Mimi-ACC anything-also no eat-PART persuade-PAST-DECL
'(We) persuaded Mimi not to eat anything.'

b. * Mimi-lul amwukes-to mek-tolok selkhuta-ci anh-ass-ta.
Mimi-ACC anything-also eat-PART persuade-PART not-PAST-DECL
Intended: 'We did not persuade Mimi to eat anything.'

Accordingly, we follow Kim (2016) in not analyzing control verbs as heads of complex predicates. They take VP complements, and scrambling in (53) must be due to a different process (domain union, Reape 1994).

4.2.2 Korean auxiliaries and the verbal complex

It has been shown in this paper that different structures could be associated with argument attraction. Korean auxiliaries are the head of a verbal complex (Chung 1998; Kim 2016). The main fact is that nothing can intervene between the two verbs, for instance no parenthetical expression, such as *hayekan* 'anyway' (56) (examples from Chung 1998: 162). This contrasts with control verbs. In (57), the adverb *cengmal* 'really' can occur before the embedded verb, or between the two verbs (example (57) from Kim 2016: 93).

(56) a. Mary-ka hayekan sakwa-lul mek-ko iss-ta.
Mary-NOM anyway apple-ACC eat-PART be.in.the.process.of-DECL
'Anyway, Mary is eating an apple.'

b. * Mary-ka sakwa-lul mek-ko hayekan iss-ta.
Mary-NOM apple-ACC eat-PART anyway be.in.the.process.of-DECL
Intended: 'Anyway, Mary is eating an apple.'

(57) Mimi-nun Haha-lul (cengmal) ttena-tolok (cengmal)
Mimi-TOP Haha-ACC really leave-PART really
seltukha-yess-ta.
persuade-PST-DECL
'Mimi (really) persuaded Haha to (really) leave.'

In addition, there is evidence that the verb complement of an auxiliary and its complement do not form a constituent. While an NP may occur after the head verb in a construction called afterthought (58a), this is not possible for the embedded verb *mekko* with its complement (58b) (from Chung 1998: 162).

(58) a. Mary-ka mek-ko iss-ta, sakwa-lul.
Mary-NOM eat-PART be.in.the.process.of-DECL apple-ACC
'Mary is in the process of eating an apple.'

b. * Mary-ka iss-ta, sakwa-lul mek-ko.
Mary-NOM be.in.the.process.of-DECL apple-ACC eat-PART
Intended: 'Mary is in the process of eating an apple.'

These data point to a verbal complex (see section 3.2). However, before coming to this conclusion, we must show that the two verbs do not form a compound word. No (1991) (summarized in Chung 1998, Kim 2016) presents two arguments to the effect that they combine in the syntax. First, a delimiter (such as *-man* 'only', *-to* 'also' can combine with the embedded verb (e.g. *mekkoman issta* 'to be only eating'). Delimiters are a syntactic phenomenon, not limited to verbal morphology. Second, compounding cannot account for the productivity of the construction: it may involve several auxiliaries, since there is no limit to the number of auxiliaries which may combine, with the appropriate ending. It is not plausible to list all the possible combinations in the lexicon. (59a), which repeats (50a), exemplifies a sentence with one auxiliary, and (59b) a sequence of two auxiliaries (Chung 1998: 172).

(59) a. Mary-ka ku chayk-ul ilk-ko iss-ta.
Mary-NOM the book-ACC read-PART be.in.the.process.of-DECL
'Mary is in the process of reading the book.'

b. Mary-ka ku chayk-ul ilk-ke po-ko
Mary-NOM the book-ACC read-PART try-PART
iss-ta.
be.in.the.process.of-DECL
'Mary is in the process of giving the book a trial reading.'

Thus, the head auxiliary and the complement verb do not form a compound word, but a verbal complex.

4.2.3 Korean auxiliaries in HPSG

Given the free word order in Korean (except for the verb), there are two ways of representing the sentence: either there is a flat structure (except for the verbal complex), where all the arguments, subject and complements, are sisters of each other (a. o. Chung 1998 for Korean, Hinrichs & Nakazawa 1998 for German), or there is a binary branching structure (Kim 2016 for Korean). We adopt the flat structure here. Nothing crucial hinges on the choice.

The general schema for the sentence is given in (60), adapted from Chung (1998: 178). The sign / indicates that, by default, the value of SUBJ and COMPS is the empty list.

(60) *head-subject-complements-phrase* (Korean) ⇒

$$\begin{bmatrix} \textsc{synsem} \mid \textsc{cat} & \begin{bmatrix} \textsc{subj} / \langle\rangle \\ \textsc{comps} / \langle\rangle \\ \textsc{light} - \end{bmatrix} \\ \textsc{head-dtr} \mid \textsc{synsem} \mid \textsc{cat} & \begin{bmatrix} \textsc{head} \begin{bmatrix} \textit{verb} \\ \textsc{aux} - \end{bmatrix} \\ \textsc{subj} \boxed{1} \\ \textsc{comps} \boxed{2} \\ \textsc{light} + \end{bmatrix} \\ \textsc{non-head-dtrs} & (\boxed{1} \circ \boxed{2}) \text{ nonempty list} \end{bmatrix}$$

The verbal complex is headed by an auxiliary verb, which is [AUX +] while other verbs are [AUX −]. Thus only auxiliaries can enter this structure, and they must, since the head of the head-subject-complements phrase is specified as [AUX −]. The schema for the verbal complex is given in (61). Auxiliaries attract both the subject (1) and the complements of their verbal complement (list 3). The subject value is indicated as 1, rather than ⟨1⟩, because the subject is not always realized in Korean. The verbal complex is [LIGHT +], and made up of two verbs, also [LIGHT +] (see section 3.3).

(61) *Verbal-complex-phrase* (Korean) ⇒

$$\begin{bmatrix} \textsc{synsem} \mid \textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{verb} \\ \textsc{aux} + \end{bmatrix} \\ \textsc{subj} & \boxed{1} \\ \textsc{comps} & \boxed{3} \\ \textsc{light} + \end{bmatrix} \\ \textsc{head-dtr} \mid \textsc{synsem} \mid \textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{verb} \\ \textsc{aux} + \end{bmatrix} \\ \textsc{subj} & \boxed{1} \\ \textsc{comps} & \left\langle \boxed{2} \right\rangle \oplus \boxed{3} \\ \textsc{light} + \end{bmatrix} \\ \textsc{non-head-dtrs} & \left\langle \boxed{2} \begin{bmatrix} \textit{verb} \\ \textsc{subj} & \boxed{1} \\ \textsc{comps} & \boxed{3} \\ \textsc{light} + \end{bmatrix} \right\rangle \end{bmatrix}$$

The description of an auxiliary such as *issta* in (59a) is given in (62). To indicate which ending it imposes on its complement, we use the feature VFORM, thus allowing for the selection of the appropriate ending by the auxiliary (Chung 1998, Kim 2016). So, the verb *issta* 'be in the process of' selects the ending *–ko* for the verbal complement, and *ilkko* 'read', whose VFORM value is *ko*, is appropriate. The auxiliary attracts the whole ARG-ST of its verbal complement, that is the subject [1], and the complements [2] (see (9)).

(62) Description of *issta*

$$\begin{bmatrix} \textsc{form} & \left\langle \text{iss-ta} \right\rangle \\ \textsc{head} & \begin{bmatrix} \textit{verb} \\ \textsc{aux} + \end{bmatrix} \\ \textsc{arg-st} & \boxed{1} \oplus \left\langle \textsc{v} \begin{bmatrix} \textsc{vform} & \textit{ko} \\ \textsc{arg-st} & \boxed{1} \oplus \boxed{2} \\ \textsc{light} + \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix}$$

The structure of sentence (59a) is represented in Figure 13.

The structure of (59b), with a series of two auxiliaries, is represented in Figure 14 (adapted from Chung 1998: 171). The verb *issta* 'be in the process of' takes as its complement the verbal complex *ilke poko* 'try to read', whose head is *poko* 'try'. The verb *poko*, being an auxiliary like *issta*, takes as its complement the verb *ilke*, attracting its subject and complements, which are transmitted to the verbal

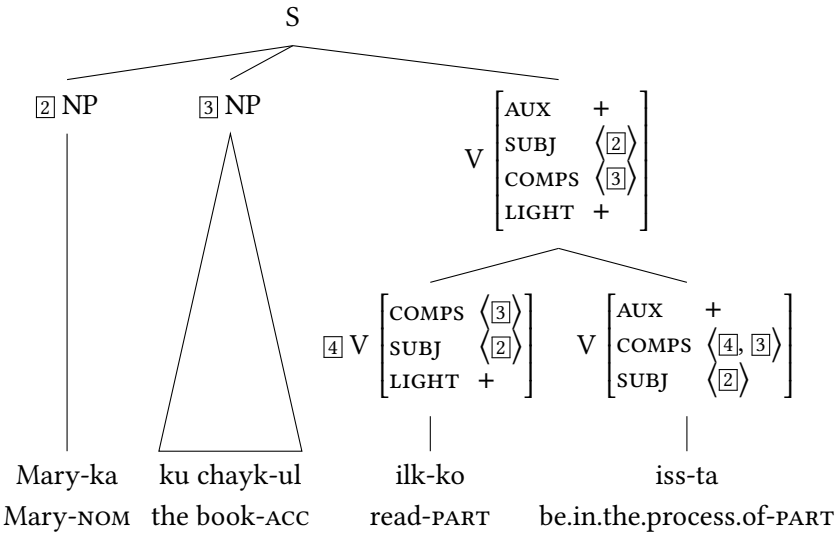


Figure 13: Clause structure with a verbal complex in Korean

complex *ilke poko*; *ilke poko* saturates the verbal complement expected by *issta*, and transmits the subject and complements to the head auxiliary (see (62)).

The head comes last in Korean, except in the afterthought construction (58), which requires an additional mechanism. Constraint (63) mirrors constraint (28) for Romance languages.

(63) $\left[\text{SYNSEM } \boxed{1}\right] < \left[\text{COMPS } \langle \ldots, \boxed{1}, \ldots \rangle\right]$ (Korean)

This constraint holds for the verbal complex, in which the head verb follows the complement verb.

5 Light verb constructions in Persian: syntax and morphology, syntax and semantics

Light verb constructions constitute the third guise of complex predicates. They are characterized semantically: the verb and the second predicate constitute together a semantic predicate. For instance, the French expression combining a semantically light verb and a noun *faire une proposition* 'to make a proposal' is close to *proposer* 'to propose'. They have been studied in HPSG for Korean (Ryu

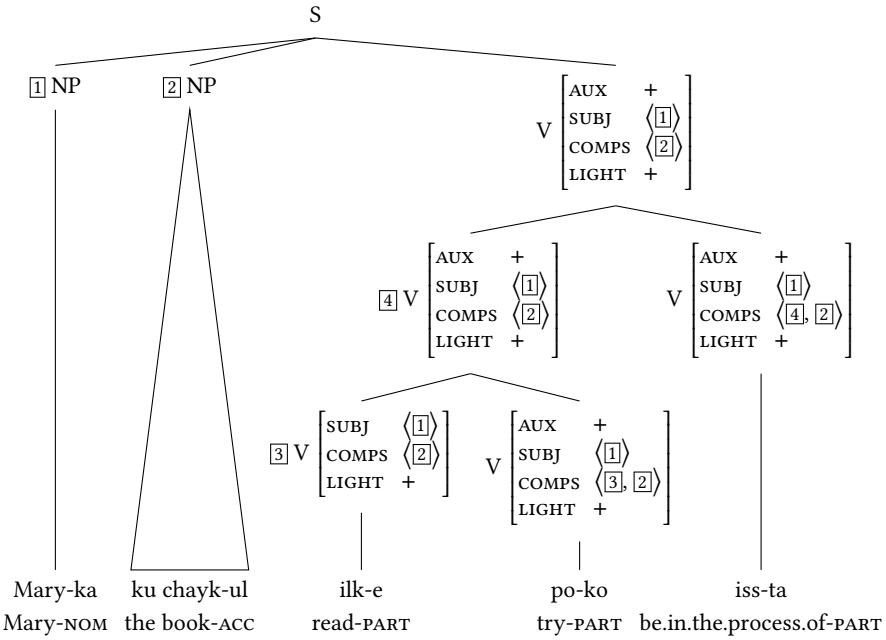


Figure 14: Clause structure with verbal complexes in Korean.

1993; Lee 2001; Choi & Wechsler 2002; Kim 2016). We focus here on Persian light verb constructions, which form a rich class and tend to replace simplex verbs.

5.1 What are complex predicates in Persian?

Persian verbs (simplex verbs) constitute a small closed class of about 250 members, only around 100 of which are commonly used. Speakers resort to complex predicates, sequences of a light verb and a preverbal element, belonging to different categories (adjective, noun, particule, prepositional phrase). Following Bonami & Samvelian (2010), Samvelian (2012), such sequences are 'multi-word expressions', that is, they are made up of several words, which, together, form a lexeme.

Several properties show that the elements are independent syntactic units (Karimi-Doostan 1997; Megerdoomian 2002; Samvelian 2012). We concentrate on noun + verb combinations. All inflection is prefixed or suffixed on the verb, as is the negation in (64), and never on the noun. The two elements can be separated by the future auxiliary, or even by clearly syntactic constituents, like the complement PP in (64). Both the noun and the verb can be coordinated, as shown in

(65) and (66) respectively (from Bonami & Samvelian 2010: 3). The noun can be extracted, as in the topicalization in (67), where the sign – indicates where the non extracted noun would have occurred. Complex predicates can be passivized. In this case, the nominal element of the complex predicate (*tohmat* 'slander' in (68a)) becomes the subject of the passive construction (68b), as does the object of a transitive construction (from Samvelian 2012: 251). The nominal part of the complex predicate is italicized in the examples.

(64) *Dast* be gol-hā na-zan.
hand to flower-PL NEG-hit
'Don't touch the flowers.'

(65) Mu-hā=yāš=rā [*boros yā šāne*] zad.
hair-PL=3SG=RA brush or comb hit
'He/she brushed or combed his/her hair.'

(66) Omid *sili* [zad va xord].
Omid slap hit and stroke
'Omid gave and received slaps.'

(67) *Dast* goft=am [be gol-hā – na-zan].
hand said=1SG to flower-PL NEG-hit
'I told you not to touch the flowers.'

There is a mistake in (68b) I do not know how to fix.

(68) a. Maryam be Omid *tohmat* zad.
Maryam to Omid slander hit
'Maryam slandered Omid.'
b. Be Omid *tohmat* zade šod.
to Omid slander hit.PST-PTCP become
'Omid was slandered.'

There is evidence that the verb and the noun in a complex predicate share one argument structure. In (69a), the verb *dādan* takes two complements, the noun *āb* and the PP *[be bāqče]*, while in (69b) the combination of *dādan* and the noun *āb* takes a direct objet, which is marked with *-rā*: in (69b), the noun *āb* and the verb *dād* form a complex predicate.

(69) a. Maryam [be bāqče] āb dād.
Maryam to garden water gave
'Maryam watered the garden.'

b. Maryam [bāqče=rā] *āb* dād.
Maryam garden=RA water gave
'Maryam watered the garden.'

Other properties show that the combination of the two elements, here a noun and a verb, behaves like a lexeme (Bonami & Samvelian 2010). Such combinations feed lexeme formation rules: for instance, the suffix *–i* forms adjectives from verbs: *xordan* 'eat' > *xordani* 'edible', and the same is possible with complex predicates (70); perfect participles can be converted into adjectives, and this applies to complex predicates (71) (see also Section 5.2) (from Bonami & Samvelian 2010: 5).

(70) a. dust dāštan > dustdāštani
friend have lovely
'Love.'

b. xat xordan > xatxordani
scratch strike scratchable
'Be scratched.'

(71) a. dast xordan > dastxorde
hand strike sullied
'To be sullied.'

b. xat xordan > xatxorde
scratch strike scratched
'Be scratched.'

The meaning of the complex predicate is often a specialization of the predictable meaning of the combination: *dast dādan* (lit. 'hand give') means 'shake hands', *čāqu zadan* (lit. 'knife hit') means 'stab', *šāne zadan* (lit. 'comb hit') means 'comb'. Each specialized meaning has to be learned in the same way as that of a lexeme. Analogy often plays a role in the creation of new lexemes, and this is also true of complex predicates. For instance, the family of complex predicates expressing manners of communication goes from *telegrām zadan* 'telegraph' where hitting *(zadan)* is involved to cases where hitting is not clearly involved: *telefon zadan* 'phone', *imeyl zadan* 'email', *esemes zadan* 'text' etc.

These complex predicates raise two problems, a morpho-syntactic one and a semantic one. To solve them, we rely crucially on the same property of HPSG as in the preceding syntactic cases, that is, the view of heads as sharing information with their expected complements.

5.2 Complex predicates and derivational processes

Although Persian complex predicates are combinations of words, they feed some derivational rules; see Section 5.1, examples (70), (71). We analyze here the case of agent nominalization, studied in Müller (2010).[19] The example he examines is especially interesting in that the nominalization does not exist independently of the complex predicate: as shown in (72), although no agent noun can be derived from the verb *kon* 'do', an agent noun can be derived from the complex predicate formed with the verb *kon* and the adjective *bāz* 'open'.

(72) a. kon 'do', * kon-ande 'do-er'
b. bāz kon (open + do 'open'), bāz-konande 'opener'

The lexeme *bāz-konande* 'opener' can be analyzed as a compound lexeme to which the suffix *–ande* is added, indicating agent nominalization. Compound lexemes are made of two lexemes. A simple rule for noun-noun compounds is given in (73) (Bonami & Crysmann 2018: 178), where the elements of the compound are the value of the feature M-DTRS (morphological daughters):[20]

(73)
$$\left[\begin{array}{l} \textit{lexeme} \\ \textsc{phon}\ \boxed{1} \oplus \boxed{2} \\ \textsc{synsem} \mid \textsc{loc} \mid \textsc{cat} \mid \textsc{head}\ \textit{noun} \\ \textsc{m-dtrs} \left\langle \left[\begin{array}{l} \textit{lexeme} \\ \textsc{phon}\ \boxed{1} \\ \textsc{synsem} \mid \textsc{loc} \mid \textsc{cat} \mid \textsc{head}\ \textit{noun} \end{array}\right], \left[\begin{array}{l} \textit{lexeme} \\ \textsc{phon}\ \boxed{2} \\ \textsc{synsem} \mid \textsc{loc} \mid \textsc{cat} \mid \textsc{head}\ \textit{noun} \end{array}\right] \right\rangle \end{array}\right]$$

Similarly, a noun can be formed from the elements of the complex predicate, in this instance an adjective and a verb. The verb *kon* in the complex predicate *bāz kon* 'to open' is described in (74). It expects a subject NP, the agent, and two complements, an adjective and an NP, the latter being attracted from the adjective. The content of the adjective is included in the content of the verb, as the nucleus of the caused *soa* 'make something be adj' (see Müller 2010: 642).

[19] Müller's analysis is couched in a slightly different framework.
[20] For a similar approach to Morphology in HPSG, see Orgun (1996), Riehemann (1998), Koenig (1999), Sag et al. (2003).

(74)
$$\begin{bmatrix} \textsc{cat}|\textsc{head}\ \textit{verb} & \\ \textsc{arg-st} & \left\langle \mathrm{NP}_k, \mathrm{NP}_j, \textsc{Adj} \begin{bmatrix} \textsc{prd} & + \\ \textsc{arg-st} & \langle \mathrm{NP}_j \rangle \\ \textsc{cont} & \boxed{1} \begin{bmatrix} \textit{open-relation} \\ \textsc{theme}\ j \end{bmatrix} \end{bmatrix} \right\rangle \\ \textsc{cont} & \begin{bmatrix} \textit{soa} & \\ \textsc{nucleus} & \begin{bmatrix} \textit{cause-relation} \\ \textsc{causer}\ k \\ \textsc{soa-arg}|\textsc{nucleus}\ \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The compound lexeme *bāz-konande* is made of the adjective and the verb, which are the morphological daughters, very similar to what they are in the complex predicate. The verbal element is expecting two complements, an adjective and an NP, and they have the same semantics as in the complex predicate: the verb denotes a cause relation taking as argument the adjective content, and the adjective content is a relation taking the nominal complement as its argument (SS abbreviates SYNSEM).

(75)
$$\begin{bmatrix} \textit{lexeme} & \\ \textsc{phon}\ \boxed{1} \oplus \boxed{2} \oplus \langle \textit{ande} \rangle & \\ \textsc{synsem} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head}\ \textit{noun} \\ \textsc{comps} \langle \boxed{3}\ \mathrm{NP} \rangle \end{bmatrix} \\ \textsc{cont}\ [\textsc{ind}\ k] \end{bmatrix} \\ \textsc{m-dtrs} & \left\langle \begin{bmatrix} \textsc{phon}\ \boxed{1} & \\ \textsc{ss}\ \boxed{4} & \begin{bmatrix} \textsc{cat}\ \textit{adjective} \\ \textsc{arg-st} \langle \boxed{3}\ \mathrm{NP}_j \rangle \\ \textsc{cont}\ \boxed{5}\ \textit{adj-rel (j)} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \textsc{phon}\ \boxed{2} & \\ \textsc{ss} & \begin{bmatrix} \textsc{cat}\ \textit{verb} & \\ \textsc{arg-st} \langle \mathrm{NP}_k, \boxed{3}, \boxed{4} \rangle & \\ \textsc{cont} \mid \textsc{nucl} & \begin{bmatrix} \textit{cause-rel} \\ \textsc{causer}\ k \\ \textsc{soa-arg} \mid \textsc{nucl}\ \boxed{5} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

The agent noun keeps as a complement the NP expected by the verb, as illustrated in (76).

(76) [dar-e botri] bāz-konande
lid-EZ bottle opener
'a bottle opener'

This compound nominalization is accompanied with the appropriate changes: the noun denotes the causer, the first argument of the verb m-daughter, and a

derivational suffix *(-ande)* is appended to the sequence of the two elements. Nothing in the rule requires that the agent noun (**kon-ande*) exist independently of the elements of the complex predicate. Hence, the data in (72a) are accounted for.

5.3 The semantics of light verb constructions

The nouns found in complex predicates are used either as referential, or as part of a complex predicate, where they are analyzed as predicative complements, participating in its semantics (with the feature [PRD +]). We assume that such nouns come in two guises, as [PRD +] and as [PRD −].

These complex predicates do not have a homogeneous semantics. The general idea is that the verb serves to turn a noun into a verb (Bonami & Samvelian 2010), but there is a spectrum, going from a (relatively) semantically compositional combination, to idioms whose semantics is not predictable from the components. Complex predicates exploit different schemas, which can be extended to new nouns, describing new situations. We will exemplify certain common cases, drawing on the detailed study of *zadan* 'to hit' in Samvelian (2012). The uses of *zadan* as a light verb are numerous and varied. We will not try to investigate them exhaustively; rather, we indicate different patterns for the combination of this verb with the noun.

The semantics of a complex predicate is often a specialization of that of the simplex verb. For instance, *lagad zadan* (lit. kick hit) means 'kick', and *sili zadan* (slap hit) means 'slap'.

(77) Olāq be Omid lagad zad.
donkey to Omid kick hit
'The donkey kicked Omid.'

Within a hierarchical organization of the lexicon (Davis & Koenig 2020, Chapter 4 of this volume), the content of the simplex verb is higher and less specialized than that of the predicative noun. Thus, the content of the complex predicate can be simply that of the noun. This is reminiscent of the way Wechsler (1994) represents the import of a PP with a verb like *talk*; the verb content itself is represented as an *soa* with one participant, the talker; the verb can take a number of PP complements (headed by *to*, *about*, ...), which add semantic information describing the situation. The result is a description of an *soa* which conjoins partial descriptions. Similarly here, the conjunction of the two contents is in effect, but, the content of *lagad* 'kick' being more specialized than that of *zadan*, the con-

tent of the complex predicate identifies with it. The complement of the complex predicate may be an NP or a PP headed by *be* 'to' (the preposition is optional).

(78)
$$\begin{bmatrix} \textit{zadan1-lexeme} \\ \textsc{cat}\ \textit{verb} \\ \textsc{arg-st}\left\langle \textsc{np}_k,\ \textit{(be)}\ \textsc{np}_m,\ \textsc{n}\begin{bmatrix}\textsc{cat}\ [\textsc{prd}\ +] \\ \textsc{cont}\ \boxed{1}\end{bmatrix}\right\rangle \\ \textsc{cont}\begin{bmatrix}\textit{soa} \\ \textsc{nucleus}\ \boxed{1}\begin{bmatrix}\textit{kick-relation} \\ \textsc{agent}\ k \\ \textsc{undergoer}\ m\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Another case where the combination gives more information than the simplex verb is when this verb takes as its predicative complement a noun denoting an instrument crucially involved in the situation (Bonami & Samvelian 2010). Such are, in different domains, *čāqu zadan* (lit. knife hit) 'stab', *telefon zadan* (lit. phone hit) 'phone', *piāno zadan* (lit. piano hit) 'play the piano'. We illustrate here *šāne zadan* (lit. comb hit) 'comb'.

(79) Maryam mu-hā=yaš=rā šāne zad.
Maryam hair-PL=3SG=RA comb hit
'Maryam combed her hair.'

Stefan: Avoid bold weherever possible

(80)
$$\begin{bmatrix} \textit{zadan2-lexeme} \\ \textsc{synsem}\begin{bmatrix}\textsc{cat}\ \textit{verb} \\ \textsc{arg-st}\left\langle \textsc{np}_k,\ \textsc{np}_m,\ \textsc{n}\begin{bmatrix}\textsc{cat}\ [\textsc{prd}\ +] \\ \textsc{cont}\ \boxed{2}\end{bmatrix}\right\rangle \\ \textsc{cont}\ \boxed{2}\begin{bmatrix}\textit{soa} \\ \textsc{sit}\ \boxed{1} \\ \textsc{nucleus}\begin{bmatrix}\textit{comb-relation} \\ \textsc{agent}\ k \\ \textsc{undergoer}\ m\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \textsc{background}\left\{\textbf{involves}\ (\boxed{1},\ \exists\, x\ [\textbf{comb}\ (x) \wedge \textbf{use}\ (\boxed{1},\ k, x)]\right\}\end{bmatrix}$$

The condition in the background can be read as follows: the situation $\boxed{1}$ involves that there exists a comb and that the agent k uses it in that situation. Although the complex predicate takes on the content of the predicative complement, the semantics does not rely solely on identifying with that of the more

specialized element, as in the preceding case, but exploits a schema based on information present in the background: the existence of an object (the [PRD −] correspondent of the predicative complement), and the fact that, in the situation, such an object is used (see Bonami & Samvelian 2010: 10).

Further from a compositional or recoverable meaning is the use of *zadan* or more precisely *xod=rā zadan* (lit. self zadan) with a series of nouns denoting illnesses, handicaps or problematic states (like stupidity, ignorance…): it means 'to pretend, feign' the illness or state in question (example (81) from Samvelian 2012: 223).

(81) Maryam xod=rā be divānegi zad.
Maryam self=RA to madness hit
'Maryam feigned madness.'

This use of *zadan* may be seen as an extension of its use with nouns denoting some sort of deceit, such as *gul zadan* (lit. deceit hit) 'to deceive': as in (78), the noun imposes its content to the combination, with a metaphorical use of the verb, retaining from the physical violence meaning of *zadan* 'hit') the idea of an action to the detriment of someone. Nevertheless, nothing in the actual combination in (81) indicates deception. Not all nouns for illnesses are acceptable, only those which cannot be really verified in the situation: a state of fatigue, but not a heart attack. We group them as *internal-problematic-states*. Here the combination of the verb and the noun is standard, in that the noun is a semantic argument of the verb, but the meaning of the verb is unpredictable.

(82)
$$\left[\begin{array}{l} \textit{zadan3-lexeme} \\ \textsc{cat}\ \textit{verb} \\ \textsc{arg-st}\left\langle \textsc{np}_k, \textsc{pro}_k, \textsc{pp}\left[\begin{array}{l} \textsc{cat}\left[\begin{array}{l}\textsc{pform be}\\ \textsc{prd} +\end{array}\right] \\ \textsc{cont}\ \boxed{1} \mid \textsc{nucleus}\left[\begin{array}{l}\textit{internal-problematic-state}\\ \textsc{experiencer}\ k\end{array}\right]\end{array}\right]\right\rangle \\ \textsc{cont}\left[\begin{array}{ll}\textit{soa} & \\ \textsc{nucleus} & \left[\begin{array}{l}\textit{pretend-relation}\\ \textsc{agent}\ k\\ \textsc{theme}\ \boxed{1}\end{array}\right]\end{array}\right] \end{array}\right]$$

Note that, contrary to *zadan1-lexeme*, with which *be* is optional, the *zadan3-lexeme* requires the predicative complement to be in fact a PP, headed by *be*. We assume that the preposition *be* (frequent in the complement of a complex

predicate) is contentless and shares syntactic (the [PRD ±] value) and semantic information with its complement, the predicative N ([CONT 1]); this is indicated by treating *be* as the value of the feature P(REPOSITION) FORM (Pollard & Sag 1987: Chapter 3).

Finally, we turn to an idiom: *dast zadan* (lit. hand hit) meaning 'start'. The combination may mean, in a more recoverable way 'to touch' with PP complements denoting concrete objects (as in (64)), or 'to applaud' with a PP complement denoting a person (83a) (from Samvelian 2012: 45). However, it means 'to start' with a PP complement denoting an event as in (83b) (from Samvelian 2012: 185).

(83) a. Barā=yaš xeyli dast zad-im.
for=3SG a.lot hand hit-1PL
'We applauded him a lot.'

b. Kārgar-ān be e'tesāb dast zad-and.
worker-PL to strike hand hit-3PL
'The workers went on strike.'

To represent the idiom, we resort to the feature LID (lexical identifier) which is associated with lexemes in the lexicon, and contains morpho-syntactic as well as semantic information, and allows the verb to select a specific form (Sag 2007; 2012). Thus, the noun *dast* in the idiom has a LID value *dast*. The preposition *be*, which heads the other complement, is the same as in *zadan-3*: it identifies its content with that of its complement.

The description of *zadan-4*, which occurs in the idiom *dast zadan* 'to start' is given in (84). The predicative noun complement being specified with the LID value *dast*, it is only in combination with the noun *dast* that *zadan* acquires this meaning.

(84)
$$\left[\begin{array}{l} \textit{zadan4-lexeme} \\ \textsc{cat}\ \textit{verb} \\ \textsc{arg-st}\left\langle \textsc{np}_k,\ \textsc{pp}\left[\begin{array}{l}\textsc{pform be}\\ \textsc{cont}\ \boxed{1}\end{array}\right],\ \textsc{n}\left[\begin{array}{l}\textsc{cat}\ [\textsc{prd}\ +]\\ \textsc{lid}\ \textit{dast}\end{array}\right]\right\rangle \\ \textsc{cont}\left[\begin{array}{l}\textit{soa}\\ \textsc{nucleus}\left[\begin{array}{l}\textit{start-relation}\\ \textsc{agent}\ k\\ \textsc{soa-arg}\ \boxed{1}\ |\ \textsc{nucleus}\ \textit{event-relation}\end{array}\right]\end{array}\right] \end{array}\right]$$

6 Conclusion

Following the usual definition of complex predicates in HPSG, as a series of (at least) two predicates, of which one is the head attracting the complements of the other, we have studied them in different languages, Romance languages, German, Korean, and Persian. These languages illustrate three ways in which argument attraction (or composition) manifests itself: bounded dependencies (such as clitic climbing), flexible word order, mixing the arguments of the two predicates, and special semantic combination, which builds a lexeme out of the two predicates (particularly from the verb and the noun in light verb constructions).

HPSG is well equipped to represent this phenomenon. The feature structure associated with a predicate specifies which complements it is waiting for, and the feature structure associated with a phrase allows it to be non-saturated regarding its complements, a possibility exploited by a number of verbs, which are or can be the head of a complex predicate: the phenomenon is lexically driven. Certain verbs have two entries, one which takes a saturated complement, one which is the head of a complex predicate; but a head can be itself flexible, accepting a complement which is saturated, partially saturated or not saturated at all: this is the case of the copula in Romance languages.

Crucially, the mechanism of argument attraction is not tied with a specific syntactic structure; on the contrary, it is compatible with different structures: it is shown that the properties of a verbal complex (where the two predicates form a syntactic unit by themselves) differ from those of a flat structure (where the two predicates form a unit with the complements). The structures can characterize a language as opposed to another one (Spanish restructuring verbs contrast with Italian ones), but they can also be present in the same language (as in Romanian, for instance, see Monachesi 1999).

Similarly, the mechanism of argument attraction does not induce a specific semantic combination: it is compatible with a compositional semantics (as in a verb + adjective combination in Persian, or modal verb + infinitive complement in Romance languages), as well as a variety of senses specific to the combination of the verb with a class of complements. The description in HPSG can exploit the hierarchical organization of the lexicon and the mechanism of conjuction of descriptions (informally referred to as unification) (as with combinations specializing the meaning of the verb in Persian), the richness of the feature structure, appealing to a background feature (as when the noun corresponds to an instrument implied in the action), or to a special feature allowing to point to a specific form (for representing idioms).

Abbreviations

Stefan: Only add stuff that is not in Leipzig Glossing rules. The word particule is not English, it seems. Add to the table below. Make sure it is complete. EZ is missing. RA as well.

CAUS	causative
DAT	dative
DECL	particule indicating a declarative
PST-PTCP	past participle

CAUS (causative), DAT (dative), DECL (particule indicating a declarative), NEG (negation), NOM (nominative), PART (particule), PASS (passive), PL (plural), PROGR (progressive), SG (singular), PST-PTCP (past participle); = indicates an enclitic.

Acknowledgements

We thank Anne Abeillé, Gabriela Bîlbîe, Olivier Bonami, Caterina Donati, Jean-Pierre Koenig, Kil Soo Ko, Paola Monachesi, Stefan Müller, Tsuneko Nakazawa, Daniel Rojas Plata, and Stephen Wechsler.

References

Abeillé, Anne. 2020. Control and Raising. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 447–491. Berlin: Language Science Press. DOI: ??.

Abeillé, Anne & Danièle Godard. 1994. The complementation of tense auxiliaries in French. In *Proceedings of WCCFL*, vol. 13, 157–172. Stanford: CSLI Publications.

Abeillé, Anne & Danièle Godard. 2000. French word order and lexical weight. In R. Borsley (ed.), *The Nature and Function of Syntactic Categories*, 325–360. New York: Academic Press.

Abeillé, Anne & Danièle Godard. 2001a. Deux types de prédicats complexes dans les langues romanes. *Linx* 45. 167–175.

Abeillé, Anne & Danièle Godard. 2001b. Varieties of ESSE in Romance languages. In D. Flickinger & A. Kathol (eds.), *Proceedings of 7th Conference of HPSG*, 2–22. Stanford: CSLI Publications.

Abeillé, Anne & Danièle Godard. 2002. The syntactic structure of French auxiliaries. *Language* 78(3). 404–452.

Abeillé, Anne & Danièle Godard. 2010. Complex predicates in the Romance languages. In Danièle Godard (ed.), *Fundamental Issues in the Romance Languages*, 107–170. Stanford: CSLI Publications.

Abeillé, Anne, Danièle Godard & Philip Miller. 1995. La double structure des causatifs et des verbes de perception. In J. Labelle (ed.), *Actes 2d Colloque des lexiques-grammaires comparés*, 129–150. Montréal: UQAM.

Abeillé, Anne, Danièle Godard, Philip H. Miller & Ivan A. Sag. 1998. French bounded dependencies. In Sergio Balari & Luca Dini (eds.), *Romance in HPSG* (CSLI Lecture Notes 75), 1–54. Stanford, CA: CSLI Publications.

Abeillé, Anne, Danièle Godard & Ivan A. Sag. 1998. Two kinds of composition in French complex predicates. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 1–41. San Diego: Academic Press.

Aissen, Judith & David M. Perlmutter. 1983. Clause reduction in Spanish. In D. Perlmutter (ed.), *Studies in Relational Grammar*, vol. 1, 360–403. Chicago: Chicago University Press.

Andrews, Avery D. & Christopher D. Manning. 1999. *Complex predicates and information spreading in lfg*. Stanford, CA: CSLI Publications.

Bech, Gunnar. 1955. *Studien über das deutsche Verbum infinitum. Det Kongelige Danske videnskabernes selskab. Historisk- Filosofiske Meddelelser*. 2nd unrevised edition published 1983 by Max Niemeyer Verlag, Tübingen (Linguistische Arbeiten 139). Kopenhagen: Munksgaard.

Bonami, Olivier & Berthold Crysmann. 2018. Lexeme and flexeme in a formal theory of grammar. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer (eds.), *The lexeme in descriptive and theoretical morphology* (Empirically Oriented Theoretical Morphology and Syntax 4), 175–202. Berlin: Language Science Press.

Bonami, Olivier & Pollet Samvelian. 2010. Persian complex predicates: Lexeme formation by itself. In *Décembrettes 7. Colloque international de morphologie*. Toulouse, 2-3 décembre 2010. http://www.llf.cnrs.fr/fr/node/3496.

Bonami, Olivier & Pollet Samvelian. 2015. The diversity of inflectional periphrasis in persian. *Journal of Linguistics* 51(2). 327–382.

Bonami, Olivier & Gert Webelhuth. 2012. The phrase structural diversity of periphrasis: A lexicalist account. In Marina Chumakina & Greville G. Corbett (eds.), *Periphrasis: The Role of Syntax and Morphology in Paradigms* (Proceedings of the British Academy 180), 141–167. Oxford: Oxford University Press/ British Academy.

Borsley, Robert D. & Berthold Crysmann. 2020. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 493–546. Berlin: Language Science Press. DOI: ??.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language & Linguistic Theory* 19(1). 1–65. DOI: 10.1023/A:1006473306778.

Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press. (25 February, 2018).

Butt, Miriam. 1995. *The Structure of Complex Predicates in Urdu*. Stanford: CSLI Publications.

Butt, Miriam. 2010. The light verb jungle: Still hacking away. In Mengistu Amberber, Brett Baker & Mark Harvey (eds.), *Complex predicates: cross-linguistic perspectives on event structure*, 48–78. Cambridge University Press.

Choi, Incheol & Stephen Wechsler. 2002. Mixed categories and argument transfer in the Korean light verb construction. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, 103–120. Stanford: CSLI Publications.

Chung, Chan. 1998. Argument composition and long-distance scrambling in Korean. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 159–220. San Diego: Academic Press.

Davis, Anthony & Jean-Pierre Koenig. 2020. The nature and role of the lexicon in HPSG. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 105–152. Berlin: Language Science Press. DOI: ??.

De Kuthy, Kordula & Walt Detmar Meurers. 2001. On partial constituent fronting in German. *Journal of Comparative Germanic Linguistics* 3(3). 143–205.

Geach, Peter Thomas. 1970. A program for syntax. *Synthese* 22. 3–17.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Gunji, Takao & Kôiti Hasida (eds.). 2012. *Topics in Constraint-Based Grammar of Japanese* (Studies in Linguistics and Philosophy 68). Springer.

Hale, Kenneth & Samuel Jay Keyser. 1997. On the complex nature of simple predicators. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates* (CSLI Lecture Notes 64), 29–65. Stanford, CA: CSLI Publications.

Haspelmath, Martin. 2016. The serial verb construction: Comparative concept and cross linguistic generalizations. *Language and Linguistics* 17(3). 291–319.

Hinrichs, Erhard & Tsuneko Nakazawa. 1999. VP relatives in German. In Gosse Bouma, Erhard Hinrichs, M. Kruiff Geert-Jan & Richard Oehrle (eds.), *Constraints and Resources in Natural Language Syntax and Semantics*, 63–80. Stanford: CSLI Publications.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989. Flipped out: AUX in German. In *Aspects of German VP structure* (SfS-Report-01-93). Tübingen: Eberhard-Karls-Universität Tübingen.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 11–38. Stanford, CA: CSLI Publications.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1998. Third Construction and VP extraposition in German: An HPSG analysis. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 115–157. San Diego: Academic Press.

Höhle, Tilman N. 2018. Spuren in HPSG. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur deutschen Grammatik: Gesammelte Schriften*, 491–498. Berlin: Language Science Press.

Karimi-Doostan, Gholamhossein. 1997. *Light verb constructions in Persian*. Department of Language & Linguistics, University of Essex. (Doctoral dissertation).

Kathol, Andreas. 1998. Constituency and linearization of verbal complexes. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 221–270. San Diego: Academic Press.

Kathol, Andreas. 2000. *Linear syntax*. New York, NY: Oxford University Press.

Kim, Jong-Bok. 2016. *The syntactic structures of Korean: A Construction Grammar perspective*. Cambridge, UK: Cambridge University Press.

Kim, Jong-Bok. 2020. Negation. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 753–787. Berlin: Language Science Press. DOI: ??.

Kiss, Tibor. 1994. Obligatory coherence: The structure of German modal verb constructions. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in*

Head-Driven Phrase Structure Grammar (CSLI Lecture Notes 46), 71–108. Stanford, CA: CSLI Publications.

Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Max Niemeyer Verlag.

Koenig, Jean-Pierre. 1999. *Lexical relations* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Kraak, Esther. 1998. A deductive account of French object clitics. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 313–368. San Diego: Academic Press.

Lee, Sun-Hee. 2001. Argument composition and linearization: Korean complex predicates and scrambling. In *OSUWPL* (56), 53–78. Ohio State University. Department of Linguistics.

Manning, Christopher, Ivan A. Sag & Masayo Iida. 1999. The Lexical Integrity of Japanese Causatives. In R. Levine & G. Green (eds.), *Readings in Modern Phrase Structure Grammar*, 39–79. Cambridge: Cambridge University Press.

Megerdoomian, Karine. 2002. *Beyond words and phrases: A unified theory of predicate composition.* University of Southern California. (Doctoral dissertation). http://www.zoorna.org/publications.html (23 December, 2015).

Meurers, Detmar. 2000. *Lexical generalizations in the syntax of German non-finite constructions.* Arbeitspapiere des SFB 340 No. 145. Universität Tübingen.

Meurers, Detmar. 2001. To Flip or not to Flip: On the Nature of Irregularities in the German Verbal Complex. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, 235–246. Stanford: CSLI Publications.

Miller, Philip H. & Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language & Linguistic Theory* 15(3). 573–639.

Monachesi, Paola. 1998. Italian restructuring verbs: A lexical analysis. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 313–368. San Diego: Academic Press.

Monachesi, Paola. 1999. On certain properties of Romanian auxiliary (and modal) verbs. In Gosse Bouma, Erhard W. Hinrichs, Geert-Jan M. Kruijff & Richard T. Oehrle (eds.), *Constraints and resources in natural language syntax and semantics* (Studies in Constraint-Based Lexicalism 3). Stanford, CA: CSLI Publications.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan. 2003. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *Journal of Linguistics* 39(2). 275–325. DOI: 10.1017/S0022226703002032.

Müller, Stefan. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655. DOI: 10.1017/S0022226709990284.

Müller, Stefan. 2012. *On the copula, specificational constructions and type shifting*. Ms. Freie Universität Berlin.

Müller, Stefan. 2018. *German Clause Structure: An Analysis with Special Consideration of Apparent Multiple Frontings*. To appear. hpsg.hu-berlin.de/~stefan/Pub/gs.html.

Müller, Stefan. 2020. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 335–379. Berlin: Language Science Press. DOI: ??.

No, Yongkyoon. 1991. *Case alternations on verb-phrase internal arguments*. The Ohio State University. (Doctoral dissertation).

Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology*. University of California, Berkeley. (Doctoral dissertation).

Pollard, Carl & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Poornima, Shakthi & Jean-Pierre Koenig. 2009. Hindi Aspectual Complex Predicates. In *Proceedings of the International Conference on Head-Driven Phrase Structure Grammar*, 276–296.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Rentier, Gerrit. 1994. Dutch cross serial dependencies in HPSG. In Makoto Nagao (ed.), *Proceedings of COLING 94*, 818–822. Kyoto, Japan: Association for Computational Linguistics. http://www.aclweb.org/anthology/C94-2130 (25 February, 2018).

Riehemann, Susanne Z. 1998. Type-based derivational morphology. *Journal of Comparative Germanic Linguistics* 2(1). 49–77. DOI: 10.1023/A:1009746617055.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Reidel.

Ryu, Byong-Rae. 1993. *Structure sharing and argument transfer: An HPSG approach to verbal noun constructions*. SfS Report 04-93. Eberhard-Karls-Universität Tübingen: Seminar für Sprachwissenschaft.

Sag, Ivan A. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 394–414. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2007/ (25 February, 2018).

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 69–202. Stanford, CA: CSLI Publications.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal Introduction*. 2nd edition. Stanford: CSLI Publications.

Salvi, Giampaolo. 1980. Gli ausiliari 'essere' e 'avere' in italiano. *Acta Linguistica Academiae Scientiarum Hungaricae* 30. 137–162.

Samvelian, Pollet. 2012. *Grammaire des prédicats complexes: les constructions nom-verbe*. Paris: Lavoisier.

Sells, Peter. 1991. Complex verbs and argument structures in korean. In Susumo Kuno et al. (eds.), *Harvard Studies in Korean Linguistics IV*, 395–406. Seoul: Hanshin.

Suñer, Margarita. 1982. *The Syntax and Semantics of Presentational Types*. Washington, D.C.: Georgetown University Press.

Webelhuth, Gert. 1998. Causatives and the nature of argument structure. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 369–422. San Diego: Academic Press.

Wechsler, Stephen. 1994. Preposition selection outside the lexicon. In Raul Aranovich, William Byrne, Susanne Preuss & Matha Senturia (eds.), *Proceedings of 13th West Coast conference on Formal Linguistics*, 416–431. UCSD.

Yoo, Eun-Jung. 2003. Case marking in Korean auxiliary verb constructions. In Jong-Bok Kim & Stephen Wechsler (eds.), *The proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 413–438. Stanford, CA: CSLI Publications.