Chapter 12

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. It characterizes auxiliaries, copulas and predicative verbs, certain control or raising and causative verbs, and light verbs. It has been studied in HPSG in different languages, Romance and Germanic languages, Korean and Persian, which illustrate different aspects of the phenomenon. Romance languages, as well Korean, show that argument inheritance is distinct from structure: it is compatible with two 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?

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 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 (Hin-

richs & 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:

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

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

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

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

     b. Paul veut   le lire. | * Paul le veut lire.[1]
         Paul wants it read

This definition of complex predicate 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) 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

---

[1]Possible in an earlier stage of French.

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 1996). 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
        '(someone1) 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 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* ('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).

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

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 (**AGS1998**; Abeillé & Godard 1994; 2000; 2001a,b; 2002; 2010; Abeillé et al. 1995; 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* or *être*) 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).

---

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

(6) a. Lola a acheté des pommes, et Alice (a) cueilli des pêches.

'Lola has bought apples, and Alice (has) picked peaches'

b. Lola a acheté des pommes, et Alice (a acheté) des pêches.

'Lola has bought apples, and Alice (has bought) peaches'

c. # Lola a acheté des pommes, et Alice a des pêches.

'Lola has bought apples, and Alice has peaches'

d. Lola pense acheter des pommes, et Alice (pense) cueillir des pêches.

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

e. Lola pense acheter des pommes, et Alice (pense acheter) des pêches.

'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 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'

b. * Cette technique est impossible à réussir à maîtriser en un jour.

(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'

These two properties follow if the complements of the participle become those of *avoir*: the auxiliary "attracts" the complements of the participle. In addition, the tense auxiliary *avoir* is a subject raising verb (see chap ?): 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*) 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: $\left[\;\textsc{arg-st}\;\boxed{1}\oplus\left\langle\left[\textsc{subj}\;\boxed{1}\right]\right\rangle\oplus\boxed{2}\;\right]$

 b. Tense auxiliary: $\left[\textsc{arg-st}\;\boxed{1}\oplus\left\langle\begin{bmatrix}\textsc{light}\;+\\\textsc{arg-st}\;\boxed{1}\oplus\boxed{2}\end{bmatrix}\right\rangle\oplus\boxed{2}\right]$

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.

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; 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

$$\textit{word}\Rightarrow\left[\textsc{synsem}\mid\textsc{loc}\begin{bmatrix}\textsc{cat}\begin{bmatrix}\textsc{subj}&\boxed{1}\\\textsc{comps}&\boxed{2}\\\textsc{spr}&\boxed{3}\end{bmatrix}\\\textsc{arg-st}\;\boxed{1}\oplus\boxed{2}\oplus\boxed{3}\;\bigcirc\;\text{list}\;(\textit{non-canon})\end{bmatrix}\right]$$

In (10a), the participle *lu* selects the argument *son livre*, which is attracted by the auxiliary *a*. 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'

 b. Marie l'a   lu.

 Marie it-has read

 'Mary has read it'

$$
\begin{array}{c}
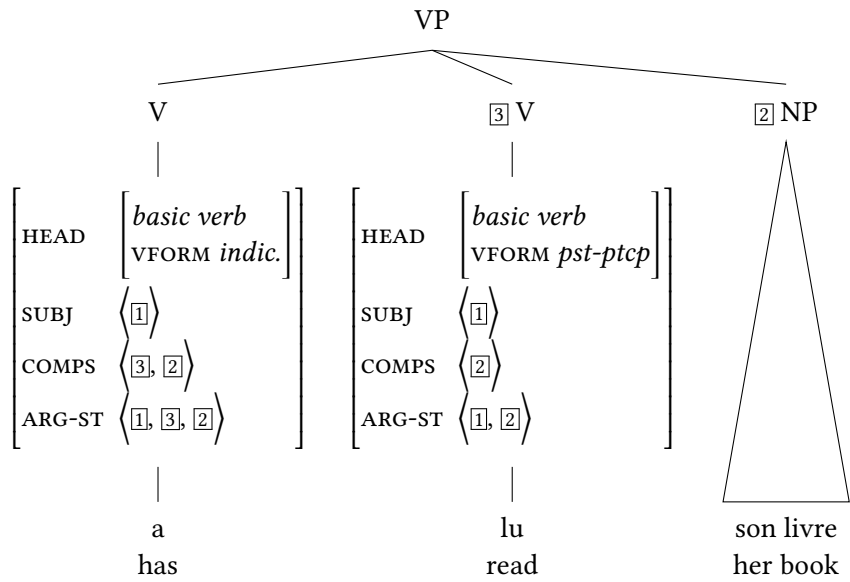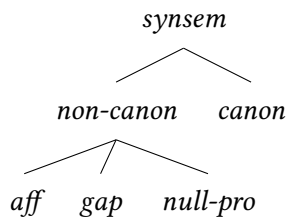\text{VP} \\
\end{array}
$$

Figure 1

Figure 2: Provide caption

Let us turn to pronominal clitics. The arguments are *synsems*, which can have different subtypes (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 **AGS1998** With basic verbs, the argument list is simply the concatenation of the subject and complements, while reduced verbs have at least one affix argument 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)*.

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.
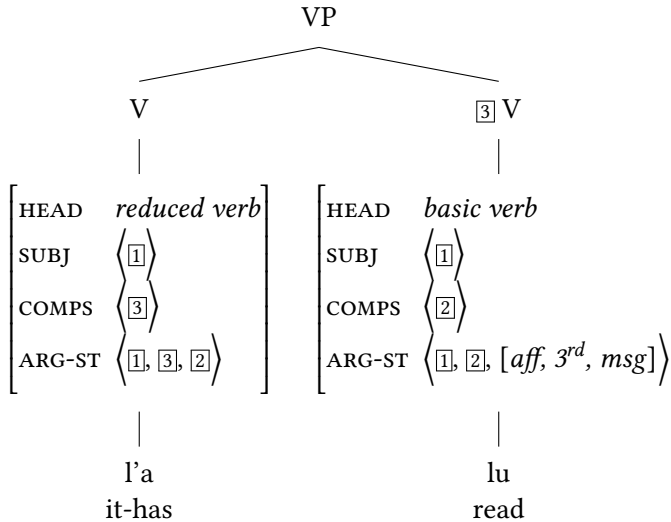

Figure 3

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. We focus here on restructuring verbs and the copula. An analysis of causative and perception verbs is proposed in (Abeillé et al. 1995; 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). 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). 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

b.   Giovanni vuole mangiar*le*.
Giovanni wants eat-them

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

d.   Juan quiere comer*las*.
Juan wants eat-them

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

f.   O João quer comê-*las*.
O João wants eat-them

g.   En Joan *les* vol menjar.   (Catalan)
En Joan them-wants eat

h.   En Joan vol menjar-*les*.
En Joan wants 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* (12c), (12d), but not with verbs only taking an infinitival complement such as *parere* (12e).

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

     b.   Queste camicie si stirano facilmente.
         These shirts sɪ-iron easily
         'These shirts iron easily'

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

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

     e.   * Queste camicie si paiono stirare facilmente.
         These shirts sɪ-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) (**AGS1998**; Monachesi 1998).

(13) Medio-Passive Lexical Rule

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

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* or *parere* 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* from *stirare*, 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) relies on *cominciare* being a restructuring verb, while *promettere* is not (14c).

(14) a. Questa canzone è facile da apprendere.   (Italian)

   'This song is easy to learn'

  b. Questa canzone è facile da cominciare a apprendere.

   'This song is easy to begin to learn'

  c. * Questa canzone è facile da promettere a apprendere.

   (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 (2)) (see **AGS1998**; Monachesi 1998). The value of MARKING is that of Italian.

$$(15) \quad \left[\begin{array}{l} \text{HEAD} \; adjective \\[4pt] \text{ARG-ST} \left\langle \text{XP}_j, \; \text{VP} \left[\begin{array}{l} \text{VFORM} \; infinitive \\ \text{MARKING} \; da \\ \text{COMPS} \left\langle null\text{-}pro \, [acc]_j \right\rangle \oplus \boxed{2} \end{array} \right] \right\rangle \end{array} \right]$$

Complex predicates can occur in this construction because their head attracts the complement of their complement. Thus, in (14b), *cominciare* is expecting the

same object as *apprendere*, 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), 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.

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

    a.   −Parlare a Maria, certamente lo vuole.   (Italian)

    b.   −Hablarle a María, seguramente quiere (pero no a su madre).

                                                 (Spanish)

        'Talk to Maria, certainly he wants to (but not to her mother)'

(17)   a.   * −Parlare, certamente glielo   vuole.   (Italian)
               Talk   certainly   to.her-it wants

    b.   * −Hablar, le-quiere   (pero no mucho   tiempo).   (Spanish)
               Talk   to.her-wants 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).[3]

(18)   Argument Attraction Lexical Rules for Romance restructuring verbs

    a.   Subject control verbs

$$\begin{bmatrix} \text{HEAD } \textit{verb} \\[2ex] \text{ARG-ST}\left\langle \text{XP}_i, \boxed{2}\begin{bmatrix} \text{HEAD}\begin{bmatrix} \textit{verb} \\ \text{VFORM } \textit{inf.} \end{bmatrix} \\ \text{SUBJ}\left\langle \text{XP}_i \right\rangle \\ \text{COMPS}\left\langle \right\rangle \end{bmatrix} \right\rangle \oplus \boxed{3} \end{bmatrix}$$

---

[3]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 et al. 1998; Abeillé & Godard 2010), concentrating on the case of so-called Restructuring verbs.

$$\Rightarrow$$

$$\left[\text{ARG-ST}\left\langle \text{XP}_i, \text{V}\begin{bmatrix} basic\text{-}verb \\ \text{LIGHT +} \\ \text{COMPS } \boxed{4} \end{bmatrix}\right\rangle \oplus \boxed{4} \oplus \boxed{3}\right]$$

b. Subject raising verbs

$$\begin{bmatrix} \text{HEAD } verb \\ \\ \text{ARG-ST } \boxed{1} \oplus \left\langle \boxed{2}\begin{bmatrix} \text{HEAD} \begin{bmatrix} verb \\ \text{VFORM } inf. \end{bmatrix} \\ \text{SUBJ } \boxed{1} \\ \text{COMPS } \langle\rangle \end{bmatrix}\right\rangle \oplus \boxed{3} \end{bmatrix}$$

$$\Rightarrow$$

$$\left[\text{ARG-ST } \boxed{1} \oplus \left\langle \text{V}\begin{bmatrix} basic\text{-}verb \\ \text{LIGHT +} \\ \text{COMPS } \boxed{4} \end{bmatrix}\right\rangle \oplus \boxed{4} \oplus \boxed{3}\right]$$

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 $\boxed{3}$. 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* is a raising verb, and the infinitive *piacere* is an impersonal 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).[4] Note that there is speakers' variation: *sembrare* 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*:

(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/her'

---

[4]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 Composition Lexical Rule (as in Monachesi 1998).

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

'To go away on vacation seems to please him/her'

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 enter either a flat structure or a verbal complex. We contrast Italian and Spanish.[5] 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).

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

          'Giovanni wants to read it often'

      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'

---

[5]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).

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

       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) (from Abeillé & Godard 2010). Again, the presence of a clitic on the head verb (*lo vuole, le volvió*) shows that this is a complex predicate construction.

(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 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 +][6], 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.
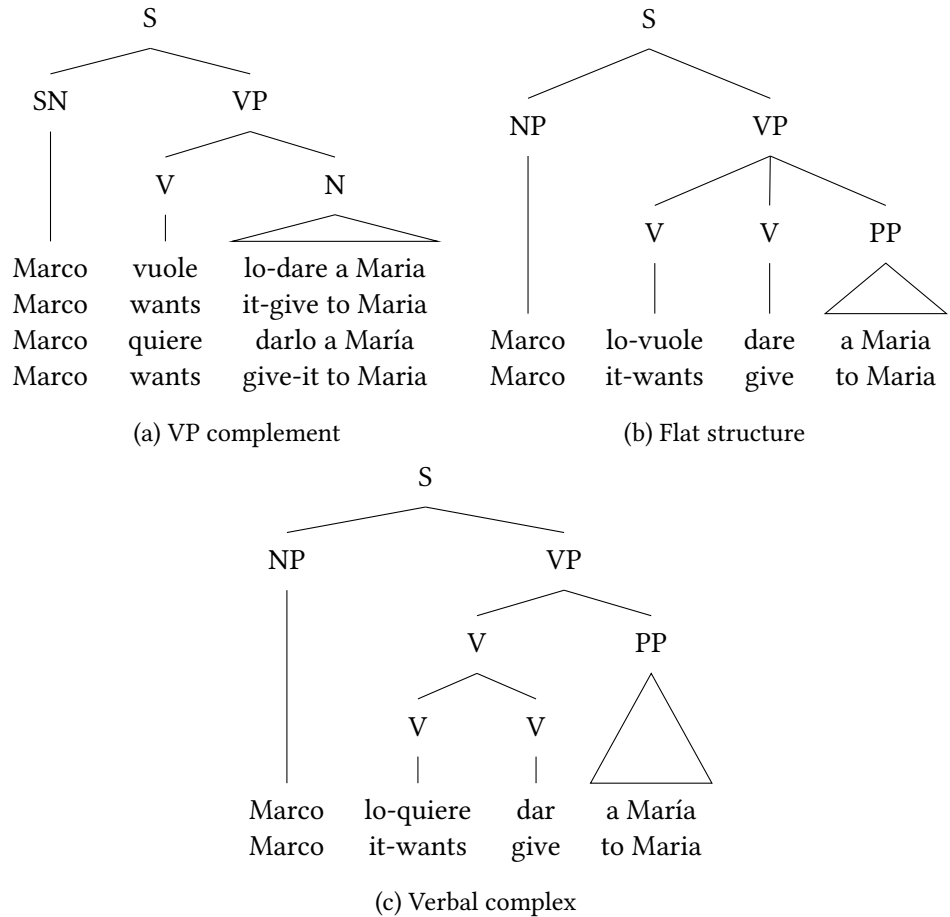

(a) VP complement (b) Flat structure

(c) Verbal complex

Figure 4

---

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

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? We propose, following (Abeillé & Godard 2010), 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, which is similar to that of (22b), repeated in (23c), if the structure is that of a verbal complex.

(23) a. Juan da [el libro de Proust] [a María] y [el (libro) de Camus] [a
Pablo]. (Spanish)

b. ?? Juan da [el libro de Proust] [a María] y [de Camus] [a Pablo].

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

In (23a) the NP *el de Camus* is parallel to and at the same level as *el libro de Proust*, the PP *a Pablo* is parallel to and at the same level as *a María*, the NP and the PP are both complements of *da*. But, in (23b), *de Camus* is parallel to *de Proust*, and not at the same level as *el libro de Proust* or as *a Pablo*: *a Pablo* corresponds to the complement of *da* while *de Camus* corresponds to the complement of the noun *libro*. 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 co-ordination 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). 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 structures 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) ⇒

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

The head-complements-phrase is usually saturated for the expected complements, but not always: list $\boxed{3}$ 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.

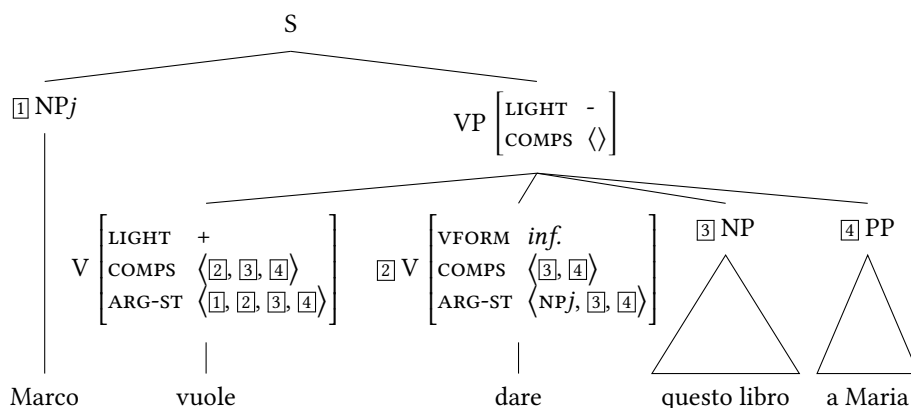The verbal complex corresponds to another kind of Head-complements-phrase,

$$
\begin{array}{c}
\text{S} \\
\end{array}
$$

Figure 5

called the Head-cluster phrase, given in (25) (adapted from Müller 2018).[7]

(25)   *Head-cluster phrase* (Spanish) ⟹

$$
\begin{bmatrix}
\text{MOTHER} \mid \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{COMPS} & \boxed{1} \\ \text{LIGHT} & + \end{bmatrix} \\[2ex]
\text{HEAD-DTR} \mid \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{COMPS} & \boxed{1} \oplus \langle \boxed{2} \rangle \\ \text{LIGHT} & + \end{bmatrix} \\[3ex]
\text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM} & \boxed{2} \begin{bmatrix} \text{LIGHT} + \end{bmatrix} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

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.[8]

---

[7]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).

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

Draft of 8th March 2020, 11:25

But some phrases can be light if they are composed of light constituents. Such is 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*) corresponds to the usual Head-complements-phrase in (24).
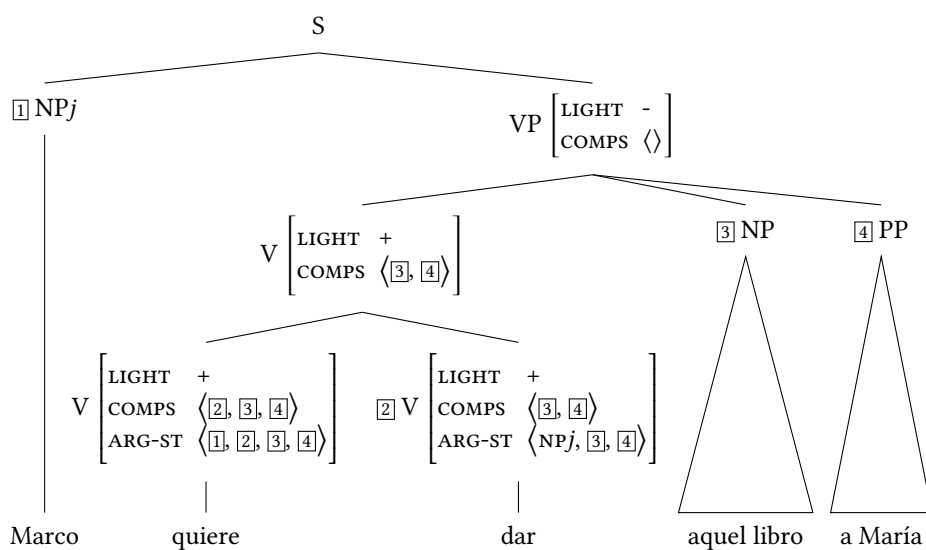
Figure 6

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* and Spanish *Marco lo-quiere dar a María* in Figures 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* on the head verb *vuole* or *quiere*.

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

VP [COMPS ⟨ ⟩]

[4] PP

V
$$\begin{bmatrix} \textit{reduced verb} \\ \text{LIGHT} \quad + \\ \text{COMPS} \quad \langle \text{[2]}, \text{[4]} \rangle \\ \text{ARG-ST} \quad \langle \text{NP}_j, \text{[2]}, \text{[3]}, \text{[4]} \rangle \end{bmatrix}$$

[2] V
$$\begin{bmatrix} \textit{basic verb} \\ \text{LIGHT} \quad + \\ \text{COMPS} \quad \langle \text{[3]}, \text{[4]} \rangle \\ \text{ARG-ST} \quad \langle \text{NP}_j, \text{[3]}\textit{aff}, \text{[4]} \rangle \end{bmatrix}$$

a Maria

lo-vuole   dare

(a)

VP [COMPS ⟨ ⟩]

V
$$\begin{bmatrix} \text{LIGHT} \quad + \\ \text{COMPS} \quad \text{[4]} \end{bmatrix}$$

[4] PP

V
$$\begin{bmatrix} \textit{reduced verb} \\ \text{LIGHT} \quad + \\ \text{COMPS} \quad \langle \text{[2]}, \text{[4]} \rangle \\ \text{ARG-ST} \quad \langle \text{NP}_j, \text{[2]}, \text{[3]}, \text{[4]} \rangle \end{bmatrix}$$

[2] V
$$\begin{bmatrix} \textit{basic verb} \\ \text{LIGHT} \quad + \\ \text{COMPS} \quad \langle \text{[3]}, \text{[4]} \rangle \\ \text{ARG-ST} \quad \langle \text{NP}_j, \text{[3]} \textit{aff}, \text{[4]} \rangle \end{bmatrix}$$
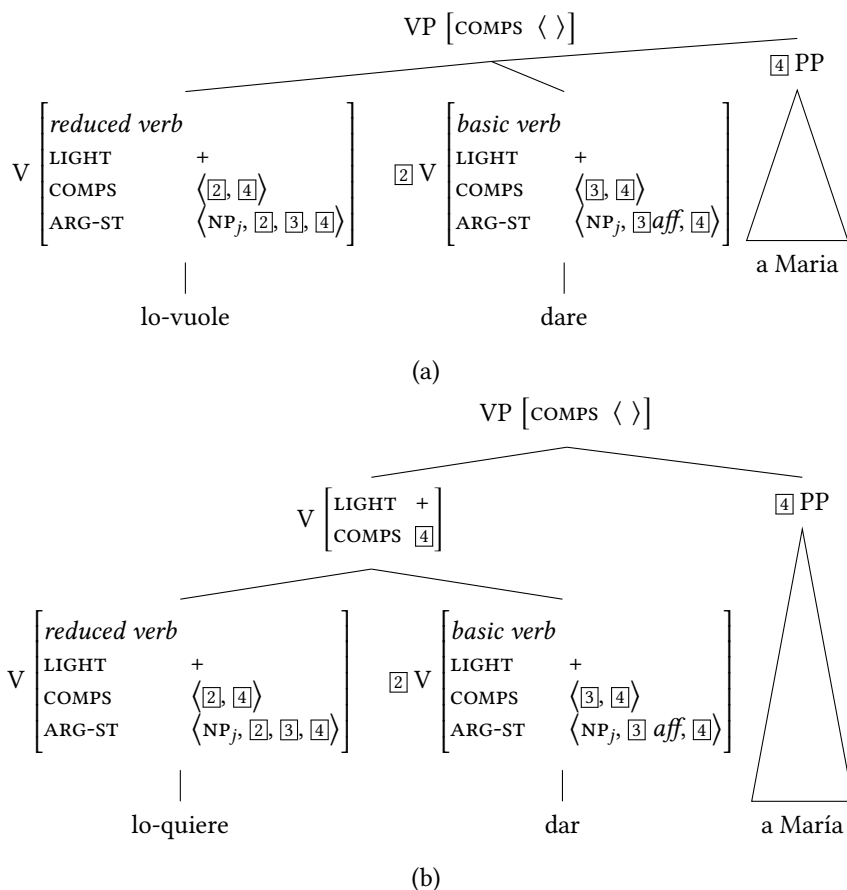
lo-quiere   dar   a María

(b)

Figure 7

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) $\begin{bmatrix} \textit{phrase} \\ \text{LIGHT} \; \text{[1]} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{NON-HEAD-DTRS} \mid \text{LIGHT} \; \text{[1]} \end{bmatrix}$ \hfill (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 dif-

ferent additional constraint.

(27) $\begin{bmatrix} phrase \\ \text{LIGHT} + \end{bmatrix} \Rightarrow \begin{bmatrix} \text{NON-HEAD-DTRS list } ([\text{HEAD} \neg \ verbal]) \end{bmatrix}$ (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) $\begin{bmatrix} \text{V} \begin{bmatrix} \text{COMPS} \langle ... \boxed{1} ... \rangle \end{bmatrix} \end{bmatrix}$ precedes $\begin{bmatrix} \text{SYNSEM } \boxed{1} \end{bmatrix}$ (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)[9]. 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.

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).

(29) a. Jean lui était fidèle. (French)
John to.him was faithful

b. Giovanni le era fedele. (Italian)

c. Juan le era fiel. (Spanish)

d. O Joãn era-lhe fiel. (Portuguese)

e. En Joan li era fidel. (Catalan)

---

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

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

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).

(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

    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

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

        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)

        c. Fiel,   lo es más a sus amigos que a sus convicciones
            Faithful, it is more to his friends than to his political
            políticas.   (Spanish)
            convictions
            '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 its complements behind. In the latter case, this complement 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*), as part of a concessive adjunct.

(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 intérêts. (French)

        '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
        As   faithful as he is   to his friends, he does not lose sight of
        ses intérêts.
        his interests.

    c. Aussi fidèle qu'il leur   soit, il ne   perd pas de   vue ses
        As   faithful as he to.them is,   he does not lose sight of   his
        intérêts.
        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)
      b. Romeo le sarà probabilmente fedele.   (Italian)
      c. Romeo le será probablemente fiel.   (Spanish)

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

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

Like tense auxiliaries, the copula is a subject raising verb, hence the identical value $\boxed{1}$ for its subject and that of the complement, which allows it to be empty. Its complement differs from that of a tense auxiliary (7) 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.[10]

$$
\text{VP} \begin{bmatrix} \text{COMPS} & \langle \rangle \end{bmatrix}
$$

$$
\text{V} \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{3} \rangle \\ \text{ARG-ST} & \langle \boxed{1}, \boxed{3} \rangle \end{bmatrix}
\qquad
\boxed{3}\ \text{AP} \begin{bmatrix} \text{HEAD} & [\text{PRD} +] \\ \text{LIGHT} & - \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \rangle \end{bmatrix}
$$

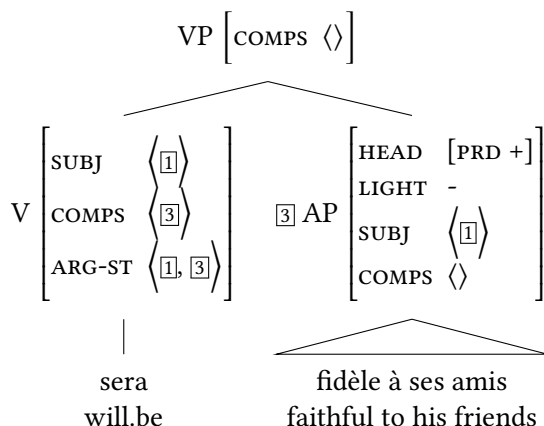|	
sera	fidèle à ses amis
will.be	faithful to his friends

Figure 8

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.

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 &

---

[10]Note that the complements included in a predicative PP are not attracted to the copula in a general way.
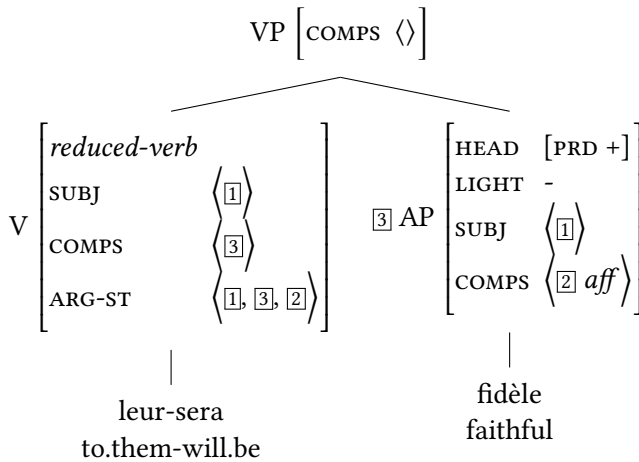
Draft of 8th March 2020, 11:25

Figure 9

Nakazawa 1989; 1994; Kiss 1994; 1995; Hinrichs & Nakazawa 1998; Kathol 1998; 2000; Meurers 2000; 2001; De Kuthy & Meurers 2001; Müller 2002; 2003; 2013) 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 auxiliaries and control verbs in Korean.

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).

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* may occur between the two verbs (35a) (from Müller 2002), 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).

(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

      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* (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 from Müller 2002), or pied-piping of the infinitive in relative clauses (36g,f).[11]

(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
          'that Karl will not read the book'

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

      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

---

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

Draft of 8th March 2020, 11:25

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

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

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 (Peter)* and *kaufen (das Auto)* are not interleaved. In (37b), *Peter*, the complement of *sehen*, occurs between *das Auto*, which is the complement of *kaufen*, and *kaufen* (examples from Hinrichs & Nakazawa 1998).

(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

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 trigger a modification of the ordering domain. More precisely, it would induce domain union of the two ordering domains associated with the two verbal projections (Reape 1994), thus allowing the order in (38b), for instance, while the structure would remain the same with a VP complement (as in Figure 10a) (chapter ??). 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*) alternates with a coherent construction such as (38b), with a corresponding nominative (examples from Müller 2002).
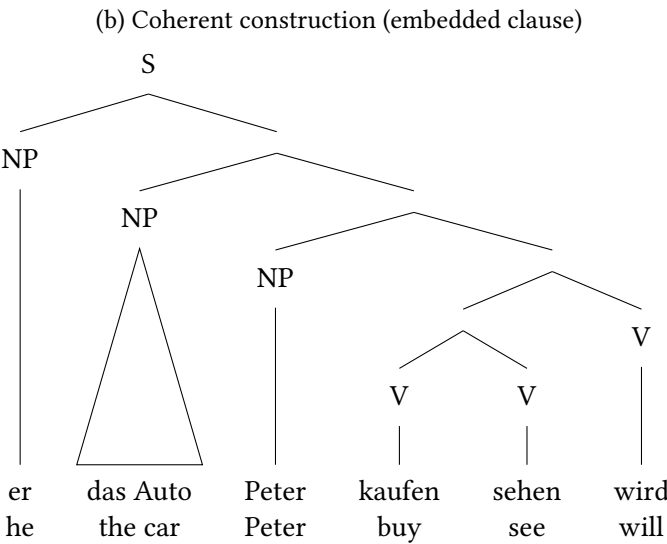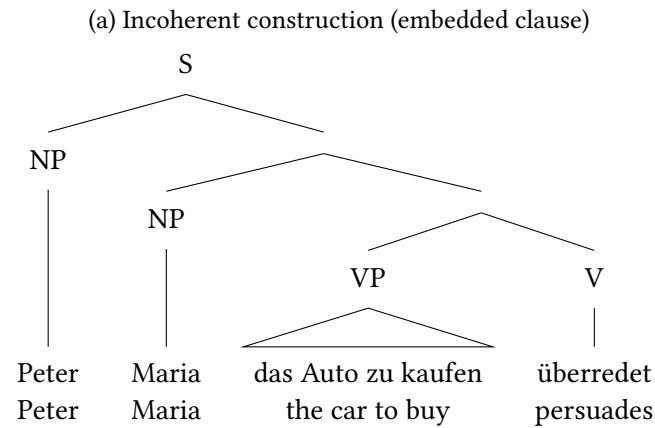
(a) Incoherent construction (embedded clause)

```
                        S
           ┌────────────┴──────────┐
          NP               ┌───────┴───────┐
           │              NP           ┌────┴────┐
           │               │          VP         V
           │               │       ┌───┼───┐      │
         Peter           Maria  das Auto zu kaufen  überredet
         Peter           Maria    the car to buy    persuades
```

(b) Coherent construction (embedded clause)

```
                        S
          ┌─────────────┴────────┐
         NP              ┌────────┴───────┐
          │             NP         ┌──────┴──────┐
          │              │        NP        ┌─────┴────┐
          │              │△        │      ┌──┴──┐       V
          │             ╱  ╲       │      V     V       │
          │            ╱    ╲      │      │     │       │
          er        das Auto    Peter   kaufen sehen   wird
          he         the car    Peter    buy    see    will
```

Figure 10

Draft of 8th March 2020, 11:25

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

      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 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*, which occurs before *zu reparieren*, while modifying *versucht*. In a coherent construction, an adverb can scope over any of the verbs that belong to it. In (38c), *zu schlafen* is not part of the coherent construction, because it is extraposed; *nicht* can have scope over *darf* or *versuchen*, not over *schlafen*. 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).

(39) *Head-complement-phrase* (German) $\Rightarrow$

$$\begin{bmatrix} \text{MOTHER} \mid \text{SYNSEM} & \begin{bmatrix} \text{LOC} \mid \text{CAT} \mid \text{COMPS } \boxed{1} \oplus \boxed{3} \\ \text{LIGHT -} \end{bmatrix} \\ \text{HEAD-DTR} \mid \text{SYNSEM} & \begin{bmatrix} \text{LOC} \mid \text{CAT} \mid \text{COMPS } \boxed{1} \oplus \langle \boxed{2} \rangle \oplus \boxed{3} \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \langle [\text{SYNSEM } \boxed{2}] \rangle \end{bmatrix}$$

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 -].[12] The structure of (40) is exemplified in Figure 11.

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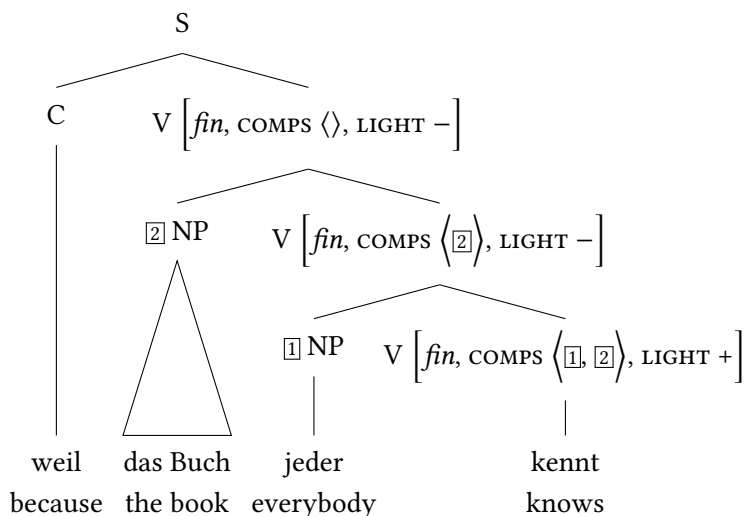
    'because everybody knows the book'

Figure 11

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).

The phrase structure constraint allowing complex predicates is as in (41) (Müller 2013; 2018). It is called *the head-cluster-phrase*, rather than *the verbal-complex-phrase*, because it is not specialized for verbal heads.[13]

---

[12]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 (see chapter ?? for extraction) (Müller 2018).

[13]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

(41)   *Head-cluster-phrase* (German) ⇒

$$
\begin{bmatrix}
\text{MOTHER} \mid \text{SYNSEM} & \begin{bmatrix} \text{LOC} \mid \text{CAT} \mid \text{COMPS} \; \boxed{1} \\ \text{LIGHT} + \end{bmatrix} \\
\text{HEAD-DTR} \mid \text{SYNSEM} & \begin{bmatrix} \text{LOC} \mid \text{CAT} \mid \text{COMPS} \; \boxed{1} \oplus \langle \boxed{2} \rangle \\ \text{LIGHT} + \end{bmatrix} \\
\text{NON-HEAD-DTRS} & \langle [\text{SYNSEM} \; \boxed{2} \; [\text{LIGHT} +]] \rangle
\end{bmatrix}
$$

We illustrate 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), and that of *sehen*, an object raising verb and an obligatorily coherent verb, in (43). The 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* is included in list $\boxed{1}$, and that of *sehen* is included in list $\boxed{2}$. 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)

$$
\begin{bmatrix}
\text{HEAD} \; verb \\
\text{COMPS} \; \boxed{1} \oplus \langle \text{V} \; [bse, \text{COMPS} \; \boxed{1}, \text{LIGHT} +] \rangle
\end{bmatrix}
$$

(43)   *sehen* (obligatory coherent verb)

$$
\begin{bmatrix}
\text{HEAD} \; verb \\
\text{COMPS} \; \boxed{1} \oplus \boxed{2} \oplus \langle \text{V} \; [bse, \text{COMPS} \; \boxed{2}, \text{LIGHT} +] \rangle
\end{bmatrix}
$$

Sentence (37b) is represented in Figure 12.

4.1.3 The German copula

The copula in German, with an adjectival argument, is also the head of a complex predicate.[14] The subject of the copula and the complements of the adjectives can be permuted (examples from Müller 2002; 2013) (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

---

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.

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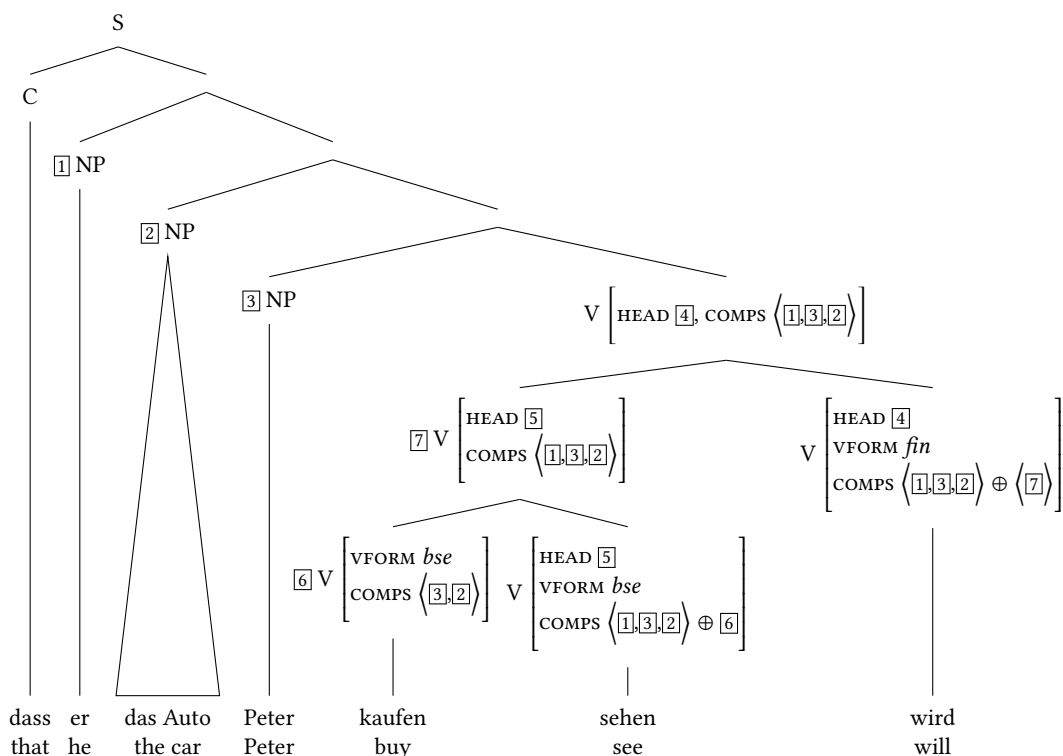
S

C

[1] NP

[2] NP

[3] NP

V [HEAD [4], COMPS ⟨[1],[3],[2]⟩]

[7] V [HEAD [5], COMPS ⟨[1],[3],[2]⟩]

V [HEAD [4], VFORM *fin*, COMPS ⟨[1],[3],[2]⟩ ⊕ ⟨[7]⟩]

[6] V [VFORM *bse*, COMPS ⟨[3],[2]⟩]

V [HEAD [5], VFORM *bse*, COMPS ⟨[1],[3],[2]⟩ ⊕ [6]]

dass	er	das Auto	Peter	kaufen	sehen	wird
that	he	the car	Peter	buy	see	will

Figure 12

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), *immer* 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 true   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 com-

plement. Contrary to a construction with an incoherent verb, this AP cannot be extraposed (46b), or pied piped with a relative pronoun (46d) (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.

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

       d.   * der Sohn, auf den   stolz Karl gewesen ist
            the son   on whom proud Karl been   is

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:

(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:[15]

$$
(48) \quad
\begin{bmatrix}
sein \\
\text{HEAD } verb \\
\text{COMPS } \boxed{1} \oplus \left\langle \begin{bmatrix} \text{HEAD } [\text{PRD } +] \\ \text{COMPS } \boxed{1} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

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.

---

[15]Müller (2013) generalizes the description of the copula to English and Danish. Given that list $\boxed{1}$ can be empty, the description is also valid for the copula taking a saturated complement (an NP or a PP complement in German, as well as for English and Danish, where the copula does not enter a complex predicate). However, NP and PP complements of the copula are independently specified as complement saturated in German, and more generally in English and Danish.

4.2 Argument attraction with Korean auxiliaries and control verbs

4.2.1 Scrambling in Korean

Korean resembles German in that a complex predicate signals itself mainly by its word order properties. The examples as well as the analysis are from Chung (1998) (see also Sells 1991; Yoo 2003; Kim 2016).

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'. Scrambling with auxiliaries is illustrated in (49). There is no evidence of scrambling in (49a): the subject *Maryka* starts the sentence, and the complement of the verb *ilkko* immediately precedes it. However, in (49b), the subject of the head verb *issta* occurs between the complement of *ilkko* and the verb *ilkko* itself.

(49) 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

Control verbs such as *seltukha-* 'to persuade', *cisi-* 'to order' (object control), *yaksokha-* 'to promise', *sito-* 'to try' (subject control), as well as causative *ha-*, allow complements of the embedded verb to be interleaved with their own arguments: the subject of the head verb *Maryka* occurs between the complement of the embedded verb and this verb in (50b) (50d).

(50) a. Mary-ka   John-hantley [ku chayk-ul ilk-ula-ko]
             Mary-NOM John-DAT   the book-ACC read-PART-PART
             seltukha-yss-ta.
             persuade-PAST-DECL

          'Mary persuaded John to read the book'

    b. Ku chayk-ul Mary-ka   John-hantley ilk-ula-ko
             the book-ACC Mary-NOM John-DAT   read-PART-PART
             seltukha-yss-ta.
             persuade-PAST-DECL

    c. Mary-ka   John-hantley [ku chayk-ul pilyecwu-kess-tako]
             Mary-NOM John-DAT   the book-ACC lend-PART-PART

yaksokha-yss-ta.
promise−PAST-DECL

'Mary promised John to lend the book'

d. Ku chayk-ul Mary-ka John-hantley pilyecwu-kess-tako
the book-ACC Mary-NOM John-DAT lend-PART-PART
yaksokha-yss-ta.
promise−PAST-DECL

A priori, these data could be explained in two ways: either the auxiliary or the control verb 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 or the control verb. Assuming a flat structure for the Korean sentence, as does Chung (1998)[16] the head-subject-complements-phrase is as in (51), the subject occurs at the same level as the complements, and they can be reordered.

(51)   *Head-subject-complements-phrase* (Korean) ⇒

$$
\begin{bmatrix}
\text{MOTHER} \mid \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{LIGHT} - \\ \text{SUBJ} / \langle \rangle \\ \text{COMPS} / \langle \rangle \end{bmatrix} \end{bmatrix} \\
\text{HEAD-DTR} \mid \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{LIGHT} + \\ \text{SUBJ} \; \boxed{1} \\ \text{COMPS} \; \boxed{2} \end{bmatrix} \end{bmatrix} \\
\text{NON-HEAD-DTRS} & (\boxed{1} \bigcirc \boxed{2}) \; \textit{non empty list}
\end{bmatrix}
$$

However, the same argument can be levelled against an analysis which appeals to linearization, as above for German (Section 4.1.2): so-called long passivization is possible, which cannot be accounted for by appeal to linearization and domain union.

Certain auxiliary and control verbs can be passivized, so that the expected complement of the embedded verb becomes the subject of the construction. (52a) exemplifies the active sentence, and (52b), the passive one. In (52a), *malsseng-manhun solul* ('the troublesome cow') is the complement of the embedded verb *phala*. In (52b), *malssengmanhun soka* is the subject of the passivized head verb *chiwe ciessta*. In the same way, while *ku cengchaykul* is the complement of the embedded verb *sihaynghalako* in (53a), *ku cengchayki* is the subject of the head passive verb *cisitoyessta* in (53b).

---

[16]However, see Kim (2016) for a binary branching structure.

(52) a. Ku mongpwu-ka malssengmanhun so-lul   phal-a
          The farmer-NOM   troublesome   cow-ACC sell-PART
          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)'

(53) a. Nay-ka Mary-hantley ku cengchayk-ul sihayngha-lako
          I-NOM Mary-DAT   the policy-ACC   carry.out-PART
          cisiha-yss-ta.
          order-PAST-DECL
          'I ordered Mary to carry out the policy'

      b. Ku cengchayk-i naey-uyhayse Mary-hantley sihayngha-lako
          The policy-NOM I-by   Mary-DAT   carry.out-PART
          cisitoy-ess-ta.
          be.ordered-PAST-DECL
          (lit. the policy was ordered by me for Mary to carry out)
          'Mary was ordered by me to carry out the policy'

Since passivization only affects the complement of the verb which is itself pas-sivized, it follows that *malssengmanhun solul* is the complement of the auxiliary in (52a) (*chiwesta*), and *ku cengchaykul* the complement of the control verb in (53a) (*cisihayssta*).[17]

4.2.2 The Structures of Complex Predicates in Korean

As shown by Chung (1998), in Korean, the structure of a complex predicate is different depending on whether the head verb is an auxiliary or a control verb. Auxiliaries are the head of a verbal complex (as in German complex predicates), while control verbs can either take a saturated VP complement, or be the head of a flat structure (like restructuring verbs in Italian). Thus, complex predicates

---

[17]In addition, Yoo (2003) shows that some variation regarding the case of the attracted comple-ment can be explained if the auxiliary and the verbal complement form a verbal complex.

Draft of 8th March 2020, 11:25

in Korean confirm the observation made in Romance Languages, although the complex predicate manifests itself by different properties: argument attraction is not correlated with a specific structure.

A characteristic property of Korean auxiliary constructions is that nothing can separate the two verbs: no parenthetical expression, such as *hayekan*, can intervene (54) (examples from Chung 1998). In addition, the embedded verb cannot occur by itself. Thus, while an NP may occur after the head verb in a construction called afterthought (55a), this is not possible for the embedded verb *mekko* alone (55b) or with its complement (55c). This behavior follows if the two verbs form a verbal complex (see Section 3.2).

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

      b. * Mary-ka sakwa-lul mek-ko hayekan issta.

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

      c. * Mary-ka   iss-ta,   sakwa-lul mek-ko.
          Mary-NOM be.in.the.process.of-DECL apple-ACC eat-PART

The Korean verbal-complex-phrase is shown in (56). This structure characterizes the class of auxiliaries. This is captured by the head feature [AUX +], which differentiates auxiliaries from control verbs, which are [AUX-]. Auxiliaries attract both the subject (list ①) and the complements of their verbal complement (list ③). Note the use of '○', the shuffle operator, which allows reordering of the elements of the two lists.

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

$$
\begin{bmatrix}
\text{MOTHER} \mid \text{SYNSEM} \mid \text{CAT} \begin{bmatrix} \text{HEAD} \begin{bmatrix} verb \\ \text{AUX} + \end{bmatrix} \\ \text{LIGHT} + \\ \text{SUBJ} \langle \boxed{1} \rangle \\ \text{COMPS} \boxed{3} \end{bmatrix} \\
\text{HEAD-DTR} \mid \text{SYNSEM} \mid \text{CAT} \begin{bmatrix} \text{AUX} + \\ \text{LIGHT} + \\ \text{SUBJ} \langle \boxed{1} \rangle \\ \text{COMPS} \langle \boxed{2} \rangle \bigcirc \boxed{3} \end{bmatrix} \\
\text{NON-HEAD-DTRS} \left\langle \boxed{2} \begin{bmatrix} verb \\ \text{LIGHT} + \\ \text{SUBJ} \langle \boxed{1} \rangle \\ \text{COMPS} \boxed{3} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

The auxiliary *iss-* in (54a), (55a) takes as its complement a light verb which is constrained to end in *−ko* (different auxiliaries put different restrictions on their verbal complement, Yoo 2003), from which it attracts all arguments.

We follow Chung (1998) in assuming that the sentence in Korean has a flat structure, which is allowed by the Head-subject-complements-phrase in (51). Sentence (49a) is represented in Figure 13.

Nothing prevents the verbal complex from embedding a verbal complex: it suffices that the complement of the auxiliary be an auxiliary itself. This is illustrated in (57).

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

     b. Ku chayk-ul Mary-ka   ilk-e   po-ko
          the book-ACC Mary-NOM read-PART try-PART
          iss-ta.
          be.in.the.process.of-DECL

The structure of (57a), with a series of two auxiliaries, is represented in Figure 14. *Issta* takes as its complement the verbal complex *ilke poko*, whose head is *poko*. *Poko*, being an auxiliary like *issta*, takes as its complement the verb *ilke*, attracting its subject and complements, which are transmitted to the verbal complex *ilke poko*; *ilke poko* saturates the verbal complement expected by *issta*, and transmits the subject and complements to the head auxiliary (see (56)).
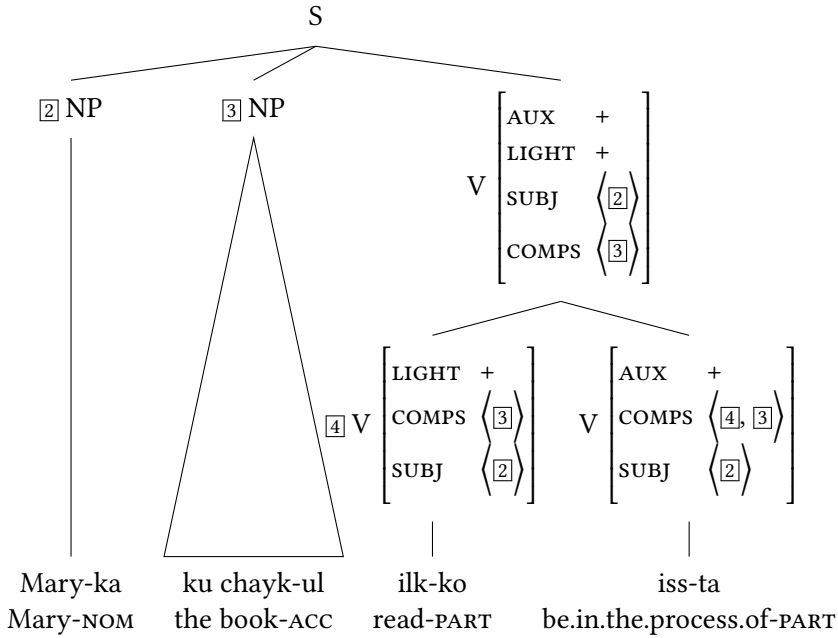
Figure 13

Contrary to auxiliaries, control verbs such as *seltukha-* (50a), (50b) or *cisi-* (53) can be separated from their verbal complement, for instance by an adverb as in (58a). They also allow for the infinitive and its complement to form a VP constituent as in (58b), where it is an afterthought. Thus, control verbs are analyzed in the same way as Italian restructuring verbs: they either take a saturated VP complement (58a) (58b), or are the head of a complex predicate (58c) (58d). They contrast with Korean Raising verbs which only take a VP complement.

(58) a. Mary-ka   ku mwuncey-lul phwulye-ko (kkuncilkikey)
          Mary-NOM the problem-ACC solve-PART   ceaselessly
          sito-hayss-ta.
          try-PAST-DECL
          'Mary tried (ceaselessly) to solve the problem'

      b. Mary-ka   sito-hayess-ta, [ku mwuncey-lul phwulye-ko].
          Mary-NOM try-PART-DECL the problem-ACC solve-PART

S

☐1 NP   ☐2 NP

V
$$\begin{bmatrix} \text{AUX} & + \\ \text{LIGHT} & + \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix}$$

☐4 V
$$\begin{bmatrix} \text{LIGHT} & + \\ \text{AUX} & + \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix}$$

V
$$\begin{bmatrix} \text{LIGHT} & + \\ \text{AUX} & + \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{4}, \boxed{2} \rangle \end{bmatrix}$$

☐3 V
$$\begin{bmatrix} \text{LIGHT} & + \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix}$$

V
$$\begin{bmatrix} \text{LIGHT} & + \\ \text{AUX} & + \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{3}, \boxed{2} \rangle \end{bmatrix}$$

Mary-ka	ku chayk-ul	ilk-e	poko	issta
Mary-NOM	the book-ACC	read-PART	try-PART	be.in.the.process.of-PART
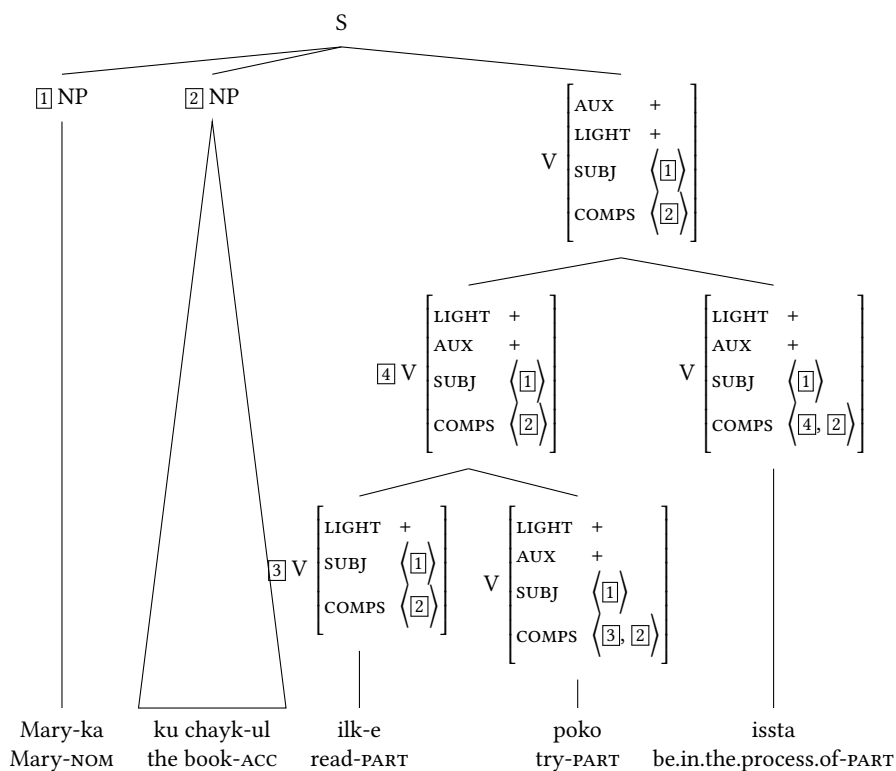
Figure 14

c. Ku chayk-ul Mary-ka ilku-lako sito-hayss-ta.
   The book-ACC Mary-NOM read-PART try-PAST-DECL

   'Mary tried to read the book'

d. Ku chayk-ul Mary-ka ilku-lako John-hantley
   The book-ACC Mary-NOM read-PART John-DAT
   seltukha-yss-ta.
   persuade-PAST-DECL

   'Mary persuaded John to read the book'

The scrambling data, together with the possibility of long passivization (53), show that there is a complex predicate. The subject of the head verb occurs between the complement of the infinitive and the infinitive in (58c), (58d). More precisely, there is no verbal complex in this case: the two verbs do not have to be contiguous, but can be separated, for instance by a complement as in (58d)

(John-hantley). Thus, they are the head of a flat structure as in Figure 15 corresponding to (58d). The head of verbal complexes in Korean is [AUX +], see (56). Since control verbs are [AUX −], they cannot enter verbal complexes, and the structure for complex predicates whose head is a control verb is a flat structure, corresponding to the Head-subject-complements-phrase in (51).

Figure 15

Ku chayk-ul	Mary-ka	ilku-lako	John-hantley	seltukha-yss-ta
The book-ACC	Mary-NOM	read-PART	John-DAT	persuade-PAST-DECL
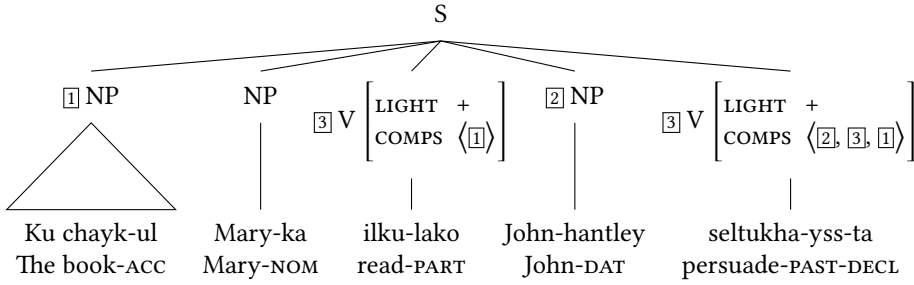
Control verbs which may be the head of a complex predicate are subject or object control verbs. They are the target of an argument attraction rule, in a way parallel to Italian Restructuring verbs.

(59)   Argument Attraction Rules for Korean control verbs

a.   Subject control verbs

$$
\begin{bmatrix}
\text{HEAD} & \begin{bmatrix} verb \\ \text{AUX-} \end{bmatrix} \\
\text{LIGHT} & + \\
\text{SUBJ} & \langle \text{NP}_i \rangle \\
\text{COMPS} & \left\langle V \begin{bmatrix} \text{SUBJ} & \langle \text{NP}_i \rangle \\ \text{COMPS} \langle\rangle \end{bmatrix} \right\rangle
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
\text{COMPS} & \left\langle V \begin{bmatrix} \text{LIGHT} & + \\ \text{COMPS} & \boxed{2} \end{bmatrix} \right\rangle \oplus \boxed{2}
\end{bmatrix}
$$

b.   Object control verbs

$$
\begin{bmatrix}
\text{HEAD} & \begin{bmatrix} verb \\ \text{AUX-} \end{bmatrix} \\
\text{LIGHT} & + \\
\text{SUBJ} & \langle \text{NP}_i \rangle \\
\text{COMPS} & \left\langle \text{NP}_j, V \begin{bmatrix} \text{SUBJ} & \langle \text{NP}_j \rangle \\ \text{COMPS} \langle\rangle \end{bmatrix} \right\rangle
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
\text{COMPS} & \left\langle \text{NP}_j, V \begin{bmatrix} \text{LIGHT} & + \\ \text{COMPS} & \boxed{2} \end{bmatrix} \right\rangle \oplus \boxed{2}
\end{bmatrix}
$$

An example with the verb *seltukha-* is given in Figure 15, which represents (58d).

The head comes last in Korean, except in the afterthought construction ([58b]), which requires an additional mechanism. This is true in the verbal complex, in the embedded VP and in the flat structure, as well. The alternative ordering for ([58a]) *Maryka phwulyeko ku mwunceylul sitohayssta*, where the verb *phwulye-ko* precedes its complement *ku mwuncey-lul* is thus impossible. In ([58d]), the NP *John-hantley* can follow the complement verb *ilkulako* because it is the complement of the head verb *seltukhayssta*, not of *ilkulako*.

Constraint ([60]) mirrors constraint ([28]) for Romance Languages.

(60) $\left[\text{SYNSEM }\boxed{1}\right]$ precedes $\left[\text{COMPS }\langle...\boxed{1}...\rangle\right]$ (Korean)

[Chung (1998)](#) extends the possibility of argument attraction to adjuncts, as well as to constructions with a S complement, although in the latter case the data are somewhat marginal. The behavior of adjuncts is easily accounted for, if adjuncts can be treated as complements ([Bouma et al. 2001](#)) by a verb. For the second case, Chung proposes a flat structure, in which the subject, the complements and the (lexical) verbs are all sisters. In this analysis, the definition of a complex predicate, which relies on a syntactic relation between words can then be maintained (but see [Lee 2001](#) for a different proposal based on linearization).

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 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 (61), 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 (61). Both the noun and the verb can be coordinated, as shown in (62) and (63) respectively. The noun can be extracted, as in the topicalization in (64), 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* in (65a)) becomes the subject of the passive construction (65b), as does the object of a transitive construction. The nominal part of the complex predicate is italicized in the examples.

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

(62) Mu-hā=āš=rā   [*boros yā šāne*] zad
      Hair-PL=3SG=RA brush or comb hit
      '(S)he brushed or combed her hair'

(63) Omid *sili* [zad va   xord]
      Omid slap hit   and strike
      'Omid gave and received slaps'

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

(65) 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 (66a), the verb *dādan* takes two complements, the noun

āb and the PP *[be bāqče]*, while in (66b) the combination of *dādan* and the noun *āb* takes a direct objet, which is marked with *-rā*: in (66b), the noun *āb* and the verb *dād* form a complex predicate.

(66)   a.   Maryam [be bāqče]   āb   dād
            Maryam to   garden water give.3sg
            'Maryam watered the garden'

       b.   Maryam [bāqče=rā]   *āb*   dād
            Maryam garden.ra   water give.3sg
            '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 (67); perfect participles can be converted into adjectives, and this applies to complex predicates (68) (see also Section 5.2).

(67)   a.   dust   dāštan   > dustdāštani
            friend have 'love'   lovely

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

(68)   a.   dast   xordan   > dastxorde
            hand strike 'to be sullied'   'sullied'

       b.   xat   xordan   > xatxorde
            scratch strike 'be scratched'   '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 (67), (68). We analyze here the case of agent nominalization, studied in Müller (2010).[18] The example he examines is especially interesting in that the nominalization does not exist independently of the complex predicate: as shown in (69), 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'.

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

The lexeme *bāz-konande* 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 (70) (Bonami & Crysmann 2018), where the elements of the compound are the value of the feature MDTRS (morphological daughters):[19]

$$(70) \quad \begin{bmatrix} lexeme \\ \text{PHON } \boxed{1} \oplus \boxed{2} \\ \text{SYNSEM | CAT | HEAD } noun \\ \text{MDTRS} \left\langle \begin{bmatrix} lexeme \\ \text{PHON } \boxed{1} \\ \text{SYNSEM | HEAD } noun \end{bmatrix}, \begin{bmatrix} lexeme \\ \text{PHON } \boxed{2} \\ \text{SYNSEM | HEAD } noun \end{bmatrix} \right\rangle \end{bmatrix}$$

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* is described in (71). 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').

---

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

(71)
$$
\begin{bmatrix}
\text{CAT } verb \\[4pt]
\text{ARG-ST } \left\langle \text{NP}_k,\ \text{NP}_j,\ Adj
\begin{bmatrix}
\text{PRD +} \\
\text{ARG-ST } \langle \text{NP}_j \rangle \\
\text{CONT } \boxed{1}
\begin{bmatrix}
open\text{-}relation \\
\text{THEME } j
\end{bmatrix}
\end{bmatrix}
\right\rangle \\[20pt]
\text{CONT}
\begin{bmatrix}
soa \\
\text{NUCLEUS}
\begin{bmatrix}
cause\text{-}relation \\
\text{CAUSER } k \\
\text{SOA-ARG } | \text{ NUCLEUS } \boxed{1}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
$$

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).

(72)
$$
\begin{bmatrix}
lexeme \\
\text{PHON } \boxed{1} \oplus \boxed{2} \oplus \langle ande \rangle \\[4pt]
\text{SYNSEM}
\begin{bmatrix}
\text{CAT}
\begin{bmatrix}
\text{HEAD } noun \\
\text{COMPS } \langle \boxed{3} \text{ NP} \rangle
\end{bmatrix} \\
\text{CONT } [\text{IND } k]
\end{bmatrix} \\[20pt]
\text{MDTRS } \left\langle
\begin{bmatrix}
\text{PHON } \boxed{1} \\
\text{SS } \boxed{4}
\begin{bmatrix}
\text{CAT } adjective \\
\text{ARG-ST } \langle \boxed{3} \text{ NP}_j \rangle \\
\text{CONT } \boxed{5}\ adj\text{-}rel\ (j)
\end{bmatrix}
\end{bmatrix},
\begin{bmatrix}
\text{PHON } \boxed{2} \\
\text{SS}
\begin{bmatrix}
\text{CAT } verb \\
\text{ARG-ST } \langle \text{NP}_k, \boxed{3}, \boxed{4} \rangle \\
\text{CONT } | \text{ NUCL}
\begin{bmatrix}
cause\text{-}rel \\
\text{CAUSER } k \\
\text{SOA-ARG } | \text{ NUCL } \boxed{5}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\right\rangle
\end{bmatrix}
$$

The agent noun keeps as a complement the NP expected by the verb, as illustrated in (73).

(73) [dar-e quti] bāz-konande
      lid-EZ box   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-andēr*) exist independently

of the elements of the complex predicate. Hence, the data in (69a) 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 have two realizations, 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* 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* (kick hit) means 'kick', and *sili zadan* (slap hit) means 'slap'.

(74) Olāq   be Omid lagad zad
      Donkey to Omid kick   hit

      'The donkey kicked Omid'

Within a hierarchical organization of the lexicon (chapter ?????), 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, if it unifies with it. 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…*), and its content is then unified with that of the PP. The result is a description of an *soa* with more information than is present in the verb alone. Similarly, here, the content of *lagad* is more specialized than that of *zadan*; hence, it takes over if it is unified with it. The complement of the complex predicate may be an NP or a PP headed by *be* (the preposition is optional).

(75)
$$
\begin{bmatrix}
\textit{zadan1-lxm} \\
\text{CAT } \textit{verb} \\
\text{ARG-ST} \left\langle \text{NP}_k,\ \textit{(be)}\ \text{NP}_m,\ \text{N} \begin{bmatrix} \text{CAT } [\text{PRD } +] \\ \text{CONT } \boxed{1} \end{bmatrix} \right\rangle \\
\text{CONT} \begin{bmatrix} \textit{soa} \\ \text{NUCLEUS } \boxed{1} \begin{bmatrix} \textit{kick-relation} \\ \text{AGENT } k \\ \text{UNDERGOER } m \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$

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* (knife hit) 'stab', *telefon zadan* (phone hit) 'phone', *piāno zadan* (piano hit) 'play the piano'. We illustrate here *šāne zadan* (comb hit) 'comb'.

(76) Maryam mu-hā=yaš=rā *šāne* zad
      Maryam hair-PL=3SG=RA comb hit

      'Maryam combed her hair'

(77)
$$
\begin{bmatrix}
\textit{zadan2-lexeme} \\
\text{SYNSEM} \begin{bmatrix}
\text{CAT } \textit{verb} \\
\text{ARG-ST} \left\langle \text{NP}_k,\ \text{NP}_m,\ \text{N} \begin{bmatrix} \text{CAT } [\text{PRD } +] \\ \text{CONT } \boxed{2} \end{bmatrix} \right\rangle \\
\text{CONT } \boxed{2} \begin{bmatrix} \textit{soa} \\ \text{SIT } \boxed{1} \\ \text{NUCLEUS} \begin{bmatrix} \textit{comb-relation} \\ \text{AGENT } k \\ \text{UNDERGOER } m \end{bmatrix} \end{bmatrix} \\
\text{BACKGROUND } \big\{ \textbf{involves}\,(\boxed{1}, \exists\, x\, [\textbf{comb}\,(x) \wedge \textbf{use}\,(\boxed{1}, k, x)]) \big\}
\end{bmatrix}
\end{bmatrix}
$$

Although the complex predicate takes on the content of the predicative complement, the semantics does not rely on unification 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.

Further from a compositional or recoverable meaning is the use of *zadan* or more precisely *xod=rā zadan* (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.

(78) 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* (deceit hit) 'to deceive': as in (75), 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 (78) 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.

(79)
$$
\begin{bmatrix}
\textit{zadan3-lexeme} \\
\text{CAT } \textit{verb} \\
\text{ARG-ST} \left\langle \text{NP}_k, \text{pro}_k, \text{PP} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{PFORM BE} \\ \text{PRD } + \end{bmatrix} \\ \text{CONT } \boxed{1} \mid \text{NUCLEUS} \begin{bmatrix} \textit{internal-problematic-state} \\ \text{EXPERIENCER } k \end{bmatrix} \end{bmatrix} \right\rangle \\
\text{CONT} \begin{bmatrix} \textit{soa} \\ \text{NUCLEUS} \begin{bmatrix} \textit{pretend-relation} \\ \text{AGENT } k \\ \text{THEME } \boxed{1} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$

Note that, contrary to *zadan1-lexeme*, with which *be* is optional, the *zadan3-lexeme* requires the 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 $\boxed{1}$]); this is indicated by treating *be* as the value of the feature PREP(OSITION) FORM (Pollard & Sag 1987: Chapter 3).

Finally, we turn to an idiom: *dast zadan* (hand hit) meaning 'start'. The combination may mean, in a more recoverable way 'to touch' with PP complements denoting concrete objects (as in (61)), or 'to applaud' with a PP complement denoting a person (80a). However, it means 'to start' with a PP complement denoting an event as in (80b).

(80) 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-an
    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 (81). 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.

(81)
$$
\begin{bmatrix}
\textit{zadan4-lexeme} \\
\text{CAT } \textit{verb} \\
\text{ARG-ST} \left\langle \text{NP}_k, \text{PP} \begin{bmatrix} \text{PFORM BE} \\ \text{CONT } \boxed{1} \end{bmatrix}, \text{N} \begin{bmatrix} \text{CAT } [\text{PRD +}] \\ \text{LID } \textit{dast} \end{bmatrix} \right\rangle \\
\text{CONT} \begin{bmatrix} \textit{soa} \\ \text{NUCLEUS} \begin{bmatrix} \textit{start-relation} \\ \text{AGENT } k \\ \text{SOA-ARG } \boxed{1} \mid \text{NUCLEUS } \textit{event-relation} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$

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 contrasts with Italian), but they can also be present within the same language (Korean auxiliaries contrast with Korean control verbs).

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 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

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

References

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 Miller & Ivan A Sag. 1998. French bounded dependencies. In S. Balari & L. Dini (eds.), *Romance in HPSG*, 1–54. Stanford: CSLI Publications.

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: CSLI Publications.

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

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*, 175–202. Berlin: Language Science Press.

Draft of 8th March 2020, 11:25

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.

Bouma, Gosse, Robert Malouf & Ivan A Sag. 2001. Satisfying constraints on extraction andadjunction. *Natural Language & Linguistic Theory* 19(1). 1–65.

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.

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: An extension of the complex predicate analysis. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 159–220. San Diego: Academic Press.

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

Geach, Peter. 1970. A program for syntax. *Synthèse* 22. 3–17. Reprint in Donald Davidson & Gilbert Harman (eds.), 1972, *Semantics of Natural Language*, 483–497, Springer.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative Investigations.* Stanford: CSLI Publications.

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

Hale, Ken & Samuel Jay Keyser. 1996. On the complex nature of simple predicators. In Alsina Alex, Joan Bresnan & Peter Sells (eds.), *Complex Predicates*, 29–65. Stanford: CSLI Publications.

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

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989. Flipped out: AUX in German. *Chicago Linguistic Society* 24. 193–202.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl J. 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 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.* Oxford: Oxford University Press.

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

Kiss, Tibor. 1994. Obligatory coherence: The structure of German modal verb constructions. In John Nerbonne, Klaus Netter & Carl J. 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.* Tübingen: Niemeyer.

Koenig, Jean-Pierre. 1999. *Lexical Relations.* Stanford: 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 dissertation.

Meurers, Detmar. 2000. *Lexical generalizations in the syntax of German non-finite constructions*. Arbeitspapiere des SFB 340 145. Tübingen: Eberhard-Karls-Universität.

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.

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.

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

Müller, Stefan. 2013. On the Copula, Specificational Constructions and Type Shifting. Freie Universität Berlin. Ms.

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

Orgun, Cemil Orhan. 1996. *Sign-Based Morphology and Phonology with Special Attention to Optimality Theory*. University of California Berkeley dissertation.

Pollard, Carl & Ivan A. Sag. 1987. Information-based syntax and semantics. *CSLI lecture notes* 13.

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 J. 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, accessed 2018-2-25.

Riehemann, Susanne Z. 1998. Type-based Derivational Morphology. *The Journal of Comparative Germanic Linguistics* 2. 49–77.

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/, accessed 2018-2-25.

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*. Stanford: CSLI Publications. 2nd edition.

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 Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Syntax and seman-*

tics vol. 30: complex predicates in non-derivational syntax, 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.