Chapter 12

Complex predicates

Danièle Godard
Université Paris Diderot

Pollet Samvelian
Université Sorbonne Nouvelle

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

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 (CP) is composed of two or more words, which are predicates, of which one is a verb, head of the construction, and the other belongs to diverse categories: non-finite verb, noun, adjective, preposition. Both predicates merge their argument structures: the head inherits the arguments of its complement.

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 mono-clausality can manifest itself differently in different languages (Butt (2010) , *pace* Haspelmath (2016)). In the case of Romance auxiliary constructions, the first verb (the auxiliary) hosts the clitics which pronominalize the arguments of the main predicate: corresponding to the NP complement in (1a) the pronominal clitic *l(e)* is hosted by the auxiliary *a* (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.

(2)   a.   Paul veut lire son livre.

           'Paul wants to read his book.'

   b.   Paul veut le lire.

           Paul wants to it-read

   c.   * Paul le veut lire.[1]

Complex predicates are sometimes given a semantic definition: the two elements constitute one predicate which describes one situation. However, a semantic definition is not sufficient. It is true that the head verb 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 (3a) or not (3b).

(3)   a.   Anna lo vuole comprare. (Italian, Monachesi 1998)
           Anna it wants to-buy
           'Anna wants to-buy it.'

---

[1]Possible in an earlier stage of French.

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

In other cases, the term is used to describe the complex content of a word. For instance, the verb *dance* has been analyzed as incorporating the noun *dance* and considered a complex predicate (Hale & Keyser 1997). In the sense adopted here, complex predicates involve at least two words.

It is more difficult to distinguish CPs from *Serial verb constructions* (SVC) which are composed of verbs which belong to the same clause (4). However, following Haspelmath (2016), there are a number of differences: (i) SVC are sequences of verbs, and cannot involve various categories, contrary to CP. (ii) These verbs are independent of each other: they have the same form and meaning when they occur outside the SVC, and they do not involve predicate argument relationship between them, in particular infinitival complements and causative constructions are excluded (see 3a). (iii) Finally, they tend to each describe a subevent, being frequently directionals or resultatives (Prince 2017), while a tense auxiliary as in the French example in (1) is not associated with a subevent.

(4) Òzó sàán rrá ógbà.
    Ozo jump cross fence
    'Ozo jumped over the fence.' (edo, Haspelmath 2016)

1.2 Constructions based on complex predicates

Complex predicates enter into a number of constructions across languages. The following have been particularly studied in HPSG:

- Tense auxiliaries, the copulas and predicative verbs, restructuring verbs, and certain causative and perception verbs in Romance languages (Abeillé & Godard 2002; 2005; 2010; Abeillé et al. 1998; Monachesi 1998);

- Coherent constructions in German and Dutch, headed by tense auxiliaries, copulas, certain raising and control verbs, certain verbs with predicative complements, and particle verbs (De Kuthy & Meurers 2001; Hinrichs & Nakazawa 1994; 1998; Kathol 1998; Kiss 1994; Meurers 2002; Müller 2002; Bouma & van Noord 1998; Rentier 1994);

- Korean auxiliaries, control verbs and causative verb *ha* (Chung 1998; Kim 2016; Sells 1991; Yoo 2003);

- 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; Samvelian & Faghiri 2016) and Korean (Choi & Wechsler 2002; Kim 2016; Lee 2001; Ryu 1993).

In this chapter, we examine some of these constructions which illustrate different aspects of the phenomenon of complex predicates.

2 The basic mechanism in HPSG: Argument inheritance

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

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 a general way (1b), (1c). Moreover, it can be gapped alone (5d), while the participle can only be gapped with the auxiliary (5e), (5f); this is expected if the auxiliary is the head, since it behaves like *pense* in (5a) and (5b), while the participle behaves like the infinitive.

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

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

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

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

c. * Lola pense cueillir des pommes et Alice pense des pêches.

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

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

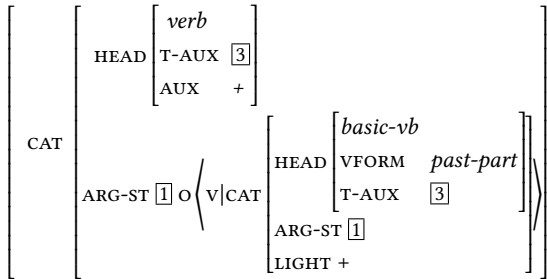'Lola has bought apples, and Alice (has) bought picked peaches.'

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

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

    f. * Lola a acheté des pommes, et Alice a des pêches.

        Lola has bought apples, and Alice has peaches

The auxiliary construction in French is a CP: The clitic corresponding to a complement of the participle is hosted by the auxiliary (it is said to "climb"), as in (1); moreover, it occurs in bounded dependencies such as the infinitival complement of adjectives like *facile* 'easy', whose nominal complement is unexpressed (6a). This unexpressed complement can be that of a participle (6c), but not that of an embedded infinitive (6b).

(6)   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 promettre de maîtriser en un jour.

           This technique is impossible to promise 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

The syntactic description of the auxiliary *avoir* is given in (7):[2]

(7)   Lexeme of the French tense auxiliary *avoir*

$$
\left[\text{CAT} \begin{bmatrix} \text{HEAD} \begin{bmatrix} verb \\ \text{T-AUX} \quad \boxed{3} \\ \text{AUX} \quad + \end{bmatrix} \\ \text{ARG-ST } \boxed{1} \, \bigcirc \left\langle \text{V} \middle| \text{CAT} \begin{bmatrix} \text{HEAD} \begin{bmatrix} basic\text{-}vb \\ \text{VFORM} \quad past\text{-}part \\ \text{T-AUX} \quad \boxed{3} \end{bmatrix} \\ \text{ARG-ST } \boxed{1} \\ \text{LIGHT } + \end{bmatrix} \right\rangle \end{bmatrix} \right]
$$

The auxiliary takes the participle as its complement, from which it inherits its syntactic arguments. The value for ARG-ST uses the shuffle operator O, which allows the elements of the lists to be reordered (Reape 1994). Thus, it is the auxiliary itself, as head of the construction, which is responsible for argument inheritance:

---

[2]The feature [AUX ±] indicates whether the infinitival can occur before the negative adverb (*pas*, *jamais*) or not. The feature T-AUX ensures that the appropriate auxiliary (*avoir* or *être*) combines with a participle: it has the same value for the auxiliary and the participle, and is defined lexically.

the phenomenon is lexically driven. These arguments are realized as indicated in (8) (see Ginzburg & Sag 2000: for the Argument Realization Principle). The argument structure contains canonical and non-canonical elements, but the valence only contains canonical ones. The valents in turn are realized are the subject and the complements (and, possibly, the specifier for nouns).

(8) Argument Realization (Bonami & Samvelian 2015)

$$word \Rightarrow \left[\text{SS|LOC|CAT} \left[\begin{array}{l} \text{ARG-ST } \boxed{1} \circ list\big(non\text{-}canon\big) \\ \text{VAL } \boxed{1} \; list\big(canon\big) \end{array} \right] \right]$$

The arguments are *synsems*, which can have different subtypes as in (9). Usually, they are not specified on the lexeme description, but they are on words. Thus, (8) says that if an argument is non-canonical, it stays on the argument list, but is not realized as a valent.

(9) *synsem* hierarchy:

synsem
non-canon   canon
aff   gap   null-pro
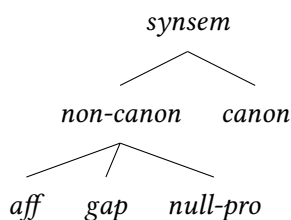
In (1a), the participle *lu* selects the argument *son livre*, which is inherited by the auxiliary *a*. Accordingly, it is realized as the complement of the auxiliary *a* in (1a). The structure is as in Figure 1.

Romance clitics are analyzed as affixes (Miller & Sag 1997). As arguments, they are inherited by the auxiliary. A morphological function may apply to verbs containing affixes in their argument list, including auxiliaries, which induces a certain morphology: a value of FORM (final form) different from that of I-FORM (inflected form). Hence the description of *l'a* in *l'a lu* differs from that of *a* in *a lu* (10) (Abeillé et al. 1998).

Add example with glosses and translation, add glosses in tree

Figure 1: Add caption

(10)   Description of *l'a* in *l'a lu*:

$$
\begin{bmatrix}
\text{MORPH} \begin{bmatrix} \text{FORM} & l'a \\ \text{I-FORM} & a \\ \text{STEM} & av\text{-} \end{bmatrix} \\[3em]
\text{SS|CAT} \begin{bmatrix}
\text{HEAD} \begin{bmatrix} reduced\text{-}verb \\ \text{VFORM} & indic \\ \text{T-AUX} & \boxed{5} \end{bmatrix} \\
\text{SUBJ} \langle \boxed{1} \rangle \\
\text{COMPS} \langle \boxed{2} \rangle \oplus \boxed{4} \\
\text{ARG-ST} \left\langle \boxed{1}, \boxed{2} \text{V} \left[\text{CAT} \begin{bmatrix} \text{LIGHT} + \\ \text{HEAD} \begin{bmatrix} basic\text{-}verb \\ \text{FORM} & past\text{-}part \\ \text{T-AUX} & \boxed{5} \end{bmatrix} \\ \text{ARG-ST} \langle \boxed{1}, \boxed{3}[\mathit{aff}] \rangle \oplus \boxed{4} \end{bmatrix} \right], \boxed{3} \right\rangle \oplus \boxed{4}
\end{bmatrix}
\end{bmatrix}
$$

In (10) , the complements of the auxiliary are the participle ②, and the list ④; the argument list contains in addition ③, an affix, 3rd person, masculine singular. In the sequence *l'a lu*, as in (1b), [4] is the empty list; in the sequence *l'a lu à son père* 'has read it to his father', ④ is the PP complement.

This description is sufficient to ensure that the clitic is always realized on the auxiliary in French, and not on the participle: *Paul l'a lu* contrasts with **Paul a le-lu*. In French, past participles never host clitics, which we assume to be a morphological property. But, in Italian, past participles may host clitics, although

never when they combine with the auxiliary. To account for this, we distinguish between basic verbs, with no cliticized argument, and reduced verbs, with at least one cliticized argument which belongs to the argument list, but not to the COMPS list. The verb complement of *avoir* is specified as *basic-verb* (see (7), (10)). We distinguish between *basic-verbs* and *reduced-verbs*. Basic verbs can have arguments typed as affixes, but not cliticized, while reduced verbs have at least one cliticized argument: present in the argument list, but not in the complement list. The complement of *avoir* is specified as *basic-verb* (see (7), (10)).

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

In addition to tense auxiliaries, Romance languages have Restructuring verbs as heads of complex predicates. A comparison of their properties illustrates an important aspect of the phenomenon: argument inheritance is compatible with different syntactic structures, in this case, a flat structure, and a verbal complex (Abeillé & Godard 2010).

3.1 Romance Restructuring verbs

Certain verbs in Romance Languages, called *Restructuring verbs*, exhibit two behaviors: either as ordinary verbs taking a VP complement or as complex predicates (Aissen & Perlmutter 1983; Rizzi 1982). They are: modal verbs, aspectuals, or movement verbs. However, it must be kept in mind that this behavior is lexical: verbs which are close semantically may not be heads of CPs.

Several properties show that they head CPs. First, clitic climbing, which is optional (while it is obligatory with tense auxiliaries):

gloss and translate everything

(11) a. Giovanni *le* vuole mangiare.
          Giovanni vuole mangiar*le*. (Italian)

      b. Juan *las* quiere comer.
          Juan quiere comer*las*. (Spanish)

      c. O Joãn quere-*as* comer.
          O Joãn quer comê-*las*. (Portuguese)

      d. En Joan *les* vol menjar.
          En Joan vol menjar-*les*. (Catalan)

          'John want to eat them.' (Abeillé & Godard 2010)

Second, the medio-passive or middle *si* construction, where the verb hosts the reflexive clitic *si* or *se* and the subject corresponds to the object of the active construction (12a), with an interpretation close to that of passive (12b). The construction is possible with restructuring verbs (12c), but not with ordinary verbs taking an infinitival complement (12d).

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

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

c. Queste camicie si possono stirare facilmente.
'These shirts can be ironed easily.'

d. * Queste camicie si paiono stirare facilmente.
These shirts appear to be ironed easily.'

Third, their acceptability in bounded dependencies, as illustrated in (6) for tense auxiliaries, and (13) for Restructuring verbs. (13b) relies on *cominciare* being a Restructuring verb, while *promettere* is not (13c).

(13) a. Questa canzone è facile da apprendere.
'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
'This song is easy to promise to learn.'

Finally, the possibility of preposing the verbal complement of a Restructuring verb disappears when there is evidence of a complex predicate. For the sake of simplification, we now concentrate on Italian and Spanish. The data in (14), with a VP, contrast with those in (15), where the head verb inherits a cliticized complement of the infinitival. Preposing of the verbal complement is associated with pronominalization (*lo*) in Italian (14a) not in Spanish (14b), and is more natural in contrastive contexts in Spanish. The infinitive cannot be preposed if its argument is cliticized onto the head verb (15).

(14) Does he want to talk to Mary?

a. – Parlare a Maria, certamente lo vuole.

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

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

(15) a. *Parlare, certamente glielo vuole.

b. *Hablar, le quiere (pero no mucho tempo).

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 Composition Lexical Rule (16) (Monachesi 1998).

(16) Argument Composition Lexical Rule for Romance restructuring verbs:

$$
\begin{bmatrix}
\text{HEAD } verb \\
\text{LIGHT } + \\
\text{SUBJ } \langle \text{NP} \begin{bmatrix} \text{IND} & i \end{bmatrix} \rangle \\
\text{COMPS } \begin{bmatrix} \text{HEAD} \begin{bmatrix} verb \\ \text{VFORM} & inf \end{bmatrix} \\ \text{SUBJ } \langle \text{NP} \begin{bmatrix} \text{IND} & i \end{bmatrix} \rangle \\ \text{COMPS } \langle \rangle \end{bmatrix}
\end{bmatrix}
\mapsto
\begin{bmatrix}
\text{COMPS } \langle \begin{bmatrix} basic\text{-}verb \\ \text{COMPS } \boxed{2} \end{bmatrix} \rangle \oplus \boxed{2}
\end{bmatrix}
$$

The left hand side represents a verb taking a VP complement. The identity of indices for the expected subjects of the verb and its infinitival complement indicates that they are control or raising verbs; the verbal complement is saturated for its complements. The right hand side represents the head of a complex predicate: it inherits the expected complement(s) of the infinitival complement (list $\boxed{2}$). The categorization of the infinitive as *basic verb* ensures that it does not host cliticized complements, and the feature [LIGHT +], shared with the input, that it has not combined with complements (see below).

The complement of adjectives such as 'easy' in RL is a bounded dependency: they take an unsaturated infinitival complement which expects a complement coindexed with its subject. Complex predicates can occur in this construction because they inherit the complement of their complement. In (13a), *apprendere* is expecting an object, co-indexed with *questa conzone*. In (13b), *cominciare* inherits the expected complement of *apprendere*, a configuration which is similar to that of (13a).

The medio-passive verb is also the result of a Lexical Rule, which takes a transitive verb like *stirare* (12a) to give a verb whose subject corresponds to the expected object of the input verb, and acquires a clitic SE (realized *si* or *se* in the different languages) as in (12b) (Monachesi 1998). While a verb taking a VP complement like Italian *potere* cannot be the input to this rule, since it lacks an NP

object, the corresponding verb which is the output of LR (16) can if it inherits a nominal complement from the infinitive, as in (12c). The verb *potere* as in *Giovanni può stirare queste camicie*, which inherits *queste camicie* from *stirare*, is the input to Rule (16), giving the verb which occurs in (12c).

3.2 The structure of CP constructions: Verbal complex and flat structure

Argument composition is compatible with different structures. In Romance Languages, complex predicates enter either a flat structure, or a verbal complex. We contrast Italian and Spanish.[3]

The impossibility of Preposing illustrated in (15) 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, regarding other properties, the two languages do not behave in the same way. Note that in Spanish, there is variation among speakers: we illustrate here Spanish-1 variety. 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 (17a), but not in Spanish-1 in a general way (17b) (a few adverbs, such as *casi* 'nearly', *ya* 'already', *apenas* 'barely' are possible).

(17) a. Giovanni *lo* vuole spesso leggere. (Italian)
      b. * Juan *lo* quiere a menudo leer. (Spanish-1)
      c. Juan quiere a menudo leer*lo*.
          'John wants to often read it.'

Second, an inverted subject NP can occur between the two verbs of a complex predicate in Italian, but not in Spanish-1. The subject can occur postverbally in interrogative sentences. In Italian, it can occur between the two verbs with a special prosody, and with speaker's variation (Salvi 2010) (18a). In Spanish-1, it is not possible (except for the pronominal subject) (Suñer 2010) (18b), while it can occur between the verb and an infinitival VP complement (18c).

---

[3]In Portuguese complex predicate constructions are also a flat structure, but with different ordering constraints from Italian; in Spanish-2, they are similar to Portuguese. Complex predicates with verbs which have only one structure also distribute between these two structures: tense auxiliaries in French, Italian, Portuguese as well as Romanian modal *a putea* 'can' are the head of a flat structure, while Spanish-1 and Romanian tense auxiliaries enter a verbal complex (Abeillé & Godard 2010).

(18)　　a.　% Lo comincia MARIA a capire, il problema, oppure no?

　　　　　　'Maria, she's beginning to understand it, the problem, yes or no.'

　　　　b.　*? Lo comienza Juan a comprender?

　　　　c.　Comienza Juan a comprenderlo?

　　　　　　'Is John beginning to understand it?'

Finally, Italian heads of complex predicates can have scope over the coordination of infinitives with their complements (19a), while this is not the case in Spanish-1 (19b) (from Abeillé & Godard 2010). Again, the presence of the clitic on the head verb (*lo vuole*, *le volvió*) shows that this is a complex predicate construction.

(19)　　a.　Giovanni lo vuole comprare subito　　　　e　　dare a　Maria
　　　　　　Giovanni it wants buy　　　immediately and give to Maria

　　　　　　'Giovanni wants to buy it immediately and give it to Maria.'

　　　　b. * Le　　　volvió　　　a　pedir un autografo y　　a　hacer
　　　　　　to-him started again to ask　　an autograph and to make
　　　　　　proposiciones
　　　　　　propositions

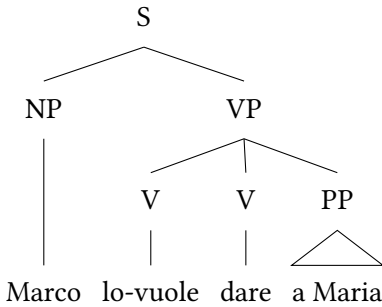　　　　　　'He started again to ask him for an autograph and to make propositions to him'

Constituency tests such as preposing (15) show that the verbal complement is not a VP. 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 non-verbal elements and adverbs (in general). This constraint can be captured by the feature [LIGHT +],[4] which has been used in Romance languages for other phenomena as well (Abeillé & Godard 2005; 2010). Hence, complex predicate constructions in Spanish-1 contain a verbal complex, while they represent a flat structure in Italian.

---

[4]The feature [LIGHT ±] is used rather than the feature WEIGHT as in Abeillé and Godard; it substitutes for [LEX ±] used in HPSG formalizations of German (Bonami & Webelhuth 2013). The adverbs admissible in the Spanish-1 verbal complex are light.
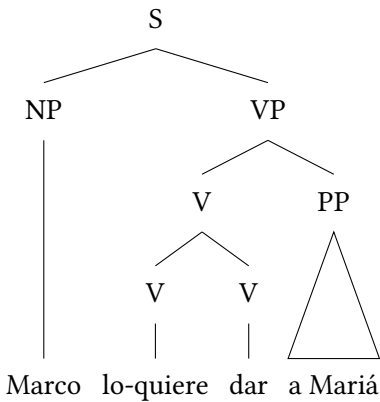
(20)   a. VP complement:

```
              S                              S
         ╱        ╲                     ╱        ╲
      NP            VP               NP            VP
       │         ╱      ╲             │         ╱      ╲
       │        V        VP           │        V        VP
       │        │      ╱───╲          │        │      ╱───╲
     Marco    vuole  lo-dare a Maria  Marco  quiere  darlo a Mariá
```

b. Flat structure:

```
              S
         ╱         ╲
      NP             VP
       │         ╱    │    ╲
       │        V     V     PP
       │        │     │    ╱──╲
     Marco   lo-vuole dare a Maria
```

c. Verbal complex:

```
              S
         ╱         ╲
      NP             VP
       │         ╱        ╲
       │        V          PP
       │      ╱   ╲       ╱──╲
       │     V     V     ╱    ╲
       │     │     │    ╱      ╲
     Marco lo-quiere dar  a Mariá
```

The possibility of the coordination in (19a) has been viewed as an argument in favor of a VP complement. However, if such sequences can be analyzed as co-ordinations of VP, they can also be Non Constituent Coordinations (NCC) (*John gives a book to Maria and discs to her brother*). So, the question is: why is (19b) not

an acceptable NCC in Spanish? We propose that 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 (21a) and (21b) in Spanish. In (21b), the constituant [*de Camus*] in the second conjunct is at the same level as [*el libro de Proust*], but not at the same level as [*de Proust*], with which it is parallel.

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

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

3.3 Analysis of Romance CP constructions in HPSG

The complex predicate in Italian is a flat structure (20b), while it is a verbal complex costruction in Spanish (20c). Abeillé & Godard (2005) do not propose a special phrase for dealing with verbal complexes, but treat them as a subtype of head-complements-phrases. The difference between flat structure and verbal complex is attributed to the feature [LIGHT ±] and linearization constraints. The head-complements-phrase is as follows:

(22)   a.   *Head-Complements-phrase* $\Rightarrow$

$$\begin{bmatrix} \text{MOTHER|SS|CAT} \begin{bmatrix} \text{HEAD } \boxed{1} \\ \text{COMPS } /\langle\rangle \end{bmatrix} \\ \text{HEAD-DTR|SS|CAT} \begin{bmatrix} \text{HEAD } \boxed{1} \\ \text{LIGHT } + \\ \text{COMPS } \boxed{2} \text{ o } \boxed{3} \end{bmatrix} \\ \text{NON-HEAD-DTRS } \boxed{2} \text{ } non\text{-}empty\text{ }list \end{bmatrix}$$

        b. $\begin{bmatrix} phrase \begin{bmatrix} \text{LIGHT } + \end{bmatrix} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{DTRS} \begin{bmatrix} \text{LIGHT } + \end{bmatrix} \end{bmatrix}$

The head-complements-phrase is usually saturated for the expected complements, but this constraint can be violated, notably in the case of a verbal complex. In this case , the infinitive combines with a head verb without being saturated, and the head verb is not saturated for the complements it inherits from the infinitive. The requirement is passed up because of the Generalized Head Feature Principle (Ginzburg & Sag 2000), which says that the *synsem* of the phrase is identical to that of the head daughter by default (each phrase description says in what way they differ).

The LIGHT feature is appropriate both for words and phrases; it has ordering as well as structural consequences. Words can be light or non-light; lexical

Add gloss and translation

verbs (finite verbs, participles or infinitives without complements) are light. Most phrases are non-light, but phrases made up of light elements can be light or non light; if the phrase is light, the daughters are light (22b). Accordingly, a verbal complex, made up of two lexical verbs can be light: this allows it to be the head of the VP (see (22a)).

In order to enforce the verbal complex in Spanish-1, we assume an additional constraint on the non-light head-complements-phrase in this language, which does not exist in Italian (or French).

(23) *Head-complements-phrase* [LIGHT −] ⟹ [NON-HEAD-DTRS list([LIGHT −]])
(Spanish)

Constraint (23) says that all the complements in a non-light phrase must be non-light. This precludes a light complement such as a (bare) infinitive from occurring at the same level as an NP or a PP, for instance, which are non-light, contrary to what happens in the flat structure in (20b).

In order to prevent a verbal complex in Italian (and French), we assume a different additional constraint (which differs from that in Abeillé & Godard (2005)).

(24) *Head-complements-phrase* [LIGHT +] ⟹ [[NON-HEAD-DTRS list([HEAD *non-verbal*]] (Italian)

The Verbal Complex has to be light, since it is the head of the VP (see (22b)). Since it is made of verbal constituents, it is excluded by constraint (24) in Italian.[5]

Romance languages follow the general constraints on ordering in head-complements-phrases in non-final languages. They are given here informally. Constraint (26) implies that the infinitival complement in complex predicates (or the participle for the auxiliary constructions), precedes the complements it subcategorizes for.[6]

(25) The head DTR precedes the complements (Romance languages)

(26) [V [COMPS < [1][ADV −]>]] precedes [1]. (Romance languages)

---

[5]Note that head-only phrases are non-light.

[6]The feature [ADV −] excludes adverbial complements from the ordering rule: they may precede the past participle as in *Paul s'est bien comporté* (Paul SE-is well behaved, 'Paul has behaved well').

4 Complex predicates and word order

In certain languages, a complex predicate construction signals itself by properties of word order. This is the case for instance in German (De Kuthy & Meurers 2001; Hinrichs & Nakazawa 1994; 1998; Kathol 1998; Kiss 1994; Meurers 2002; Müller 2002) and Dutch (Bouma & van Noord 1998; Rentier 1994), as well as Korean (Chung 1998; Kim 2016; Sells 1991; Yoo 2003).

4.1 Coherent and incoherent constructions in German

Verbs with an infinitival complement in German enter into so-called incoherent or coherent constructions: in the first case, they combine with a VP (saturated for its complements), but not in the second case (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, modals, subject and object raising, object control verbs, the copulas and predicative constructions (see Müller 2002).

Incoherent constructions are constructions where a verb takes a (saturated) VP complement. They are illustrated by the combination of *überreden* with the infinitival complement in (27). They allow for VP fronting as in (27a), VP extraposition as in (27b), and VP relatives such as (27c), where the infinitival complement is"pied-piped" with the relative pronoun (examples from Hinrichs & Nakazawa 1998).[7]

(27)   a.   [Das auto zu kaufen] wird Peter Maria *überreden.*
            to   buy the car   will Peter Maria persuade

            'Peter will persuade Maria to buy the car.'

       b.   … dass Peter Maria   *überredet,* [das Auto zu kaufen].
            that Peter persuade Marie,   the   car   to buy

            '… that Peter persuade Maria to buy the car.'

       c.   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 auxiliary *wird* or the raising verb *scheinen* with an infinitival complement are

---

[7]The verb taking the VP complement is underlined.

Draft of 7th November 2019, 11:22

typical examples, do not allow for VP extraposition (28b), (28d) (from Müller 2002), or VP relatives (28e), (28f).[8]

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

     b. * … dass Karl *wird* das Buch lesen

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

     d. * … weil Karl *scheint* das Buch zu lesen

     e. * Das ist das Buch das lesen Karl *wird*.

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

Similarly, scrambling of the complements is possible in a coherent construction, not in an incoherent one. In (29a) and (29c), the complements of *sehen* or *überreden*, respectively, do not mix. In (29b), *Peter*, the complement of *sehen*, occurs between *das Auto*, which is the complement of *kaufen*, and *kaufen*. Such a word order is possible with *sehen*, not *überreden* (29d) (examples from Hinrichs & Nakazawa 1998).

(29) a. … dass er [Peter] [das Auto kaufen] sehen wird
        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

     c. … dass er [Peter] [das Auto zu kaufen] überreden wird
        that he Peter the car to buy persuade will
        'that he will persuade Peter to buy the car'

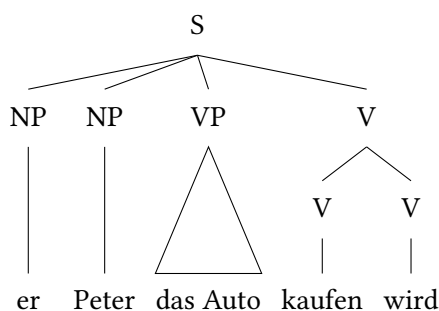   d. * … dass er das Auto Peter zu kaufen überreden wird

These data point to the following analysis: incoherent constructions involve a saturated VP complement (30a), while coherent constructions do not: they form a complex predicate, with a verb inheriting the complements of its complement. We assume here a verbal complex, a structure largely accepted for German (30b).

---

[8]The head verb of the coherent construction is underlined. We leave aside fronting of a verbal complement in a coherent construction, because the data are less straightforward (De Kuthy & Meurers 2001; Müller 2002).

(30)   a.   Incoherent construction (embedded clause)

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

   b.   Coherent construction (embedded clause)

```
                     S
        ┌─────┬──────┼─────────┐
       NP    NP     VP          V
        │     │     /\        ┌──┴──┐
        │     │    /  \       V     V
        │     │   /    \      │     │
       er   Peter das Auto  kaufen wird
```

4.2 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 (29b), for instance, while the structure would remain the same (as in (30a)). The existence of the remote (or long) passive goes against such an analysis (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 (31a) with an infinitival VP containing an accusative object (*den Wagen*) alternates with a coherent construction such as (31b), with a corresponding nominative (*der Wagen*).

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

> b. … weil der Wagen oft zu reparieren versucht wurde

In (31b), there is no infinitival VP, as shown by the occurrence of the adverb *oft*, which modifies *versucht*, between *der Wagen* and *zu reparieren*.

The German verbal complex, like the Spanish-1 one, instantiates a light head-complements-phrase. Thus, the two elements are light (22b). The constraint in (32b) renders explicit that in a verbal complex, the head inherits the complements of its complement. This complement can itself be a verbal complex (as in (29b)).[9]

Except for the verbal complex, we adopt here a flat structure for German (Uszkoreit 1987; Pollard 1996).[10] German differs from RL in not distinguishing structurally between the subject and the complements: they occur at the same level, can be interspersed, and are introduced by the same phrasal constraint. For this reason, it is useful to use the feature VAL (for valence, that is, the subject and the complements) in the description of lexemes and words, distinguishing the subject via the feature XARG on the lexemes, and the complements by COMPS only when it is necessary (Sag 2012). Thus, coherent constructions obey two phrasal constraints: the Head-Valents-phrase and the verbal-complex-phrase.

(32) a. *Head-Valents-phrase* ⇒

$$
\begin{bmatrix}
\text{MOTHER} \mid \text{SS} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{VAL } /\langle\rangle \\ \text{LIGHT } - \end{bmatrix} \end{bmatrix} \\
\text{HEAD-DTR}|\text{SS}|\text{CAT} \begin{bmatrix} \text{LIGHT } + \\ \text{VAL } \textit{nelist } \boxed{1} \end{bmatrix} \\
\text{NON-HD-DTRS } \boxed{1}
\end{bmatrix}
$$

b. *Verbal-Complex-phrase* (German) ⇒

$$
\begin{bmatrix}
\text{MOTHER} \mid \text{SS} \mid \text{CAT} \begin{bmatrix} \text{VAL } \boxed{1} \oplus \boxed{2} \\ \text{LIGHT } + \end{bmatrix} \\
\text{HD-DTR}|\text{SS}|\text{CAT}|\text{VAL } \boxed{1} \oplus \langle \boxed{3} \rangle \oplus \boxed{2} \\
\text{NON-HD-DTRS } \boxed{3} \begin{bmatrix} \textit{verb} \\ \text{LIGHT } + \\ \text{COMPS } \boxed{2} \end{bmatrix}
\end{bmatrix}
$$

The two phrases differ by the value of LIGHT, and also in that (32a) allows for a number of daughters while (32b) allows just for one non-head daughter, a verbal complex as illustrated in (29a) corresponding to a multi-level structure. The

---

[9]The representations of a verbal complex across RL, German and Korean, are homogenized. There is no need for XCOMP (Müller 2002) or GOV (Chung 1998) to distinguish this complement from the others in addition to specifying inheritance as in (32b) (Bonami & Webelhuth 2013).

[10]Alternatively, the structure can be represented as a hierarchy of binary branching structures, with a flat linearization domain (Müller 2002).

syntactic descriptions of the passive auxiliary *werden* and of *sehen* (in a coherent construction) are as in (33). *Sehen* is associated with two different descriptions: it has a incoherent version (no inheritance of valents) and a coherent one, which can be related via a lexical rule in the manner of Romance restructuring verbs as in (16).

(33)   a.   *werden* (passive auxiliary):
$$\begin{bmatrix} \text{XARG} \langle \boxed{1} \rangle \\ \text{VAL} \langle \boxed{1}, \text{V} \begin{bmatrix} \textit{passive-part}, \text{LIGHT } +, \text{VAL} \langle \boxed{1} \rangle \oplus \boxed{2} \end{bmatrix} \oplus \boxed{2} \rangle \end{bmatrix}$$

       b.   *sehen* (coherent version):
$$\begin{bmatrix} \text{X-ARG} \langle \text{NP}_i \rangle \\ \text{VAL} \langle \text{NP}_i, \text{NP}_k, \text{V} \begin{bmatrix} \textit{inf}, \text{VAL} \langle \text{NP}_k \rangle \oplus \boxed{2} \end{bmatrix} \oplus \boxed{2} \rangle \end{bmatrix}$$

We leave aside the placement of the head verb in the German clause, which is a question in itself. In a verbal complex, the head is ordered after the complement. The structure of (29b) is illustrated in Figure 2.

> Please use feature geometry of Sag (1997). People working on German usually assume binary branching structures. See Chapter on order.

Figure 2: Analysis of *dass er das Auto Peter kaufen sehen wird*

In Figure 2, the verb $\boxed{6}$ is a verbal complex, whose head is *wird*, which inherits the complements of its complement $\boxed{4}$, also a verbal complex; the head of $\boxed{4}$ is *sehen*, which identifies its subject (or xarg) with that of *wird* (see (33a)), and inherits the complement of its complement *kaufen*, that is $\boxed{2}$, in addition to its

other complement ③ (see (33b)); the verbal complex ④ inherits its valence from
its head, minus the complement which is saturated (⑤).

4.3 Argument inheritance in Korean

4.3.1 Scrambling in Korean

Korean resembles German in that a complex predicate signals itself mainly by
word order properties. Complex predicates are formed from auxiliaries, control
verbs, and causative *ha*. We follow in particular Chung (1998), from which the
examples are borrowed.

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* 'want', but also the verb of negation *anh* 'not'.
Scrambling with auxiliaries is illustrated in (34). In (34b) the subject of the auxil-
iary *issta* (*Mary-ka*) occurs between the embedded verb (*ilke*) and its complement
(*chayk-ul*).

(34) a. Mary-ka   ku chayk-ul ilke poko issta.
        Mary-NOM the book-ACC read try   be.in.the.process.of-DECL

        'Mary is giving the book a trial reading.'

    b. Ku chayk-ul Mary-ka ilke poko issta.

Control verbs such as *seltuk* 'persuade' (35a), (35b) , *cisi* 'order' (object control),
yaksok 'promise', *sito* 'try' (subject control) (35c), (35d), as well as causative *ha*,
also allow valents of the embedded verb to be interleaved with their own valents,
as in (35b) and (35d).

(35) a. Mary-ka   John-hantley [ku chayk-ul ilkulako] seltukhayssta.
        Mary-NOM John-DAT   the book-ACC read   persuade-past-DECL

        'Mary persuaded John to read the book.'

    b. Ku chayk-ul Mary-ka John-hantley ilkulako seltukhayssta.
       the book-ACC Mary-NOM John-DAT read-will persuade-PAST-DECL

       'Mary persuaded John to read the book.'

    c. Mary-ka   John-hantley [ku chayk-ul pilyecwu-kesstako]
       Marie-NOM John-DAT   the book-ACC lend-will
       yaksokhayssta.
       promise-PAST-DECL

       'Mary promised John to lend the book.'

    d. Ku chayk-ul John-hantley Mary-ka   pilyecwu-kesstako
       the book-ACC John-DAT   Marie-NOM lend-will
       yaksokhayssta.
       promise-PAST-DECL
       'Mary promised John to lend the book.'

These data can be explained in two ways: either the auxiliary or the control verb always takes a VP complement, and scrambling is due to linearization (the domains of the two verbs being unioned), or there is a complex predicate: the complement of the embedded verb (the book) is inherited by the auxiliary or the control verb, and, assuming a flat structure for the Korean sentence (with a head-valents-phrase as in (32a)), it occurs at the same level as the subject. As in German, 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: *malsseng-manhun so-ka* is subject of the passive auxiliary *ci-esta* in (36b) and *ku cengchayk-i* is subject of the passive of *order* in (37b).

(36)   a. Ku mongpwu-ka malsengmanhun so-lul   phala
         The farmer-NOM troublesome   cow-ACC sell
         chiw-essta.
         do.resultely-PAST-DECL
         'The farmer resolutely sold the troublesome cow.'

      b. Malsengmanhun so-ka   (ku nongpwu-eyuyhay) phala
         troublesome   cow-NOM the farmer-by   sell
         chiwe   ci-essta.
         do.resolutely passive-PAST-DECL
         'The troublesome cow was resolutely sold (by the farmer).'

(37)   a. Nay-ka Mary-hantley ku cengchayk-ul sihaynghalako
         I-NOM   Mary-DAT   the policy-ACC   carry.out
         cisihayssta.
         order-PAST-DECL
         'I ordered Mary to carry out the policy.'

      b. Ku cengchayk-i naey-uyhayse Mary-hantley sihaynghalako
         the policy-NOM I-by   Mary-DAT   carry.out

cisi-toy-essta.

order-PASS-PAST-DECL

(lit. the policy was ordered by me for Mary to carry out)

Since passivization only affects the complement of the verb which is itself passivized, it follows that *malssengmanhun so-lul* is the complement of the auxiliary in (36a), and *ku cengchayk-ul* the complement of the control verb in (37a).[11]

4.3.2 The structures of CP in Korean

In Korean, the structure of a complex predicate formed on an auxiliary contrasts with that formed on 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).

A characteristic property of Korean auxiliary constructions is that nothing can separate the two verbs. While an NP may occur after the verbs in a construction called afterthought (38a), this is not possible for the embedded verb alone (38b) or with its complement (38c), and no parenthetical expression can intervene as shown by *hayekan* in (39b) (examples from Chung 1998). This behavior follows if the two verbs form a verbal complex (see section 3.2).

(38) a. Mary-ka   mekko issta,   sakwa-lul.
          Mary-NOM eat   be.in.the.process.of-DECL apple-ACC
          'Mary is in the process of eating an apple.'

      b. * Mary-ka sakwa-lul issta, mekko

      c. * Mary-ka issta, sakwa-lul mekko

(39) a. Mary-ka   hayekan sakwa-lul mekko issta.
          Mary-NOM anyway apple-ACC eat   be.in.the.process.of-DECL
          'Anyway, Mary is eating an apple.'

      b. * Mary-ka sakwa-lul mekko hayekan issta

Korean has a Verbal-Complex-phrase similar to (32b). In Korean, this structure characterizes the class of auxiliaries. This is captured by the head feature AUX: auxiliaries are [AUX +], while the other verbs are [AUX −]. Moreover, unlike the

---

[11]More complicated data, showing that the case of the inherited complement depends on the agentivity of the combination of auxiliary *siph* with its verbal complement, can be explained if the auxiliary and the verbal complement form a verbal complex (Yoo 2003).

head of a verbal complex in German, which can be a control verb (like *sehen*), they inherit the whole argument structure of their complement. Note the use of 'o', the shuffle operator, which allows to reorder the elements of the two lists.

(40) *Verbal-Complex-phrase* (Korean) ⇒

$$
\begin{bmatrix}
\text{MOTHER|SS|CAT} \begin{bmatrix} \text{AUX } + \\ \text{LIGHT } + \\ \text{VAL } \boxed{1} \end{bmatrix} \\[3ex]
\text{HEAD-DTR|SS|CAT} \begin{bmatrix} \text{AUX } + \\ \text{LIGHT } + \\ \text{VAL } \boxed{1} \text{ O } \langle \boxed{2} \rangle \end{bmatrix} \\[3ex]
\text{NON-HEAD-DTRS} \left\langle \boxed{2} \begin{bmatrix} \textit{verb} \\ \text{LIGHT } + \\ \text{VAL } \boxed{1} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

An auxiliary such as *iss* 'to be in the process of' has the following syntactic description:

(41)
$$
\begin{bmatrix}
\text{HEAD} \begin{bmatrix} \text{AUX } + \end{bmatrix} \\
\text{LIGHT } + \\
\text{VAL} \left\langle \text{V} \begin{bmatrix} \text{VFORM } ko, \text{ LIGHT } +, \text{ VAL } \boxed{1} \end{bmatrix} \right\rangle \bigcirc \boxed{1}
\end{bmatrix}
$$

The auxiliary *iss* takes as its complement a light verb which is constrained to end in −*ko* (different auxiliaries put different syntactic and semantic restrictions on their verbal complement, (Yoo 2003)), from which it inherits its valence. The sentence in Korean is generally assumed to have a flat structure, which corresponds to the Head-Valents-phrase in (32a). The structure of (34a), with a series of two auxiliaries, is similar to (2). *Issta* takes as its complement *ilke poko*, whose head is *poko*, and which is also a verbal complex. *Poko*, being an auxiliary like *iss* (41), takes as its complement the verb *ilke*, from which it inherits its valence, which is transmitted to the verbal complex *ilke poko*; *ilke poko* saturates the verbal complement expected by *issta*, and transmits its valence to the head auxiliary. The structure is schematized in Figure 3.

Contrary to auxiliaries, control verbs such as *seltuk* (35) or *cisi* (37) can be separated from their verbal complement, for instance by the adverb in (42a). They also allow for the infinitive and its complement to form a VP constituent as in (42b), where they constitute an afterthought. Thus, control verbs are analyzed in the same way as Italian Restructuring verbs: they can take a saturated VP complement, or be the head of a complex predicate. They contrast with Korean Raising verbs which only take a VP complement.

Draft of 7th November 2019, 11:22

Figure 3: Add caption

(42) a. Mary-ka   ku mwuncey-lul phwulye-ko (kkuncilkikey)
         Mary-ɴᴏᴍ the problem-ᴀᴄᴄ solve   ceaselessly
         sitohayssta
         try-ᴘᴀsᴛ-ᴅᴇᴄʟ

         'Mary tried (ceaselessly) to solve the problem.'

    b. Mary-ka sitohayssta, [ku mwuncey-lul phwulye-ko].

    c. Ku chayk-ul Mary-ka   ilkulako sitohayssta.
         The book-ᴀᴄᴄ Mary-ɴᴏᴍ read   try-ᴘᴀsᴛ-ᴅᴇᴄʟ

         'Mary tried to read the book.'

    d. Ku chayk-ul Mary-ka   ilkulako John-hantley seltukhayssta.
         The book-ᴀᴄᴄ Mary-ɴᴏᴍ read   John-ᴅᴀᴛ   persuade-ᴘᴀsᴛ-ᴅᴇᴄʟ

         'Mary persuaded John to read the book.'

The scrambling data, together with the possibility of long passivization, show that there is a complex predicate (section 4.3.1). For instance, the subject of the head verb (*Mary-ka*) comes between the complement of the infinitive and the infinitive in (42c), (42d). 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 (42d) (*John-hantley*). Thus, they are the head of a flat structure as in Figure 4 corresponding to (42d).

As stated in the description of verbal complexes in Korean (40), the head of the structure is [AUX +]. Since control verbs are [AUX −], they cannot enter

Figure 4: Add caption

verbal complexes, and the structure for complex predicates is a flat Head-Valents-phrase, as described in (32a).

The head comes last in Korean, except in the afterthought construction (42b), 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 (42a) *Mary-ka [phwulye-ko] [ku mwuncey-lul] sitohayssta*, where the verb *phwulye-ko* precedes its complement *ku mwuncey-lul* is thus impossible. In (42d), the NP *John-hantley* can follow the complement verb *ilkulako* because it is the complement of the head verb *seltukhayssta*, not of *ilkulako*. We can state the constraint informally as in (43).

(43)   X precedes the HEAD

Chung (1998) extends the possibility of inheritance to adjuncts, as well as to constructions with a S complement, which 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 valents and (lexical) verbs are all sisters. In this analysis, the definition of a complex predicate, which relies on a syntactic relation between words can still 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, where the head is a verb and the second predicate belongs to different categories, depending on the language and the construction, allow us to focus on other aspects of complex predicates. Such constructions are

frequent in different languages (see Choi & Wechsler 2002; Lee 2001: for noun-verb combinations in Korean). We focus on Persian (Bonami & Samvelian 2010; Müller 2010; Samvelian 2012), examining two aspects: morphosyntactic and semantic.

Persian complex predicates are multiword expressions. More precisely, they are lexemes which are realized by several words (Bonami & Samvelian 2010). To solve the morphosyntactic and semantic problems, 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.

5.1 Complex predicates and derivational processes

Although complex predicates have properties expected of the combination of words (they can be separated by the future auxiliary, for instance), they give rise to agent nominalization. Examples (44) show that 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'.

(44) a. kon 'do', *kon-ande 'do-er'
      b. bāz kon-ande 'opener'

We follow Müller (2010), who exploits the possibility for a derived word to keep track of its formation via a daughter. Thus, the verb part of *bāz kon-ande* is the target of a nominalization rule, and is combined with an affix (*-ande*), but this is possible only if it is waiting for the adjective complement. The agent nominalization rule takes the verb as input and gives a noun which refers to the expected subject of the verb, and, syntactically takes as its complement (inherited from the verb) the adjectival predicate. This is represented in (5) (adapted from Müller 2010).

The content of the verb *kon* in the complex predicate *bāz kon* is a cause relation, with the state expressed by the adjective (noted as 5) as a result. The derived N is thus the opener.

5.2 The Semantics of Light verb constructions

Although Persian complex predicates can be viewed as lexemes, the nominal element and the verb in noun-verb combinations behave nevertheless as two independent syntactic units ((Karimi-Doostan 1997; Megerdoomian 2002; Samvelian 2012). All inflection is prefixed or suffixed on the verb, as is the negation in (45),

$$\text{N[CONT } \boxed{1}]$$

$$\boxed{2}\begin{bmatrix} adjective \\ \text{XARG} \ \langle\, \boxed{3}\, \text{NP}_k \,\rangle \\ \text{CONT} \ \boxed{5}\, open(k) \end{bmatrix} \quad \text{N} \begin{bmatrix} \text{PHON} & \boxed{7} \oplus \langle\, ande \,\rangle \\ \text{ARG-ST} & \langle\, \boxed{3}, \boxed{2} \,\rangle \\ \text{CONT} & \boxed{1}\begin{bmatrix} \text{IND} & m \end{bmatrix} \end{bmatrix}$$

-ande nominalization LR

$$\text{V} \begin{bmatrix} \text{PHON} & \boxed{7} \\ \text{XARG} & \langle\, \boxed{6}\, \text{NP}_m \,\rangle \\ \text{ARG-ST} & \langle\, \boxed{6}, \boxed{3}, \boxed{2} \,\rangle \\ \text{CONT} & cause(m, \boxed{5}) \end{bmatrix}$$

baz	kon
open	do

Figure 5: Insert caption

and never on the noun. The two elements can be separated by pronominal clitics, the future auxiliary, or even by clearly syntactic constituents, like the complement in (45). Both the noun and the verb can be coordinated, as shown in (46) and (47) respectively. The noun can be extracted, as in the topicalization in (48). CPs can be passivized. In this case, the nominal element of the CP (*tohmat* in (49a)) becomes the subject of the passive construction (49b), as does the object of a transitive construction. Finally, the noun can be modified (50). The nominal part of the CP is underlined in the examples.

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

(46) Mu-hā=yaš=rā *boros* yā *šāne* zad.
      hair-PL=3SG=RA brush or comb strike.PAST.3SG
      '(S)he brushed or combed her hair.'

Draft of 7th November 2019, 11:22

(47) Omid *sili* zad   va xord.
Omid slap hit.PAST.3SG and strike.PAST.3SG
'Omid gave and received slaps.'

(48) *Dast* goft-am [be gol-hā   —— na-zan].
hand said-1SG to flower-PL —— NEG-strike.PRS
'I told you not to touch the flowers.'

(49) a. Maryam be Omid *tohmat* zad.
Maryam to Omid slander strike.PAST.3SG
'Maryam slandered Omid.'

b. be Omid *tohmat* zade šod.
to Omid slander hit.PP become.PAST.3SG
'Omid was slandered.'

(50) [In *xabar*=e mohem]=rā   be mā dād
this news=EZ important=RA to us give.PAST.3SG
'(S)he gave us this important piece of news.'

Nevertheless, there is evidence that the verb and the noun may share one argument structure. For instance, in 51b the verb *dādan* and the noun *āb* take together a direct object.

(51) a. Maryam be bāqče āb   dād
Maryam to garden water give.PAST.3SG
'Maryam watered the gadern.'

b. Maryam bāqče=rā   āb   dād
Maryam garden-DOM water give.PAST.3SG
'Maryam watered the gadern.'

Persian 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 semantically compositional combination, with the head taking as its argument the content of its complement, like in Romance languages CP (see section §3) or "causative" predicates formed with in adjective, like *bāz kardan* 'to open' (open do) (see section 5.1), to idioms whose

semantics is not predictable from the components. Such is the case of *dast zadan (be)* 'to begin' (hand strike):[12]

(52) kārgar-ān be e'tesāb dast   zad-and
       worker-PL to strike   hand strike.PAST-3PL
       'The workers went on strike.'

We begin with idioms illustrated by *dast zadan*, following analyses in Müller (2010), which draw on Goldberg (1996), Nunberg et al. (1994), and Sag (2007): the head of the idiom is an item in the lexicon, with special selection properties and the appropriate meaning for the idiom. More precisely, the verb is associated with an idiomatic reading, an i-relation which arises in the special combination; it selects its complement via the lexical identifier or LID feature, which is associated with lexemes in the lexicon, and contains morpho-syntactic as well as semantic information (Sag 2007; 2012). The lexical entry for the verb *zadan* 'to hit' in the idiomatic combination *dast zadan* 'to start' is given in (53):

(53) Lexical entry of *zadan* 'strike' in *dast zadan* 'to start'

$$
\begin{bmatrix}
\textit{zadan1-lxm} \\
\text{CAT } \textit{verb} \\
\text{CONT} \begin{bmatrix} \textit{soa} \\ \text{NUCLEUS} \begin{bmatrix} \text{REL} & \textit{i-start-rel} \\ \text{AGENT} & k \\ \text{THEME} & m \end{bmatrix} \end{bmatrix} \\
\text{ARG-ST} \left\langle \text{NP}_k, \left(be\right)\text{NP}_m, \text{N} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{LIGHT} & + \\ \text{LID} & \textit{dast} \end{bmatrix} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

The same mechanisms can be used to describe the intermediate cases. We illustrate three cases, drawing on the detailed study of the verb *zadan* 'to strike' in Samvelian (2012). They show different ways in which the noun takes preeminence over the verb. The meaning of the verb can be close to a basic meaning that it has outside the light verb construction; the noun contributes to a more specific interpretation, and the meaning of the verb is absorbed into that of the noun. This is the case for the light verb construction involving *zadan* 'to strike' and a noun describing a type of blow, for instance, *lagad* 'kick', *sili* 'slap'. So *lagad zadan* simply means 'to kick'.

---

[12]Note that the combination *dast zadan* has several meanings. When used with a PP argument introduced by *be* 'to', it either means 'to touch' or 'to start'. In its intransitive use, it means 'to clap'.

Draft of 7th November 2019, 11:22

(54) olāq   be Omid lagad zad
      donkey to Omid kick strike.PAST.3SG
      'The donkey kicked Omid.'

The description of *zadan* (here, *zadan-2*) when it combines with a noun depicting a type of blow is as in (55).

(55)
$$
\begin{bmatrix}
\textit{zadan2-lxm} \\
\text{CAT } \textit{verb} \\
\text{CONT } \begin{bmatrix}
\textit{soa} \\
\text{NUCLEUS} \begin{bmatrix}
\text{REL} & \textit{i-inflict-}\boxed{1}\textit{-rel} \\
\text{AGENT} & k \\
\text{PATIENT} & m
\end{bmatrix}
\end{bmatrix} \\
\text{ARG-ST } \left\langle \text{NP}_k, \left(be\right)\text{NP}_m, \text{N}\begin{bmatrix}\text{CAT}\begin{bmatrix}\text{PRED} & + \\ \text{LID} & \boxed{1}\ \textit{blow-rel}\end{bmatrix}\end{bmatrix}\right\rangle
\end{bmatrix}
$$

In other cases, the meaning of the noun takes over: the construction is similar to a derivation rule, and converts the noun into a verb with the same meaning, 57. When *zadan* combines with a noun describing a kind of fraud or deception, the complex predicate means 'deceive' in the way described more precisely by the noun.

(56) Hālā mā kalak be=hešun mi-zan-im
      Now we deceit to=them   IPFFV-strike.PRS-3PL
      'Now, it is our turn to deceive them.'

(57)
$$
\begin{bmatrix}
\textit{zadan3-lxm} \\
\text{CAT } \textit{verb} \\
\text{CONT } \begin{bmatrix}
\textit{soa} \\
\text{NUCLEUS} \begin{bmatrix}
\text{REL} & \textit{i-deceit-}\boxed{1}\textit{-rel} \\
\text{AGENT} & k \\
\text{PATIENT} & m
\end{bmatrix}
\end{bmatrix} \\
\text{ARG-ST } \left\langle \text{NP}_k, \left(be\right)\text{NP}_m, \text{N}\begin{bmatrix}\text{CAT}\begin{bmatrix}\text{PRED} & + \\ \text{LID} & \boxed{1}\ \textit{deceit-rel}\end{bmatrix}\end{bmatrix}\right\rangle
\end{bmatrix}
$$

Finally, there are also cases where the meaning of the noun is preserved, but a new meaning emerges for the verb, in a non predictable manner. When *xod=rā zadan* (lit. self strike) combines with a subset of nouns describing illnesses, handicaps or problematic states (like stupidity, ignorance), it means 'to pretend, feign' the illness or state in question.

(58) Maryam xod=rā be divanegi zad
      Maryam self=RA to madness strike.PAST.3SG
      'Maryam feigned madness.'

(59)
$$
\begin{bmatrix}
\textit{zadan4-lxm} \\
\text{CAT } \textit{VERB} \\
\text{CONT} \begin{bmatrix} \text{SOA} \\ \text{NUCLEUS} \begin{bmatrix} \text{REL} & \textit{i-pretend}\boxed{1}\textit{-rel} \\ \text{AGENT} & k \\ \text{THEME} & e \end{bmatrix} \end{bmatrix} \\
\text{ARG-ST} \left\langle \text{NP}_k, \textit{pro}_k, \text{N} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{INDEX} & e \\ \text{LID} & \boxed{1}\textit{i-illness-rel} \end{bmatrix} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

Not all nouns for illnesses and problematic states can occur in this complex predicate. To account for this, we indicate that the noun, too, has an idiomatic meaning, an *i-rel*, in addition to its usual meaning, and only the nouns which have such an *i-rel* can participate in the complex Predicate with *xod=rā zadan*. Here, the case is close to that of an idiom, but the noun does not have a special meaning: it is special in that not all nouns with a comparable semantics can form such a complex Predicate with *xod=rā zadan*.

Acknowledgements

References

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

Abeillé, Anne & Danièle Godard. 2005. The diversity of complex predicates in Romance languages: argument sharing and syntactic structure. Manuscript, University of Université Paris-Diderot.

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

Aissen, Judith & David Perlmutter. 1983. Clause reduction in spanish. In David Perlmutter (ed.), *Studies in relational grammar 1*, 360–403. Chicago: The University of Chicago Press.

Bech, Gunnar. 1955. *Studien über das deutsche Verbum infinitum* (Linguistische Arbeiten 139). Tübingen: Max Niemeyer Verlag. 2nd unchanged edition 1983.

Bonami, Olivier & Pollet Samvelian. 2010. *Persian complex predicates: lexeme formation by itself*. Talk presented at Les Décembrettes 7. Colloque international de morphologie, 2-3 December, 2-3, 2010.

Bonami, Olivier & Pollet Samvelian. 2015. The diversity of inflectional periphrasis in Persian. *Journal of Linguistics* 51(1). 327–382. DOI:10.1017/S0022226714000243

Bonami, Olivier & Gert Webelhuth. 2013. The phrase-structural diversity of periphrasis : a lexicalist account. In Marina Chumakina et Greville G. Corbett (ed.), *Periphrasis : the role of syntax and morphology in paradigms*, 141–167. Oxford: Oxford University Press.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and 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. 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.), *The proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, 103–120. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2/, accessed 2018-3-20.

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

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

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

Goldberg, Adele E. 1996. Words by default: Optimizing constraints and the Persian complex predicate. *Berkeley Linguistic Society* 22. 132–146. http://roa.rutgers.edu/files/415-0900/roa-415-goldberg-2.pdf, accessed 2018-3-20.

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

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

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.

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.

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.

Lee, Sun-Hee. 2001. Argument composition and linearization: Korean complex predicates and scrambling, *OSUWPL* 56. 53–78.

Megerdoomian, Karine. 2002. *Beyond words and phrases: A unified theory of predicate composition.* University of Southern California dissertation. http://www.zoorna.org/publications.html, accessed 2015-12-23.

Meurers, Detmar. 2002. 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.), *The proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, 235–246. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2/, accessed 2018-3-20.

Draft of 7th November 2019, 11:22

Miller, Philip H. & Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language and 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. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655. DOI:10.1017/S0022226709990284

Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538.

Pollard, Carl J. 1996. On head non-movement. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 279–305. Berlin: Mouton de Gruyter. Published version of a Ms. dated January 1990.

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.

Rizzi, Luigi. 1982. *Issues in Italian syntax* (Studies in Generative Grammar 11). Dordrecht: Foris.

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. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

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.

Salvi, Giampaolo. 2010. Gli ausiliari essere e avere in italiano. *Acta linguistica Academiae scientiarum Hungaricae* 30(1-2). 137–162.

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

Samvelian, Pollet & Pegah Faghiri. 2016. In *Proceedings of the 39th Annual Meeting of the Berkeley Linguistics Society*, 212–227. Berkeley: Berkeley Linguistics Society.

Sells, Peter. 1991. Complex verbs and argument structures in Korean. In Susumo Kuno (ed.), *Harvard studies in korean linguistics 4*, 395–406. Seoul: Hanshin.

Suñer, Margarita. 2010. Two properties of clitics in clitic-doubled constructions. In C.-T. James Huang & Robert May (eds.), *Logical structure and linguistic structure*, vol. 40 (Studies in Linguistics and Philosophy), 233–252. Dordrecht: Kluwer.

Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.

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.