

## Persian complex predicates and the limits of inheritance-based analyses<sup>1</sup>

STEFAN MÜLLER

*Freie Universität Berlin*

(Received 8 August 2008; revised 21 May 2009)

Persian complex predicates pose an interesting challenge for theoretical linguistics since they have both word-like and phrase-like properties. For example, they can feed derivational processes, but they are also separable by the future auxiliary or the negation prefix.

Various proposals have been made in the literature to capture the nature of Persian complex predicates, among them analyses that treat them as purely phrasal or purely lexical combinations. Mixed analyses that analyze them as words by default and as phrases in the non-default case have also been suggested.

In this paper, I show that theories that rely exclusively on the classification of patterns in inheritance hierarchies cannot account for the facts in an insightful way unless they are augmented by transformations or some similar device. I then show that a lexical account together with appropriate grammar rules and an argument composition analysis of the future auxiliary has none of the shortcomings that classification-based analyses have and that it can account for both the phrasal and the word-like properties of Persian complex predicates.

---

[1] I thank Mina Esmaili, Fateme Nemati, Pollet Samvelian, Yasser Shakeri, and Mehran A. Taghvaipour for help with the Persian data and for comments on an earlier version of this paper. Without their help I would not have been able to write this paper. In addition I want to thank Elham Alaei for discussion of data and Daniel Hole and Jacob Maché for comments on the paper. I thank Emily M. Bender and Felix Bildhauer for discussion. Special thanks go to Bob Borsley, Karine Megerdumian, Ivan Sag, Pollet Samvelian, and several anonymous reviewers, whose comments on earlier versions of the paper improved it considerably. I also thank Philippa Cook for proof reading.

Research related to this paper was presented at the HPSG conference 2006 in Varna, at the Institut für Linguistik in Leipzig, at the Institut für Linguistik in Potsdam, at the Centrum für Informations und Sprachverarbeitung in Munich, at the center for Computational Linguistics of the Universität des Saarlandes, Saarbrücken, in 2008 at the conference Complex Predicates in Iranian Languages at the Université Sorbonne Nouvelle in Paris, and in 2008 at the conference Syntax of the World's Languages at the Freie Universität Berlin. I thank the respective institutions and the organizers of the HPSG conference for the invitation and all audiences for discussion.

This work was supported by a grant from Agence Nationale de la Recherche and Deutsche Forschungsgemeinschaft to the Franco-German project PER-GRAM: Theory and Implementation of a Head-driven Phrase Structure Grammar for Persian (DFG MU 2822/3-1).

## I. INTRODUCTION

Persian complex predicates pose an interesting challenge for theoretical linguistics, since they have both word-like and phrase-like properties. For instance, they can feed derivational processes, but they are also separable by the future auxiliary or the negation prefix.

Various proposals have been made in the literature to capture the nature of Persian complex predicates, among them analyses that treat them on a purely phrasal basis (Folli, Harley & Karimi 2005) or purely in the lexicon (Barjasteh 1983). Mixed analyses that analyze them as words by default and as phrases in the non-default case have also been suggested (Goldberg 2003).

In this paper, I show that theories that rely exclusively on the classification of patterns in inheritance hierarchies cannot account for the facts in an insightful way unless they are augmented by transformations or similar devices. I then show that a lexical account together with appropriate Immediate Dominance schemata<sup>2</sup> and an argument attraction analysis of the future auxiliary has none of the shortcomings that classification-based phrasal analyses have and that it can account for both the phrasal and the word-like properties of Persian complex predicates. While the paper focuses on the analysis of Persian complex predicates, its scope is much broader since phrasal inheritance-based analyses have been suggested as a general way to analyze language (Croft 2001: 26; Tomasello 2003: 107; Culicover & Jackendoff 2005: 34, 39–49; Michaelis 2006: 80–81).

The paper will be structured as follows: section 2 gives a brief introduction to the concept of (default) inheritance, section 3 gives an overview of the properties of Persian complex predicates that have been discussed in connection with the status of complex predicates as phrases or words, section 4 discusses the inheritance-based phrasal analysis and its problems, section 5 provides a lexical analysis of the phenomena, section 6 compares this analysis to the phrasal analysis, and section 7 draws some conclusions.

## 2. INHERITANCE

Inheritance hierarchies are a tool to classify knowledge and to represent it compactly. The best way to explain their organization is to compare an inheritance hierarchy with an encyclopedia. An encyclopedia contains both very general concepts and more specialized concepts that refer to the more general ones. For instance, concepts can be ‘living thing’, ‘animal’, ‘fish’, and ‘perch’. The encyclopedia contains a description of all the properties that living things have. This description is part of the entry for ‘living thing’;

---

[2] An Immediate Dominance (ID) schema corresponds to grammar rules in other theories. It is basically a set of constraints that has the same function as a rewrite rule in a context free grammar.

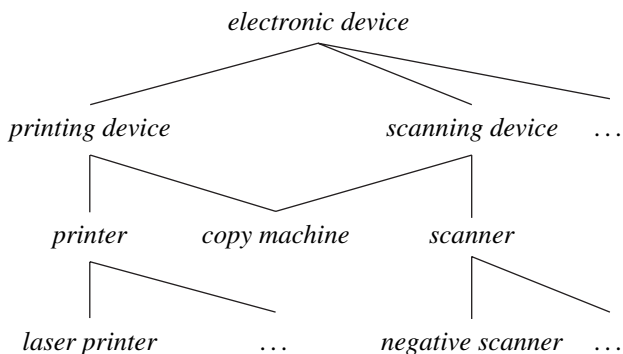


Figure 1  
Non-linguistic example for multiple inheritance.

the sub-concept ‘animal’ does not repeat this information, but instead refers to the concept ‘living thing’. The same is true for sub-concepts of ‘animal’: they do not repeat information relevant for all living things or animals, but refer to the direct super-concept ‘animal’. The connections between concepts form a hierarchy, that is there is a most general concept that dominates sub-concepts. The sub-concepts are said to inherit information from their super-concepts, hence the name *inheritance hierarchy*. From the encyclopedia example, it should be clear that some concepts can refer to several super-concepts. If a hierarchy contains such references to several super-concepts, one talks about *multiple inheritance*.

Inheritance hierarchies can be represented graphically. An example is shown in figure 1. The most general concept in figure 1 is *electronic device*. Electronic devices have the property that they have a power supply. All sub-concepts of *electronic device* inherit this property. There are several sub-concepts of *electronic device*. The figure shows *printing device* and *scanning device*. A printing device has the property that it can print information and a scanning device has the property that it can read information. *Printing device* has a sub-concept *printer*, which in turn has a sub-concept *laser printer* that corresponds to a certain class of printers that have special properties which are not properties of printers in general. Similarly there are scanners and certain special types of scanners. Copy machines are an interesting case: they have properties of both scanning devices and printing devices since they can do both, read information and print it. Therefore the concept *copy machine* inherits from two super-concepts: from *printing device* and *scanning device*. This is an instance of multiple inheritance.

Sometimes such inheritances are augmented with defaults. In default inheritance hierarchies information from super-concepts may be overridden in sub-concepts. Figure 2 shows the classical penguin example.

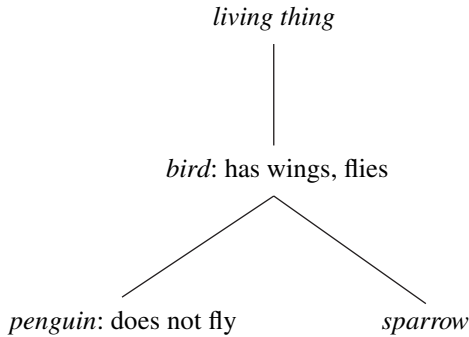


Figure 2  
Inheritance hierarchy for *bird* with default inheritance.

This hierarchy states that all birds have wings and can fly. The concepts below *bird* inherit this information unless they override it. The concept *penguin* is an example for partial overriding of inherited information: penguins do not fly, so this information is overridden. All other information (that birds have wings) is inherited.

Such hierarchies (with or without defaults) can be used for the classification of arbitrary objects, and in linguistics they are used to classify linguistic objects. For instance, Pollard & Sag (1987: section 8.1), who work in the framework of Head-driven Phrase Structure Grammar (HPSG), use type hierarchies to classify words in the lexicon. More recent work in HPSG uses type hierarchies to classify immediate dominance schemata (grammar rules); examples are Sag (1997) and Ginzburg & Sag (2000). These approaches are influenced by Construction Grammar (CxG; Fillmore, Kay & O'Connor 1988, Kay & Fillmore 1999), which emphasizes the importance of grammatical patterns and uses hierarchies to represent linguistic knowledge and to capture generalizations. Recent publications in the CxG framework that make use of inheritance hierarchies are for instance Goldberg (1995, 2003), Croft (2001) and Michaelis & Ruppenhofer (2001).

In the following I discuss Goldberg's (2003) proposal for Persian complex predicates and show the limits of such classification-based analyses. The conclusion to be drawn is that inheritance can indeed be used to model regularities in certain isolated domains of grammar, namely those domains that do not interact with valence alternations (for instance relative clauses or interrogative clauses, see Sag 1997 and Ginzburg & Sag 2000), and in the lexicon (see Pollard & Sag 1987: chapter 8) for the use of inheritance hierarchies in the lexicon, and Krieger & Nerbonne 1993, Koenig 1999, and Müller (2007b: section 7.5.2) for the limits of inheritance as far as lexical relations are concerned); but they are not sufficient to describe a language in total, since crucial properties of language (certain cases of

embedding and recursion) cannot be captured by inheritance alone in an insightful way or – in some cases – cannot be captured at all.

### 3. THE PHENOMENON

Whether Persian complex predicates should be treated in the lexicon or in the syntax, whether they are words or phrases is a controversial issue. This section repeats the arguments that were put forward for a treatment as words (by default) and shows that none of these arguments are without problems (section 3.1). Section 3.2 repeats arguments for the phrasal treatment that are relevant for the discussion of the inheritance-based approach.

Persian complex predicates consist of a preverb (sometimes also called host) and a verbal element. (1) is an example: *telefon* is the preverb and *kard* is the verbal element. (See appendix for list of abbreviations used in example glosses.)

- (1) (man) telefon kard-am  
 I telephone did-1SG  
 ‘I telephoned.’

The preverb may be a noun, an adjective, an adverb, or a preposition. If the preverb is a noun, it cannot appear with a determiner but must appear in bare form without plural or definite marking.

The literature mentions the following properties as evidence for the word status of Persian complex predicates. Persian complex predicates receive the primary stress, which is normally assigned to the main verb. They may differ from their simple-verb counterparts in argument structure properties, they resist separation – for example, by adverbs and by arguments – and undergo derivational processes that are typically restricted to apply to zero level categories. These arguments are not uncontroversial. They will be discussed in more detail below.

Apart from the word-like properties already mentioned, Persian complex predicates also have certain phrasal properties. First, Persian complex predicates are separated by the future auxiliary; second, the imperfective prefix *mi-* and the negative prefix *na-* attach to the verb and thus split the preverb and the light verb; and third, direct object clitics may separate preverb and verb.

Both the alleged word-like and the phrase-like properties will be discussed in more detail in the following two sections.

#### 3.1 *Word-like properties*

This subsection is devoted to a discussion of word-like properties of Persian complex predicates. The following phenomena will be discussed: stress, changes in argument structure, transitive complex predicates, separation of light verb and preverb, and morphological derivation. It will be shown that

only the last phenomenon provides conclusive evidence for a morphological analysis.

### 3.1.1 *Stress*

The primary stress is placed on the main verb in finite sentences with simple verbs (2a).<sup>3</sup>

- (2) (a) Ali mard-râ ZAD (simple verb)  
 Ali man-râ hit.3SG  
 ‘Ali hit the man.’  
 (b) Ali bâ Babâk HARF zad (complex predicate)  
 Ali with Babâk word hit.3SG  
 ‘Ali talked with Babâk.’

In finite sentences with a complex predicate, the primary stress falls on the preverb (2b). Goldberg claims that this is accounted for if complex predicates are treated as words; but stress is word-final in Persian, so one would have expected stress on *zad* if *harf zad* were a word in any sense that is relevant for stress assignment. In addition, Folli et al. (2005: 1391) point out that in sentences with transitive verbs that take a non-specific, non-case-marked object, the stress is placed on the object. They give the example in (3).<sup>4,5</sup>

- (3) man DAFTAR xarid-am  
 I notebook bought-1SG  
 ‘I bought notebooks.’

Since *daftar xaridam* is clearly not a word, the pattern in (2) should not be taken as evidence for the word status of complex predicates.

There is a more general account of Persian accentuation patterns that can explain the data without assuming that Persian complex predicates are words: Kahnemuyipour (2003: section 4) develops an analysis which treats complex verbs as consisting of two phonological words. Since the preverb is the first phonological word in a phonological phrase, it receives stress by a principle that assigns stress to the leftmost phonological word in a Phonological Phrase (344). Phonological Phrases are grouped into

[3] The examples are taken from Goldberg (2003: 122). This argument for the word status of complex predicates can also be found in Ghomeshi & Massam (1994: 183, 186, 189). See also Vahedi-Langrudi (1996: 43) and Dabir-Moghaddam (1997: 48–49) for similar data and a similar treatment of complex predicates as one phonological word.

[4] See also Megerdooimian (2002: 80–81) for further examples of stress placement and the conclusion that stress facts cannot be used as an argument for the word-like behavior of complex predicates in Persian.

[5] Ghomeshi & Massam (1994: 186) argue that sentences like (2b) and sentences like (3) should be treated parallel and that combinations like *zamin xordan* ‘earth collide’ = ‘to fall’ and *ketâb xaridan* ‘book buy’ = ‘to buy books’ are phonological words.

Intonational Phrases. The stress is assigned to the last phonological phrase in the Intonational Phrase (351). Thus *zad* in (3a) and *harf zad* in (3b) get stress. Since *zad* contains just one phonological word, this word receives stress. *Harf zad* contains two phonological words and according to the stress assignment principle for Phonological Phrases, the first word receives stress.

The fact that *daftar* gets the stress in (3) is explained by assuming that *daftar xaridam* forms a phonological phrase (Kahnemuyipour 2003: 353–354).

### 3.1.2 *Changes in argument structure*

Goldberg (2003: 123) provides the following examples that show that a verb in a Complex Predicate Construction may have an argument structure that differs from that of the simplex verb.

- (4) (a) ketâb-râ az man gereft (simple verb)  
 book-RÂ from me took.3SG  
 ‘She/he took the book from me.’  
 (b) barâye u arusi gereft-am (complex predicate)  
 for her/him wedding took-1SG  
 ‘I threw a wedding for her/him.’  
 (c) \*az u arusi gereftam (complex predicate)  
 from her/him wedding took

Here, while the normal verb allows for a source argument, this argument is incompatible with the complex predicate.

However, these differences in argument structure are not convincing evidence for the word status of the Complex Predicate Construction, since argument structure changes can also be observed in resultative constructions, which clearly involve phrases and should be analyzed as syntactic objects. An example of an argument-structure-changing resultative construction is given in (5):

- (5) (a) Dan talked himself blue in the face. (Goldberg 1995: 9)  
 (b) \*Dan talked himself.

Theories that assume that all arguments are projected from the lexicon have to assume that different lexical items licence (4a) and (4b), but not necessarily that *arusi gereftam* is a single word.

### 3.1.3 *The existence of transitive complex predicates*

Goldberg (2003: 123) argues that the existence of transitive complex predicates suggests that they are zero level entities. She assumes that nominal

preverbs would be treated as a direct object (DO) in a phrasal account, since the preverb has direct object semantics and it does not occur with a preposition.

She points out that there are complex predicates like the ones in (6) that govern a DO.

- (6) (a) Ali-râ setâyeš kard-am  
 Ali-RÂ adoration did-1SG  
 ‘I adored Ali.’  
 (b) Ali Babak-râ nejât dâd  
 Ali Babak-RÂ rescue gave.3SG  
 ‘Ali rescued Babak.’

This means that an account that assumes that the nominal preverb is a direct object would have to assume that the examples in (6) are double object verbs. According to Ghomeshi & Massam (1994: 194–195) there are no double object verbs in Persian and thus the assumption that the preverb is a DO would be an ad hoc solution.

There are two problems with this argumentation: first, there is an empirical problem: As (7) shows, sentences with two objects are possible.

- (7) Maryam Omid-râ ketab-i padaš dad (P. Samvelian, p. c. 2009)  
 Maryam Omid-RÂ book-INDEF gift gave  
 ‘Maryam gave a book to Omid as a present.’

In addition to the two objects in (7) we have the preverb *padaš* ‘gift’. Second, even if the claims regarding double object constructions were empirically correct, the argument is void in any theoretical framework that does not use grammatical functions as primitives, such as Government & Binding (Chomsky 1981) or HPSG (Pollard & Sag 1994). In HPSG the preverb would be treated as the most oblique argument (see Keenan & Comrie 1977 and Pullum 1977 for the obliqueness hierarchy) and the only interesting question would be the relative obliqueness of the preverb in comparison to the accusative object. The most oblique argument is the one that can form a predicate complex with the verb. Similarly, in GB analyses, the preverb would be the argument that is realized next to the verb.

### 3.1.4 *Preverb and light verb resist separation*

#### 3.1.4.1 *Separation by adverbs*

Karimi-Doostan (1997: section 3.1.1.4) and Goldberg (2003: 124) argue that the inseparability of preverb and verb is evidence for zero level status. Goldberg demonstrates the inseparability by discussing the following



examples. (8) shows that an adverb may appear immediately before the verb ('=' marks clitic boundary).

- (8) mašq=am-râ      tond      nevešt-am  
 homework=ISG-RÂ quickly wrote-ISG  
 'I did my homework quickly.'

However, in the case of complex predicates, the adverb may not separate preverb from light verb (9a). Instead, the adverb precedes the entire complex predicate (9b):

- (9) (a) ??rânanidegi tond      kard-am  
          driving      quickly did-ISG  
          Intended: 'I drove quickly.'  
       (b) tond      rânanidegi kard-am  
          quickly driving      did-ISG  
          'I drove quickly.'

The problem with this argumentation is that adjacency does not entail single wordhood. For instance, many researchers analyze German particle verbs in sentence final position as one word. Resultative constructions in German are similar in many respects to particle verb combinations, but the resultative constructions may involve PP predicates, which are clearly phrasal. Both the result phrase and the particle in verb particle constructions may be separated from the verb in final position only under very special conditions, namely if so-called focus movement is involved (see Lüdeling 1997 and Müller 2002 for examples). Thus the resultative construction is a construction involving material that has a clear phrasal status but that usually must be serialized adjacent to the verb. This means that we cannot conclude simply from adjacency that something cannot be phrasal.

As a reviewer points out, the situation is similar for (9a): the sentence is fine in a topicalization context. So, as with German particle verbs, the complex predicate may be separated depending on the information-structural properties of the utterance. Furthermore, non-specific objects in Persian behave exactly the same as complex predicates do as regards separability (Folli et al. 2005: 1391–1392):

- (10) (tond) mašq      (??tond) nevešt-am  
          quickly homework      quickly wrote-ISG  
          'I did homework quickly.'

If wordhood were used as the only means to explain adjacency requirements, it would follow that *mašq nevešt-am* has to be treated as a word.

3.1.4.2 *Separation by DO*

Goldberg (2003: 125) claims that another argument against ‘a phrasal account of the CP [complex predicate] is the fact that the host and light verb resist separation by the DO in the case of transitive CPs, even though DOs normally appear before the verb’. She gives the pair in (11): Instead of the serialization in (11a), the DO appears before the entire CPred as in (11b):

- (11) (a) ??setâyeš Ali-râ kard-am  
 adoration Ali-RÂ did-1SG  
 Intended: ‘I adored Ali.’  
 (b) Ali-râ setâyeš kard-am  
 Ali-RÂ adoration did-1SG  
 ‘I adored Ali.’

However, this argument is not valid either: The serialization of idiom parts in German shows that idiomatic elements have to be serialized adjacent to the verb (with the exception of focus movement for some idioms). If we look at the example in (12a), involving the idiom *den Garaus machen* ‘to kill’, we see a Dat < Acc order. The example in (12b) with Acc < Dat is unacceptable.

- (12) (a) daß er dem Mann den Garaus gemacht hat  
 that he.NOM the man.DAT the GARAUS.ACC made has  
 ‘that he killed the man’  
 (b) \*daß er den Garaus dem Mann gemacht hat  
 that he.NOM the GARAUS.ACC the man.DAT made has

Although the idiom part has to be realized adjacent to the verb, it is not justified to conclude that the whole idiom is a word/zero level category. Rather, one would follow insights of Nunberg, Sag & Wasow (1994) and argue that such idiomatic phrases have a complex internal syntactic structure. Evidence for this assumption is that the idiom *jemandem den Garaus machen* is flexible: it allows passivization, variable verb position (clause-initial and clause-final), and fronting of the phrase *den Garaus*.

What seems to be at stake here is not related to phrasality but rather seems to be a consequence of Behaghel’s Law (1932), that things that belong together semantically tend to be realized together.<sup>6</sup>

[6] Goldberg (1996: 135) cites Bybee (1985) and Haiman (1985) regarding the following constraint:

(i) ICONIC: A tight semantic bond between items tends to be represented by a correspondingly tight syntactic bond.

Apart from this the separation of a complex predicate by arguments and adjuncts seems to be possible, as the following Persian data indicates (see also Vahedi-Langrudi 1996: 26):

- (13) (a) *guš be man ne-mi-kon-e* (Mohammad & Karimi 1992: 197)  
 ear to me NEG-PROG-does-3SG  
 ‘She/he does not listen to me.’ (ear do = listen)  
 (b) *guš dige ne-mi-kon-e* (Mohammad & Karimi 1992: 198)  
 ear no.more NEG-PROG-do-3SG  
 ‘She/he does not listen anymore.’

In (13a) the complex predicate *guš kon* is separated by a PP and in (13b) it is separated by the emphatic particle *dige*.

### 3.1.5 Nominalizations

Vahedi-Langrudi (1996: 9) and Goldberg (2003: 124) argue that the fact that Persian complex predicates interact with derivational morphology is evidence for their zero level status. Goldberg gives the examples in (14) – (15):

- (14) (a) *bâzi kardan* (verb)  
 game do  
 ‘to play’  
 (b) *bâzi-kon* (noun)  
 play-do  
 ‘player’ (as in soccer player)  
 (15) (a) *negah dâštan* (verb)  
 look have  
 ‘to keep’  
 (b) *negah-dâr-i* (noun)  
 look-have-action.of  
 ‘maintenance’

Vahedi-Langrudi (1996: 6, 202–203, 211) and Karimi-Doostan (1997) argue that it is not just the light verb that undergoes nominalization but the whole complex predicate. Evidence for this is the fact that nominalization of the light verb alone is not possible:

- (16) (a) *pazirâ?i kon-ande*  
 entertainment do-er  
 ‘entertainer’  
 (b) \**kon-ande*  
 do-er

If one does not follow the proposals of Distributed Morphology (Marantz 1997) and wishes to maintain the principle of Lexical Integrity (Bresnan &

Mchombo 1995), this is strong evidence that Persian complex predicates can take part in word formation.

Bresnan & Mchombo (1995) and Wiese (1996) argue that morphological processes that seem to include phrasal material can be explained without recourse to allowing syntactic structures in morphology. For instance Bresnan & Mchombo (1995: 194) discuss the examples in (17), which seem to suggest that phrasal material can enter morphology:

- (17) (a) employee of the month program  
 (b) I told you so attitude

Bresnan and Mchombo point out that not all syntactic phrases can be used in such compounds, which would be expected if cases like (17) were truly productive syntactic combinations. The authors argue that the phrasal material is in fact the quotation of a phrase. This is supported by examples that involve phrases from foreign languages:

- (18) (a) a Sturm und Drang romantic  
 (b) a Heil Hitler skinhead  
 (c) a mea culpa look  
 (d) a certain je ne sais quoi quality  
 (e) his zróónat! expression  
 (f) the ich bin ein Berliner speech

In order to account for examples like (18) in a syntactic approach one would have to assume that German grammar is part of English grammar. Alternatively one could assume that the expressions are quotes; but once this option is admitted in the theory, it can be applied to examples like (17) as well.

So, if one sticks to Lexical Integrity, examples like (14) are indeed evidence for the ability of Persian complex predicates to enter the syntax as single words.

### 3.2 *Phrase-like properties*

Putative word-like properties of Persian complex predicates have been discussed in the previous section. It was shown that most of the arguments for wordhood are not conclusive. The only exception is the data from derivational morphology, if one assumes Lexical Integrity (Bresnan & Mchombo 1995). Instead of providing evidence for a treatment as words, the data reviewed above suggests that Persian complex predicates should be treated in the lexicon (if one assumes that argument-structure change is a lexical process). In what follows I discuss phrasal properties of Persian complex predicates.

3.2.1 *Coordination of the preverbs*

Vahedi-Langrudi (1996: 28), Karimi-Doostan (1997: section 3.1.2), and Megerdooian (2002: 65–66) claim that conjoinability can be used as a test to decide whether a complex predicate forms a morphological unit or not. They discuss examples like that in (19) and conclude that the example shows ‘that CPs are in fact syntactic formations, and violate syntactic atomicity attributed to CPs as morphological objects’ (Vahedi-Langrudi 1996: 28).<sup>7</sup>

- (19) tedad-e 120 qayeq towgif\_\_<sub>i</sub> =va mosadere šod<sub>i</sub>  
 number-EZ 120 boat detention =and confiscation become  
 ‘120 boats were detained and confiscated.’

It is true that an analysis that grants syntactic independence to the preverbs can account for sentences like (19) without problems. But coordination should not be used as a test for morphological status, since there are cases such as German prefix verbs that are clearly morphological objects, but that nevertheless allow the coordination of prefixes:

- (20) Man kann die Fahrzeuge hier be- und ent-laden.  
 one can the vehicles here PREFIX and PREFIX-load  
 ‘It is possible to load and unload vehicles here.’

*beladen* and *entladen* are prefix verbs. They differ from particle verbs in that they can never be separated and hence their status as unitary morphological objects is beyond any doubt. Since parts of morphological objects can be coordinated (Höhle 1982: 89–92; Wiese 1992; Artstein 2005), coordinatability should not be regarded as evidence for the phrasal status of elements. The same is true for the prefixes *kam-* and *por-* in the Persian example below (Samvelian, c. 2007):

- (21) ?mosâferat-e kam yâ por-xatar  
 journey-EZ little or full-danger  
 ‘safe or dangerous journey’

*kam-xatar* ‘safe’ and *por-xatar* ‘dangerous’ each forms a morphological unit, but the prefixes may be coordinated.

However, there is another class of coordination examples that do not have a parallel in the domain of prefix verbs. Megerdooian (2002: 65) discusses the sentence in (22), in which the light verb *kardan* has been gapped from the second predicate *ehsas kardan* ‘feeling do’ = ‘to feel’.

[7] EZ stands for Ezafe. The Ezafe is a vowel that is inserted between elements in the nominal domain. See Samvelian (2007) for a discussion of the phenomenon and an analysis that is compatible with the assumptions made in this paper.

- (22) ta diruz ne-mi-tavanest-am [har anče fekr mi-kard-am  
 until yesterday NEG-DUR-can.PST-ISG whatever thought DUR-did-ISG  
 va ehsas — ] boruz dah-am  
 and feeling reveal give-ISG  
 ‘Until yesterday, I couldn’t reveal what I thought or felt.’

In contrast to the cases discussed above, deletion of the verb is not possible in German with prefix verbs, but it is possible with particle verbs, adjective copula combinations, and resultative constructions (Zeller 1999: 57; Müller 2002: 266).

### 3.2.2 *Complex predicates are separated by the future auxiliary*

In Persian the future auxiliary is realized immediately before the main verb, which assumes the form of the past stem. (23a) shows an example with a simple verb. In the case of complex predicates, the auxiliary is realized adjacent to the light verb, i.e., it separates the Xo from the Vo. An example is given in (23b). As (23c) shows, the serialization of the future auxiliary in front of the whole complex predicate is not permitted.<sup>8</sup>

- (23) (a) Ali mard-râ xâh-ad zad (simple verb)  
 Ali man-RÂ FUT-3SG hit.PST  
 ‘Ali will hit the man.’  
 (b) (man) telefon xâh-am kard (complex predicate)  
 I telephone FUT-1SG do.PST  
 ‘I will telephone.’  
 (c) \*(man) xâh-am telefon kard (complex predicate)  
 I FUT-1SG telephone do.PST

The fact that the future auxiliary is inflected is evidence against an analysis that treats the auxiliary as an infix inside of a morphologically complex word since inflection usually applies outside of derivational morphology (Goldberg 1996: 136).

### 3.2.3 *Imperfective prefix and negation*

The indicative/aspectual, subjunctive/imperative, and negative prefixes, *mi-*, *be-*, and *na-* respectively, are attached directly to the light verb, thus intervening between preverb and light verb (Barjasteh 1983: 248). Some of the cases are illustrated by the following examples:

- (24) (a) Ali gerye ne-mi-kon-ad  
 Ali cry NEG-IPFV-do-3SG  
 ‘Ali does not cry.’

[8] It should be noted that the future auxiliary is rarely used in conversational Persian. Instead, the present tense is used to express the basic future.

- (b) Ali dâr-ad                    gerye mi-kon-ad  
 Ali PROG.AUX-3SG cry    IPFV-do-3SG  
 ‘Ali is crying.’

These morphemes may not appear as prefixes on the preverb.

3.2.4 *Complex predicates can be separated by DO clitics*

In the case of simple verbs, direct object clitics appear directly after the verb:

- (25) did-am=aš  
 saw-1SG=3SG  
 ‘I saw her/him/it.’

In the case of complex predicates, the DO clitic can attach to the light verb, but attachment to the preverb as in (26) is slightly preferred. The clitic separates the preverb from the light verb in (26):

- (26) rošan=aš kard  
 light=3SG did  
 ‘She/he turned it on.’ (for instance a radio)

Goldberg (2003: 132) observes that ‘pronominal elements may not appear in the middle of single zero level categories. That is, the clitic cannot occur between syllables in a multisyllabic single word, even after a stressed morpheme boundary. Therefore, the possibility of inserting the pronominal clitic within the CP provides a strong piece of evidence that the host [preverb] and light verb should be analyzed as two separate words’ in (26).

There are, however, languages like European Portuguese that have mesoclitics, which do appear inside of words. For instance, in Portuguese, object clitics may intervene between a verb root and the tense marker in the simple future (Crysmann 2002). However, such mesoclitics are not attested for Persian simplex verbs. The insertion of the clitic into a simplex verb like *didan* ‘to see’ is ungrammatical (both examples due to Megerdooian 2002: 65):

- (27) (a) did-im=eš  
 saw-IPL=3SG  
 ‘We saw her/him/it.’  
 (b) \*did=eš-im  
 saw=3SG-IPL  
 ‘We saw her/him/it.’

The discussion above shows that Persian complex predicates have both word-like properties (interaction with derivational morphology) and phrase-like properties (separation by adverbs, separation by the future auxiliary, separation by clitics). I have shown that some of the criteria for word- or

phrasehood that have been suggested in the literature cannot be maintained; but a discussion of them is necessary to understand the motivation for Goldberg's analysis, which is discussed in the next section.

#### 4. INHERITANCE-BASED CLASSIFICATION OF PHRASAL PATTERNS

In this section I first show how Goldberg analyzes Persian complex predicates and then go on to show where the problems with her analysis lie.

##### 4.1 Goldberg's approach to Persian complex predicates

Goldberg (2003) suggests representing Persian complex predicates in a constructional hierarchy that uses the concept of default inheritance. The default property of Persian complex predicates is that they are words (Vo), but this default property can be overridden if a complex predicate is used together with the future auxiliary. Goldberg assumes a special construction for this case. The Future Complex Predicate Construction inherits from both the Complex Predicate Construction and the Future Construction. The property of the Complex Predicate Construction of having a bar level of zero is overridden by the value of the bar level of the Future Construction, which is one. In what follows I will explain the analysis in more detail. (28) displays the Complex Predicate Construction in the box notation that is typical for CxG.<sup>9</sup>

##### (28) Complex Predicate Construction (CPVo)

Cat: V0
X0 < V0

The ' $<$ ' means that X0 has to immediately precede Vo. Goldberg argues for a constructional status of such X0–Vo combinations, since the combination of X0 and Vo has a meaning that differs from the meanings of the components.<sup>10</sup>

[9] The actual publication of Goldberg's article contains a V' as the Cat value, which is inconsistent with the text and with instances of this construction that are used in later figures in the article. Instead of quoting the book publication, I refer to the version of the paper that is available on Adele Goldberg's webpage. I omit the marking of the stress in the following pictures.

[10] There are large classes of complex predicates that are formed compositionally and are non-idiosyncratic. See section 5.7 for my analysis of compositional preverb-verb combinations.



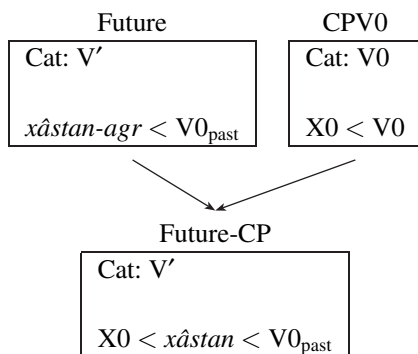
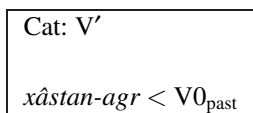


Figure 3

Combination of the Complex Predicate Construction with the Future Construction via multiple inheritance with defaults.

The future auxiliary *xâstan* is placed directly in front of the simplex verb, which has to be in the past form. Goldberg expresses this as follows:

(29) Future Construction



If complex predicates are used in the future, the auxiliary is placed between the  $X_0$  and the  $V_0$  of the complex predicate. Goldberg (2003: 128) argues that this auxiliary should not be treated as an infix, since it shows agreement and agreement information is usually regarded as belonging to inflectional morphology, which is normally applied outside of derivational morphology. Because of this we have a non-predictable morphological fact, which – according to Goldberg – justifies the stipulation of a Future Complex Predicate Construction. The commonalities between this construction and both the Future Construction and the Complex Predicate Construction are captured by a default inheritance hierarchy, which is shown in figure 3.

As is clear from looking at the figure, the Future Construction and the Complex Predicate Construction differ in important features: first, they differ in syntactic category ( $V_0$  vs.  $V'$ ) and second, the linear sequence of construction parts is incompatible as well (*xâstan-agr <  $V_0$*  vs.  $X_0 < V_0$ ) given that ‘<’ stands for immediate precedence in Goldberg’s notation. The respective values are overridden by the inheriting Future Complex Predicate Construction.

For speakers who allow for both clitic positions in (30), Goldberg (2003: 134) assumes the inheritance hierarchy in figure 4.

## Generalization over clitic constructions

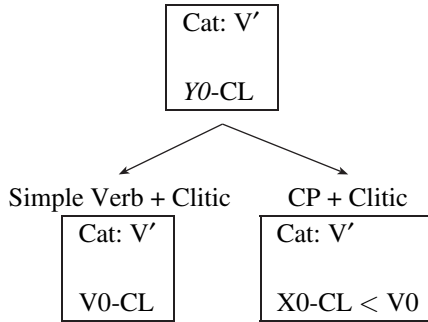


Figure 4  
Clitic constructions.

- (30) (a) masxareh=aš kard-and  
ridicule=3SG did-3PL  
'They made fun of him.'  
(b) masxareh kard-and=aš  
ridicule did-3PL=3SG

(30a) can be analyzed with the CP + Clitic Construction, while (30b) can be analyzed with the Vo-CL Construction, since the Complex Predicate Construction in (28) licenses a Vo, which can enter the Vo-CL Construction.

The interaction of the Clitic Construction with the Complex Predicate Construction is shown in figure 5.

Having sketched Goldberg's proposal, I now discuss several problematic aspects of this analysis.

#### 4.2 Problem 1: Semantics needs embedding

The first problem is a formal problem: the meaning of the complex predicate has to be embedded under the meaning of the future auxiliary, but this cannot be modeled by inheritance since the meaning of the Future-CP sub-construction overrides the meaning of the CPVo construction. Consider figure 6.

The Future Construction has the SEM value future(X), where X stands for the semantic contribution of the verb that is embedded under the future auxiliary. The semantics value of the CPVo construction is Sem<sub>CPVo</sub>, a value that is different from the SEM value of the Vo. This value has to be identified with the X in the Future Construction, but this is not possible since the SEM value of the CPVo construction is overridden in the Future-CP construction. This problem can be solved by introducing an additional bookkeeping feature, a so-called junk feature, as suggested by Kathol (1994: 262) and Koenig (1999: section 3.3) for the inheritance-based analysis of German adjectives

PERSIAN COMPLEX PREDICATES

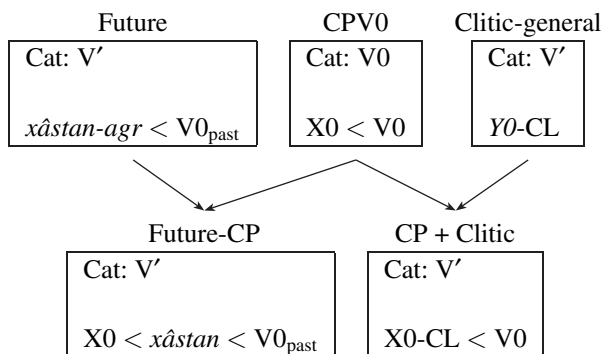


Figure 5  
Part of the hierarchy that includes Future, CP, and Clitics.

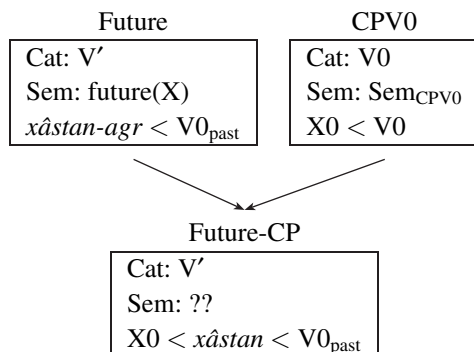


Figure 6  
Inheritance hierarchy for Persian Complex Predicate Constructions including semantics.

and for simulating the effects of argument-adding lexical rules for extra-  
position, respectively. Figure 7 shows an inheritance hierarchy that is aug-  
mented by bookkeeping features that make it possible to specify the future  
semantics.

In the Future Construction the value of SEM-H is identified with the argu-  
ment of the future functor (the identity is signaled by identical numbers  
preceded by '#', this identification is also called *structure sharing*). In the  
CPV<sub>0</sub> construction the semantics of the complex predicate is identified with  
the SEM-H value. The SEM-H values of the sub-construction are inherited from  
both super-constructions and by this the two SEM-H values are identified.  
Since the SEM value of the Future-CP differs from the SEM value of CPV<sub>0</sub>, the  
structure sharing #2 is overridden. However, the SEM-H value is inherited

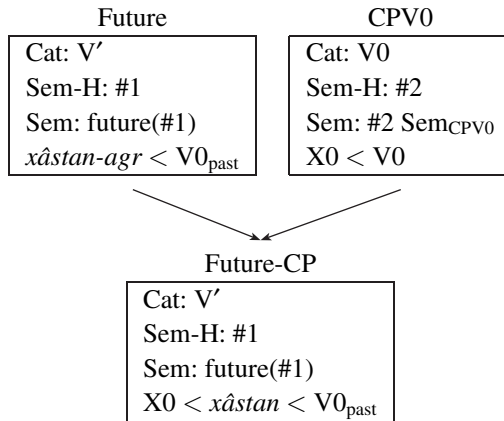


Figure 7

Inheritance hierarchy for Persian Complex Predicate Constructions including semantics and auxiliary features.

from CPV0 and therefore the SEM value of the Future-CP construction is future(Sem<sub>CPV0</sub>).

In the selection-based account that will be developed in section 5, no stipulation of auxiliary features and overriding of structure sharings is necessary.

#### 4.3 Problem 2: Interpretation of negation

Goldberg (2003: section 5.2) assumes that combinations of the negation prefix and light verbs are stored in the lexicon due to their high frequency. She writes that ‘in a usage-based hierarchy, more specific stored forms preempt or block the creation of forms based on a more general pattern. Therefore, the existence of the forms *mi-kardan*, and *na-kardan* block the possibility of adding the prefixes directly to the zero level CP as a whole. This observation is based on the simple assumption that the light verb involved in a CP is recognized as the same verb as its corresponding main verb’. It is unclear how this can be formalized since the meaning of a negated complex predicate is entirely different from the meaning of the negated main verb. In addition, in examples like (31), *kardan* is a part of *kâr kardan* and it is not clear why a stored construction that contains a negated part of another construction (*na-kardan*) should block the negation of the whole construction.

- (31) *kâr na-kard-am* (Barjasteh 1983: 62)  
 work NEG-did-1SG  
 ‘I did not work.’

To account for examples like (31) and idiomatic complex predicates like (32), Goldberg seems to need additional constructions that take care of the embedding of the idiosyncratic semantics under the negation.

- (32) u mard-râ dust na-dâšt  
 she/he man-RÂ friend NEG-have.PST.3SG  
 ‘She/he did not love the man.’

4.4 *Problem 3: Interaction of cliticization and negation*

If we agree that we need additional constructions for negated complex predicates, we need further sub-constructions to deal with the interaction of cliticization and negation as demonstrated in (33):

- (33) u dust=aš na-dâšt  
 she/he friend=3SG NEG-have.PST.3SG  
 ‘She/he did not love him/her.’

4.5 *Problem 4: Interaction of negation and progressive/subjunctive marking*

As was discussed in section 3.2.3, progressive and subjunctive markers are similar to negation in that they can separate the preverb from the light verb. To deal with the progressive and subjunctive markers one could suggest a general Xo Aff-Vo construction that allows a negation affix or the progressive/subjunctive marker in the Aff slot. However, this is not sufficient since the negation affix and the progressive marker can cooccur, as is demonstrated by the example in (24a), repeated here as (34):

- (34) Ali gerye ne-mi-kon-ad  
 Ali cry NEG-IPFV-do-3SG  
 ‘Ali does not cry.’

Since both the negation and the progressive have scope over the contribution of the complete complex predicate, another construction will be needed for cases like (34).

4.6 *Problem 5: Interaction of future and negation*

Yet another construction seems to be needed for the interaction between future and negation since the idiosyncratic semantics of the complex predicate has to be represented somehow.

- (35) râdio-râ guš na-xâh-ad kard  
 radio-RÂ ear NEG-FUT-3SG do.PST  
 ‘She/he will not listen to the radio.’

The idiosyncratic semantics has to be embedded under the future operator and this in turn has to be embedded under the negation.

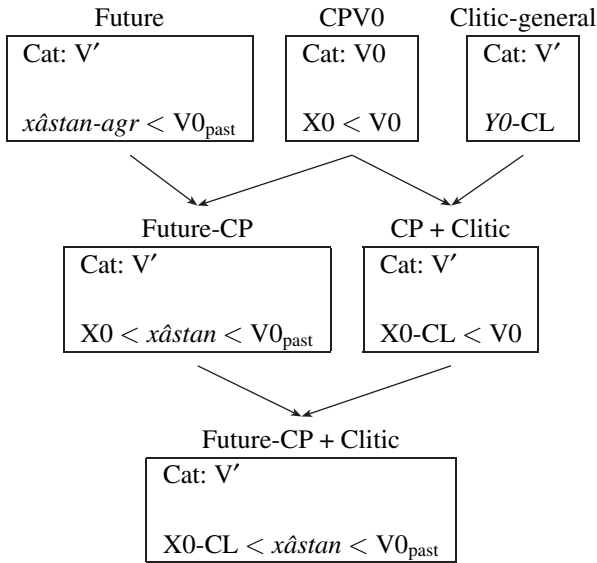


Figure 8

Extended hierarchy that includes Future, CP, and Clitics.

#### 4.7 Problem 6: Interaction between future, clitics and complex predicates

Goldberg (1996: 139) discusses the interaction between future, DO clitics, and complex predicate formation. One of her examples is shown in (36):

- (36) bâz=aş    xâh-am kard  
 open=3SG FUT-1SG do.PST  
 'I will open it.'

However, the example in (36) is not covered by her 2003 analysis. In order to deal with it, she would have to stipulate one more construction, namely the Future Complex Predicate Clitic Construction. The appropriately extended hierarchy is shown in figure 8. Note that the information regarding the serialization in the newly introduced construction is incompatible with the information in both of the super-constructions.<sup>11</sup> This means that this

[11] Jochen Trommer (p.c. 2006) remarks that one could allow for discontinuous constituents in constructions. The boxes for Future and CPV0 in figure 8 would then mean that the construction dominates *xâstan-agr* and *V0*, or *X0* and *V0*, respectively. The '<' would now stand for precedence instead of immediate precedence (see Bergen & Chang (2005: 156) for linearization statements in Embodied Construction Grammar). Such constraints would not cause a conflict during inheritance. (In fact, they could be stated independently of the construction, as is done in GPSG and HPSG.) Similar proposals have been made by Kathol (1995: 244–248) and Crysmann (2002: section 4.2) for German particle verbs. Allowing for discontinuous constituents in a grammar, however, is a very powerful extension of the basic

information has to be specified stipulatively, thereby overriding the constraints on the superordinated constructions. It is important here to point out that Goldberg assumes that the information in the CPVo construction is the default information. This information is overridden by both the Future+CP and the CP+Clitic constructions. This means that the latter constructions contain non-default information and it is unclear which construction will win if conflicting information is inherited from both of them.

Note also that the valence of the Future+CP construction and the CP+Clitic construction differs: since the clitic saturates one argument slot, the valence requirements of the respective linguistic objects differ. One must ensure that the valence information is inherited from the CP+Clitic construction and not from the Future+CP construction. There might be ways to formalize this if one assumes formal underpinnings similar to those assumed in LFG; but if one assumes the formalism of HPSG or Sign-based Construction Grammar, the inheritance of valence information causes a conflict which can only be resolved if valence information is default information in general and the result of the inheritance is explicitly represented in the Future-CP+Clitic construction.

According to my informants the order in (36) is not the only one possible. In Classical/Literary Persian, the clitic may also be attached to the future auxiliary. (37) shows an example that corresponds to (36).

- (37) *bâz xâh-am=aš kard*  
 open FUT-1SG=3SG do.PST  
 'I will open it.'

While data from Classical Persian is not relevant for a synchronic grammar, grammatical theory has to be able to account for it in principle. In section 5.6 I show how a small change in the specification of the lexical entry of the future auxiliary can account for the difference between Classical and Modern Persian.

In order to get this example in a pattern-based approach, one would need another construction that is parallel to the Future-CP+Clitic construction in figure 8 but differs in the order of the elements.

---

mechanisms for the combination of syntactic material, which is not really needed (Müller 2005). Apart from this there is fronting data that shows that particles in fronted position may form complex constituents, which cannot be explained in purely surface-oriented linearization-based approaches (Müller 2007b: section 18.3). Note that all the other problems that are mentioned in this paper are also problematic for analyses that allow for discontinuous constituents.

But this is not all. Goldberg (1996: 139) discusses the examples in (38), where (38b) repeats (36) above:

- (38) (a) ?bâz xâh-am kard=aš  
 open FUT-ISG do.PST=3SG  
 ‘I will open it.’  
 (b) bâz=aš xâh-am kard  
 open=3SG FUT-ISG do.PST

She rates the first item with two question marks.<sup>12</sup> According to my informants (38a) is fine and some even prefer it stylistically. To account for (38a) we have to introduce a new construction, since the complex *bâz xâham kard* is a *V'* and therefore cannot be combined with the clitic via the *Vo-CL* construction. Combining the light verb of the complex predicate with the clitic before the formation of the complex predicate is problematic for two reasons: firstly, complex predicate formation may change the argument structure of the verb so that we do not know about the *DO* of the whole complex, and secondly, the result of combining say *kard* and *-aš* would be *V'* and therefore this *V'* could not enter into the future constructions, which require a *Vo* in the past form. This means Goldberg would have to assume a *Xo* < *xâstan* < *Vo<sub>past</sub>-CL* construction.

In section 4.6, I discussed the interaction between future and negation, but of course there is also interaction between clitics and future and negation, adding three additional patterns:

- (39) (a) bâz=aš na-xâh-am kard  
 open=3SG NEG-FUT-ISG do.PST  
 (b) ??bâz na-xâh-am=aš kard  
 open NEG-FUT-ISG=3SG do.PST  
 ‘I will not open it.’  
 (c) ?bâz na-xâh-am kard=aš  
 open NEG-FUT-ISG do.PST=3SG  
 ‘I will not open it.’

(39a) is said to be the canonical order, but the other patterns are also attested.

#### 4.8 Problem 7: How is the Complex Predicate Construction related to stems?

As Goldberg (2003: section 3.3) notes, Persian complex predicates enter into nominalizations. But this means that one does not have to list *Xo Vo*

[12] In Goldberg (2003: 135) she writes ‘speakers only accept the clitic on the host as in ... and reject outright examples like ... in which the clitic appears after the light verb and future tense separates host and light verb’. She marks an example that corresponds to (38a) with ‘\*’.



combinations with an idiosyncratic meaning, but rather Xo V-stem combinations, since the nominalization applies to stems, not to fully inflected verbal elements. Since the Future-CP construction needs to inherit from the CPVo construction, there has to be a CPV-stem construction that is the heir of the CPVo construction. This means that inflection has to be done in the inheritance hierarchy. While this is possible in principle, it is not possible to do derivation in inheritance hierarchies: Krieger & Nerbonne (1993) observe that recursion as in the German *Vorvorvorvorversion*<sup>13</sup> ‘prepreprepreversion’ cannot be covered in inheritance networks. Since information about the prefix *vor-* is contained in *Vorversion*, inheriting a second time from *vor-* would not add anything. To go back to our introductory example in figure 1, having two connections between *printing device* and *printer* would not change anything. If we say twice that an object has certain properties, nothing is added.

Secondly, in an inheritance-based approach to derivation, it cannot be explained why *undoable* has the two readings that correspond to the two bracketings in (40), since inheriting information in different orders does not change the result.

- (40) (a) [un- [do -able]] ‘not doable’  
 (b) [[un- do] -able] ‘capable of being undone’

Again referring to figure 1, it does not matter if we first say that a *copy machine* is a *printing device* and then add that it is a *scanning device* too, or if we provide this information in the reverse order. The consequence is that one needs a morphology component in which stem affix combinations form new objects which in turn can be combined with further affixes.<sup>14</sup> With such a morphology component the two readings in (40) and the recursion in *prepreversion* is unproblematic.

Discussing inheritance-based and lexical rule-based approaches to morphology, Koenig (1999) argues for adopting a uniform analysis of inflection and derivation for reasons of parsimony. A uniform treatment of derivational and inflectional morphology is incompatible with Goldberg’s approach.

One way to avoid the problem of relating stems and phrasal constructions is to assume a word-based morphology (Becker 1993, Haspelmath 2002). Since the complex predicates are Vos in Goldberg’s proposal they could be related to other Vos by rules that map words to words. But note that such a

[13] <http://www.sgaf.de/viewtopic.php?t=21551&postdays=0&postorder=asc&start=0> (25 September 2006).

[14] This data does not discriminate between Realizational Morphology (which is usually adopted in HPSG and is modeled with lexical rules), stem/affix models, and Distributed Morphology (DM). The important point is that one needs recursion and embedding, which are not available in inheritance networks.

proposal would have problems for languages like German that allow discontinuous derivations. In German it is possible to derive nouns by circumfixing a verb stem with *Ge-* *-e*. For instance, *Gerenne* ‘repeated running’ is derived from *renn-*. If the word-based approach took the particle verb *herumrenn* ‘running around’ as input, we would expect a form \**Geherumrenne* but the correct form is *Herumgerenne* ‘repeated instances of aimless running events’. This could be fixed by assuming special word-based rules for derivations involving complex predicates, but that would add to the cost of the proposal under discussion, while the lexical proposal can account for the interaction of discontinuous derivations with complex predicates in a straightforward way (see section 5.8).

#### 4.9 Problem 8: Separation of complex predicates

The examples in (13) – repeated here as (41) – pose a serious problem for Goldberg’s analysis, since her construction would state that *guš* ‘ear’ and *nemi-kon-e* ‘NEG-PROG-DO-3SG’ are adjacent.

- (41) (a) *guš be man ne-mi-kon-e* (Mohammad & Karimi 1992: 197)  
 ear to me NEG-IPFV-DO-3SG  
 ‘She/he does not listen to me.’ (ear do = listen)
- (b) *guš dige ne-mi-kon-e* (Mohammad & Karimi 1992: 198)  
 ear no.more NEG-PROG-DO-3SG  
 ‘She/he does not listen anymore.’

Goldberg restricts the focus of her analysis to so-called inseparable complex predicates (122). This class is defined as the class in which the preverb cannot be combined with a determiner. The complex predicate *guš kardan* belongs to this class. Because the phrasal analysis does not allow for the discontinuous realization of preverb and verb, the examples in (41) pose a problem for it.

Since Goldberg works in the framework of Construction Grammar and since Construction Grammar is a non-transformational framework (Goldberg 1995: 7, 2006: 205), Goldberg cannot derive the serialization in (41) by a transformation that turns an underlying phrasal configuration in which *guš* and *nemikone* are adjacent into the configuration observed in (41).

The only option that remains (apart from allowing discontinuous realizations of constructions; see footnote 11 above) is to simulate movement/transformations by a threading mechanism such as that used in GPSG (Gazdar 1981) or by a similar technique. Usually an empty element is assumed and the information about the missing element is percolated up in the syntactic tree and bound off by a filler at a higher node. An alternative is to avoid empty elements and introduce the information about missing elements in a phrasal configuration or lexically. To get the interpretation of complex predicates right, proponents of phrasal approaches would have to assume

	Pattern	Construction	Example
1.	Xo + Vo	CPVo	(2b)
2.	Future + Simple Verb	Future	(23a)
3.	Xo + Future + Vo	Future-CP	(23b)
4.	Simple Verb + Clitic	Simple Verb + Clitic	(25)
5.	Xo + Clitic + Vo	CP + Clitic	(26)
6.	Negation + Simple Verb	<i>na-Vo</i>	
7.	Xo + Negation + Vo		(32)
8.	Xo + CL + Negation + Vo		(33)
9.	Xo + Negation + Progr + Vo		(34)
10.	Xo + Negation + Future + Vo		(35)
11.	Xo + CL + Future + Vo		(38b)
12.	Xo + Future + CL + Vo		(37)
13.	Xo + Future + Vo + CL		(38a)
14.	Xo + CL + Negation + Future + Vo		(39a)
15.	Xo + Negation + Future + CL + Vo		(39b)
16.	Xo + Negation + Future + Vo + CL		(39c)

Table 1

Complex predicates in interaction with other constructions

special constructions for cases in which the preverb is not adjacent to the verb or the future auxiliary. If fronting of the preverb turns out to be possible with all of the constructions discussed so far, this means that one will have to double the number of constructions.

#### 4.10 Summary

Taking stock for a moment, the set of constructions given in table 1 are needed. Table 1 lists the constructions proposed by Goldberg (pp. 1–6) and those that have to be stipulated in addition. I have left out the imperfective and subjunctive constructions, which are parallel to the negation case. For each case that allows a fronted preverb, another construction has to be added (see section 4.9).

Apart from the proliferation of constructions, the problems with morphology (section 4.8) and with the separability of complex predicates (section 4.9) remain unsolved: derivational morphology cannot be dealt with in inheritance networks and in order to account for the separability, one would have to considerably modify the analysis and allow for discontinuous words or stipulate additional phrasal constructions to account for separated complex predicates. As was shown in section 4.2, semantic embedding poses a

problem for inheritance-based proposals and so-called junk features would have to be assumed to derive the correct semantic representations.

Having rejected the phrasal inheritance-based proposal, I now turn to my own analysis.

## 5. THE LEXICAL ANALYSIS

In what follows I present an analysis of Persian complex predicates that has none of the problems that were discussed in the previous section. I assume that in Complex Predicate Constructions, the verb is a head that selects the preverb. In the case of non-compositional complex predicates the head of the complex predicate is present in the lexicon with special selectional properties and an appropriate meaning. Compositional complex predicates are licensed by different lexical entries for light verbs, which contribute semantically in some cases. The commonalities between light verbs can be captured in an inheritance hierarchy. Although inheritance is used to capture generalizations, the problems of the phrasal approach are avoided since it is lexical rather than phrasal material that is classified in the hierarchy and morphology is not done via inheritance.

As a framework for the analysis, I use Head-driven Phrase Structure Grammar (HPSG; Pollard & Sag 1994), but nothing hinges on this. One could cast the analysis in Categorical Grammar or in LFG. To facilitate comparison, I work out the analysis at a level of detail that matches that of Goldberg's analysis. For a more detailed analysis of Persian complex predicates the reader is referred to a forthcoming book (Müller & Samvelian in preparation). I introduce the basic assumptions in section 5.1. section 5.2 discusses the representation of lexical entries that take part in complex predicate formation and have a lexicalized, idiomatic meaning; section 5.3 provides the grammar rules that license head argument structures and predicate complexes and demonstrates how simple sentences involving complex predicates can be analyzed. Section 5.4 is devoted to predicate complexes with the future auxiliary, section 5.5 discusses negation, section 5.6 deals with clitics, section 5.7 provides an analysis of non-idiomatic complex predicates, and section 5.8 presents an account of nominalization.

### 5.1 *Basic assumptions*

HPSG is a sign-based theory in the sense of Saussure (1916), that is, all linguistic objects are form-meaning pairs. A certain phonological representation is paired with a semantic representation. HPSG uses typed feature structures for the representation of all linguistic objects (lexical entries, lexical rules, syntactic rules). A lexical head comes with a valence list that contains descriptions of elements that have to be combined with this head in order to yield a maximal projection. For instance, the ditransitive verb *dâdan*

‘to give’ selects for a nominative NP, an accusative NP and a prepositional object.

- (42) man ketâb-â-ro be Sepide dâd-am  
 I book-PL-RÂ to Sepide gave-ISG  
 ‘I gave Sepide the book.’

The valence list of *dâdam* ‘to give’ can be represented in an abbreviated form as in (43):

- (43) COMPS <NP[nom], NP[acc], PP[be]>

NP[nom], NP[acc], and PP[be] are shorthand for feature descriptions. I assume binary branching structures. Figure 9 shows the analysis of the example in (42).

The boxes with numbers mark identity of the respective information, that is, the description in the COMPS list of a head has to be compatible with the properties of the object that is combined with the head. The lexical entry for *dâdam* is combined with all its arguments in binary branching structures. Elements that have been saturated are not contained in the COMPS value of the mother node. For example *be Sepide dâdam* ‘to Sepide gave’ is the result of combining *dâdam* with its PP argument and therefore the PP argument is not contained in the COMPS list of the node that represents the complete linguistic object *be Sepide dâdam*. If a phrase (or word) has an empty COMPS list, it is fully saturated and hence a maximal projection.

Since Persian is a language with rather free constituent order (Karimi 2005), I assume that the subject of finite verbs is represented on the list of arguments, as has been suggested by Kiss (1995), Pollard (1996), Kathol (2000), De Kuthy (2002: section 4.4), and Müller (2002, 2007b) for German. This explains why the subject of finite verbs can be realized between other arguments. The authors cited above assume that the arguments in the COMPS list can be combined with the head in any order, which accounts for their permutability. Hence (44), which is a linearization variant of (42), can be analyzed with no further assumptions.

- (44) be Sepide ketâb-â-ro man dâd-am (Karimi 2005: 17)  
 to Sepide book-PL-RÂ I gave-ISG  
 ‘As for the books, TO SEPIDE I gave them.’ or  
 ‘As for Sepide, THE BOOKS I gave her.’

See Gunji (1986) for a similar approach regarding scrambling in Japanese, and Fanselow (2001) for an equivalent proposal for German in the framework of the Minimalist Program.

HPSG immediate-dominance schemata do not constrain the linearization of the daughters. This is taken care of by linearization rules (linear precedence or LP rules). The following rather trivial rules say that complements

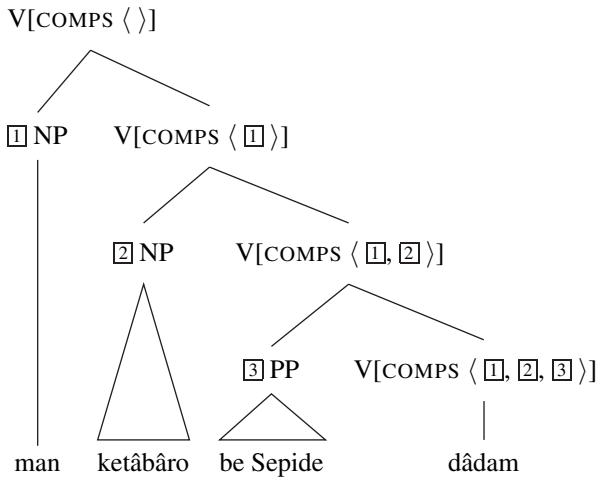


Figure 9  
Analysis of *man ketâb-â-ro be Sepide dâd-am* ‘I gave Sepide the book.’

follow their head if the head is marked as head-initial and that they precede their head if the head is marked as non-initial:

- (45) (a) HEAD [INITIAL +] < COMPLEMENT  
(b) COMPLEMENT < HEAD [INITIAL -]

Because Persian is an SOV language, verbs are marked as INITIAL – and hence the arguments are realized to the left of the verb as shown in figure 9.<sup>15</sup> Since INITIAL is a head feature, it is projected along the head path, thus ensuring that all verbal projections in Persian are INITIAL –.<sup>16</sup> Another linearization rule regarding the elements in the predicate complex is discussed in section 5.4.

So far nothing has been said about the representation of semantic information. There are various proposals in the HPSG literature, but here I use Situation Semantics (Barwise & Perry 1983, Ginzburg & Sag 2000) since it is

[15] Clitics are an exception to this rule since they follow simplex verbs and may follow the predicate complex.

[16] A linearization rule that orders lexical heads after their arguments (Pollard & Sag 1994: 172) would not apply to the daughters *ketâbâro* and *be Sepide dâdam* in figure 9 since *be Sepide dâdam* is not a word. Pollard & Sag’s linearization constraint requires a flat structure since in flat structures the lexical head is a sister of all arguments and hence its arguments can be ordered with respect to the respective word. Note also that the approach using the INITIAL feature works without further assumptions for cases of coordination in which two verbs are coordinated and the coordinated structure – which is phrasal – functions as a head:

(i) Kim [[knows and loves] this record].

sufficient for current purposes.<sup>17</sup> (46) shows the relevant features of the lexical entries for *dâdam* ‘give’:<sup>18</sup>

(46) Lexical entry for *dâdam* ‘give’ first person singular

PHON	⟨	<i>dâdam</i>	⟩								
HEAD		<i>verb</i>									
COMPS	⟨	NP <sub>[1]</sub> , NP <sub>[2]</sub> , PP <sub>[3]</sub>	⟩								
CONT		<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="padding-right: 5px;">ARG1</td> <td style="border: 1px solid black; padding: 2px 5px;">1</td> </tr> <tr> <td>ARG2</td> <td style="border: 1px solid black; padding: 2px 5px;">2</td> </tr> <tr> <td>ARG3</td> <td style="border: 1px solid black; padding: 2px 5px;">3</td> </tr> <tr> <td></td> <td><i>give</i></td> </tr> </table>	ARG1	1	ARG2	2	ARG3	3		<i>give</i>	
ARG1	1										
ARG2	2										
ARG3	3										
	<i>give</i>										

The value of PHONOLOGY (PHON) is the phonological representation of the word. The value of HEAD is its part of speech, and the value of CONTENT (CONT) is the main semantic contribution. The semantic indices of the NP and PP arguments are referred to by the lowered boxes [1], [2], and [3]. These indices are identified with the argument roles of the *give*' relation. Figure 10 shows the syntax–semantics interaction. The linking is established in the lexical entry (46). Since the elements in the valence list get identified with the NP and PP arguments, it is ensured that the referential indices of the NPs and the PP are identified with the arguments of the *give*' relation. The main contribution of each lexical head is projected to its mother node by the Semantics Principle (Pollard & Sag 1994: 56) and therefore the main semantic contribution of the whole phrase is identical to the contribution specified in the lexical entry ([4] in figure 10). The semantic contribution of NPs and PPs is not represented in figure 10 since it is irrelevant for the present discussion. See Pollard & Sag (1994) and Copestake, Flickinger, Pollard & Sag (2005) for a discussion of NP and PP semantics and quantifier scope in HPSG.

[17] In Müller & Samvelian (in preparation) we use *Minimal Recursion Semantics* (MRS; Copestake et al. 2005) since MRS allows for the underspecification of scope relations. See <http://hpsg.fu-berlin.de/Fragments/Persian/> for an implemented fragment of Persian that uses MRS.

[18] I assume a feature geometry which is basically that introduced in Pollard & Sag (1994: chapters 1–8). This means that the features HEAD and SUBCAT are part of the feature description that is the value of SYNSEM|LOCAL|CAT and that CONT is part of the description under SYNSEM|LOC. The feature SYNSEM was introduced to enforce locality of selection in the feature geometry, and LOCAL separates information that is locally relevant from information that plays a role in non-local dependencies. Since the respective phenomena are not discussed in this paper, I omit these features in the rest of the paper.

See Müller (2007b) and Richter & Sailer (2009) for an argument against adopting more recent feature geometries such as, for instance, that suggested in Sign-Based Construction Grammar (Sag 2007a).

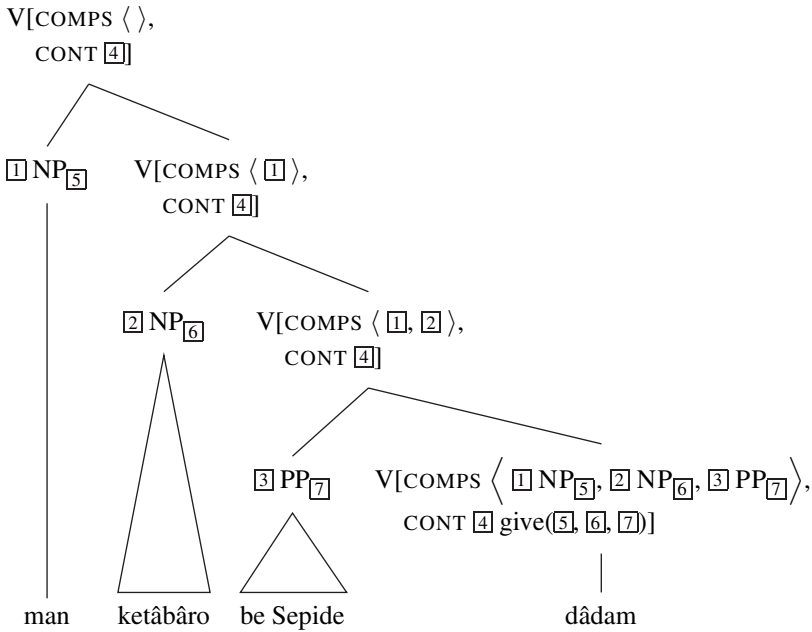


Figure 10  
 Analysis of *man ketâb-â-ro be Sepide dâd-am* ‘I gave Sepide the book.’ including semantic information.

5.2 The lexical representation of verbs in non-compositional complex predicates

The representation for the verb that is used in the non-compositional complex predicate *dust dâštan* ‘to like/love’ – as discussed by Goldberg (2003: 139) and Karimi (1997: 279) – is given in (47) in abbreviated form:<sup>19</sup>

(47) Lexical entry for past stem of *dâštan* as used in *dust dâštan* ‘to like/love’

PHON	⟨ <i>dâšt</i> ⟩						
COMPS	⟨ NP <sub>1</sub> , NP <sub>2</sub> , N[LEX+, LID <i>dust</i> ] ⟩						
CONT	<table border="1" style="border-collapse: collapse; width: 80%; text-align: left;"> <tr> <td style="padding: 2px 5px;">ARG1</td> <td style="padding: 2px 5px;">1</td> </tr> <tr> <td style="padding: 2px 5px;">ARG2</td> <td style="padding: 2px 5px;">2</td> </tr> <tr> <td style="padding: 2px 5px;"></td> <td style="padding: 2px 5px;"><i>love</i></td> </tr> </table>	ARG1	1	ARG2	2		<i>love</i>
ARG1	1						
ARG2	2						
	<i>love</i>						

The PHON value contains information on how the stem is pronounced. The COMPS list contains descriptions of the syntactic and semantic properties of the elements that have to be combined with *dâštan* in order to project a

[19] See section 5.7 for the treatment of compositional Complex Predicate Constructions.



complete phrase. I follow Vahedi-Langrudi (1996: 14) in assuming that all preverbs are subcategorized arguments of the respective light verbs. Therefore, the COMPS list contains two full NPs and a lexical noun (LEX+) which is identified as *dust*. LID stands for *lexical identifier* (Sag 2007a). The specification in (47) ensures that the noun to be combined with *dâstan* is *dust* and not something else.

The value of CONT corresponds to the semantic content of the stem when used together with the selected element *dust*. The first two elements in the COMPS list are linked to the arguments of *love*. The boxed numbers convey identity of values. NP<sub>[1]</sub> is a shorthand for a feature description of a noun phrase with the referential index [1]. Mohammad & Karimi (1992: 203) show that the noun in complex predicates is non-referential but has a predicate-like character. This is captured in (47): the referential potential of the noun is not used, but the information about the predicate is directly incorporated into the verb.

The lexical entry in (47), and the way it is used in the analyses that will be presented shortly, are basically a formalization of what Goldberg (1996: section 8.3) calls the *Idiomatic Argument Analysis*. Similar suggestions can be found in Krenn & Erbach (1994), Nunberg et al. (1994) for the analysis of idioms and in Müller (2003) for the analysis of particle verbs.

### 5.3 Simple forms

First, consider the most simple case in which the components of the complex predicate appear adjacent to each other as in (48):

- (48) u        mard-râ dust    dâšt  
           she/he man-RÂ friend have.PST.3SG  
           ‘She/he loved the man.’

In this case the stem entry in (47) has to be inflected, whereupon it can be combined with *dust* to form *dust dâšt*. The analysis of (48) is shown in figure 11. The combination of *dâšt* with its arguments is exactly parallel to the combination of *dâdam* ‘give’, that was shown in figure 9. The CONT value of the lexical item for *dâšt* is projected along the head path by the Semantics Principle (Pollard & Sag 1994: 56); therefore the main semantic contribution of the whole phrase is identical to the contribution specified in the lexical entry for *dâšt* given in (47), or rather the inflected version of this lexical entry.

The account presented here is lexical in the sense that the structure of a clause is determined by the lexical head. This does not entail that the complex predicate is formed in the lexicon, that is, the combination of *dust dâšt* is licensed by the lexical entry for *dâstan*. This account thus differs from the one suggested by Barjasteh (1983: 225), who assumes that the combination of *dust* and *dâšt* is done in the lexicon. Such lexicalist accounts were criticized

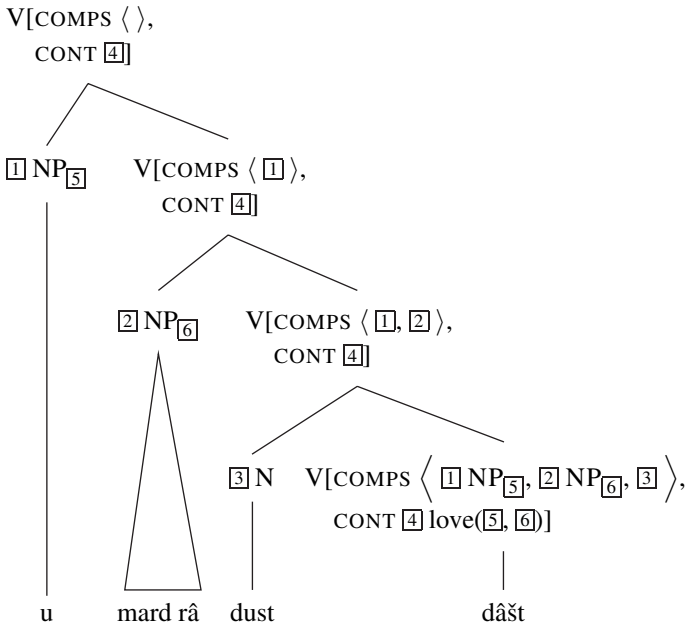


Figure 11  
Analysis of *u mard-râ dust dâšt* ‘She/he loved the man.’

by Embick (2004), but Embick’s criticism does not apply to accounts of the type presented here (see Müller 2006: 874–875 for discussion).

#### 5.4 Future forms

The future forms can be analyzed by assuming a lexical entry for the future auxiliary that makes use of the technique of argument attraction introduced into the framework of HPSG by Hinrichs & Nakazawa (1989).<sup>20</sup>

(49) *xâstan* ‘will’, future tense auxiliary:

HEAD	<i>verb</i>				
COMPS	$\bar{1} \oplus \langle V[\textit{past}, \textit{FLIP+}, \textit{COMPS} \bar{1}, \textit{LEX+}]:\bar{2} \rangle$				
CONT	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 5px;">ARG1</td> <td style="padding: 5px;"><math>\bar{2}</math></td> </tr> <tr> <td style="padding: 5px;"></td> <td style="padding: 5px;"><i>future</i></td> </tr> </table>	ARG1	$\bar{2}$		<i>future</i>
ARG1	$\bar{2}$				
	<i>future</i>				

This verb *xâstan* governs another verb in past form and attracts all arguments of this verb ( $\bar{1}$ ). The specification of the LEX value of the embedded verb

[20] This argument attraction approach is an adaption of earlier work in the framework of Categorical Grammar (Geach 1970).



future auxiliary embeds the semantic contribution of the embedded verb under the future operator and since the semantic contribution gets projected from the head, the semantic contribution of the mother node of *xâhad dâšt* is future(love( $\bar{5}$ ,  $\bar{6}$ )). This value is projected along the head path and therefore is also the main semantic contribution of the complete phrase.

The future auxiliary differs from auxiliaries like *budan* ‘to be’ in that it precedes the verbal complex that it embeds (Karimi-Doostan 1997: 179). This situation is known from Dutch and German, which are verb-final languages as well but which do not have a strictly verb-final verbal complex. While German verbal complexes are verb-final in general, there are certain situations in which verbs that are part of the verbal complex can precede the verbs they embed. For example, the perfect auxiliary *hast* ‘have’ precedes *gewinnen helfen* ‘win help’ in (51):

- (51) daß du uns diese Schlacht hast gewinnen helfen  
 that you us this battle have win help  
 ‘that you helped us win this battle’ (Haftka 1981: 723)

Hinrichs & Nakazawa (1994) develop an analysis for German Auxiliary Flip constructions that accounts nicely for the rather complex German patterns. Verbs have a feature called FLIP. The value of this feature is + if the governing verb has to be serialized to the left of the verb and – if the verb is linearized to the right. (52) shows the linearization rules that ensure the proper serialization:

- (52) (a) HEAD < [FLIP+]  
 (b) [FLIP–] < HEAD

The FLIP value of the verb that is embedded under *xâstan* is + since this is required by the specification in the lexical entry (49). The LP rule in (52a) ensures that the future auxiliary *xâstan* is serialized to the left of *dâšt* when it is combined with *dâšt*. Figure 12 shows an example analysis in which the future auxiliary appears between the parts of the complex predicate.

Thus far nothing has been said about the agreement between the future auxiliary and the subject of the embedded complex predicate. Here, I follow Müller (2002: 369), who assumes that a finite verb agrees with the first NP with structural case on its valence list, if there is any. In the case of subjectless predicates the verb shows default agreement. This agreement principle can account for the agreement facts in languages like English, German, Hindi, Persian, and Spanish. In sentences with future auxiliaries and complex predicates, the subject of the embedded complex predicate ( $\bar{1}$  in figure 12) is raised to the COMPS list of the future auxiliary. The agreement principle ensures that the future auxiliary agrees with its subject; thus, since the auxiliary’s subject is raised and is therefore identical to the subject of the complex predicate, the future auxiliary agrees with the subject of the complex predicate.

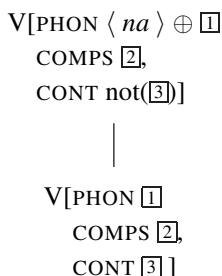


Figure 13  
Negation lexical rule depicted as unary branching tree.

### 5.5 Imperfective prefix and negation

I assume the lexical rule in figure 13 for the treatment of negation.<sup>21</sup>

This lexical rule takes a finite verb as input and licences a lexical item that has the phonology of the input item ( $\boxed{1}$ ) prefixed with *na-*. The syntactic properties of the output ( $\boxed{2}$ ) are identical to the syntactic properties of the input of the lexical rule. The semantic contribution of the input of the lexical rule ( $\boxed{3}$ ) is embedded under the negation relation.

The lexical rule-based approach makes the same predictions as a word syntax approach that assumes a negation morpheme which is combined with a verb in a binary branching structure. The negation prefix would be the semantic head in the binary branching structure and would contribute the negation relation embedding the contribution of the verbal element. This relation is represented directly in figure 13. Despite their similarity in the case at hand, there are other areas of morphology in which the two approaches differ. In the lexical rule-based approach there is no need to assume subtractive morphemes, abstract morphemes for Ablaut and similar things, or portmanteau morphemes. See, for instance, Pollard & Sag (1987: section 8.2), Orgun (1996), Ackerman & Webelhuth (1998), Riehemann (1998), Kathol (1999), Koenig (1999), and Müller (2002) for lexical rule-based approaches to morphology in HPSG and the motivation for these approaches.<sup>22</sup>

[21] See Meurers (2000) on various conceptions of lexical rules and on formalization in the framework of HPSG. The formalization of lexical rules that is usually used in HPSG nowadays is equivalent to a unary branching phrase-structure rule. The only difference between lexical rules and unary branching phrase-structure rules is that the former are applied before syntax.

[22] For non-HPSG-based approaches, see e.g. Dowty (1979: 304), Stump (1991), Aronoff (1994).

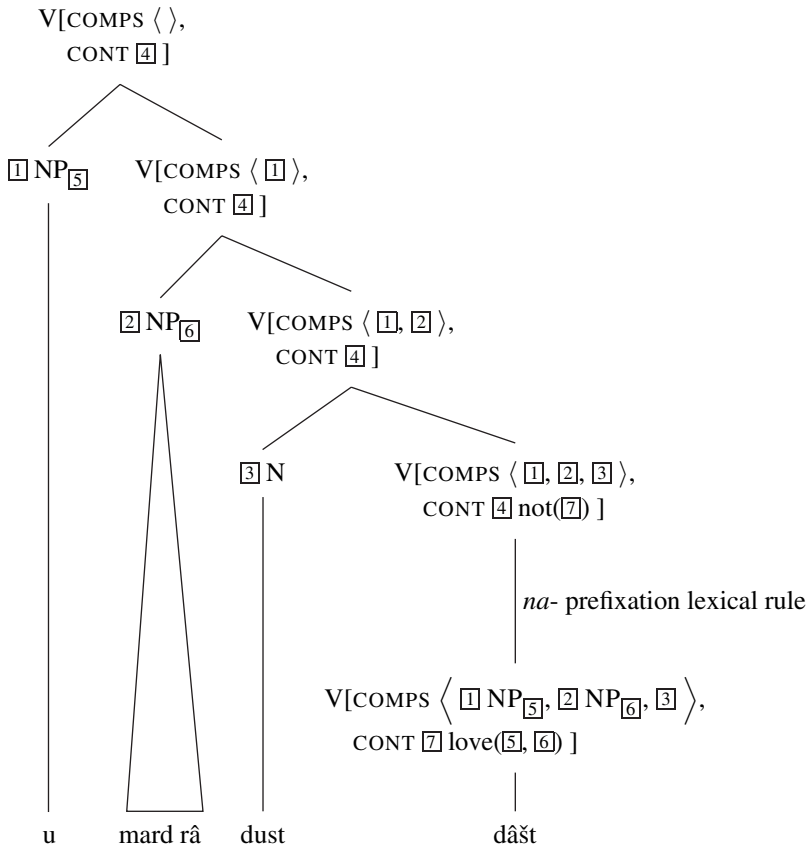


Figure 14  
 Analysis of *u mard-râ dust na-dâšt* ‘She/he did not love the man.’

The negation lexical rule can apply to simplex verbs and to the verbal part of complex predicates. The interaction between negation and non-compositional complex predicates is the most interesting one, since the negation scopes over the contribution of the whole complex predicate despite being realized in the middle of the predicate. I therefore discuss how the negation lexical rule interacts with the complex predicate *dust dâštan*. The lexical rule applies to the inflected form of the lexical item in (47) and licenses *na-dâšt* with the semantics  $\text{not}(\text{love}([1], [2]))$ , where [1] and [2] are linked to the NP arguments of *dust dâštan*. The respective items for *na-dâšt* may be used both in past and in future sentences. The details of the analyses are shown in figures 14 and 15, respectively. These figures are parallel to figures 11 and 12, the only difference being *na-* prefixation and the presence of the negation in the semantic contribution of the respective element.

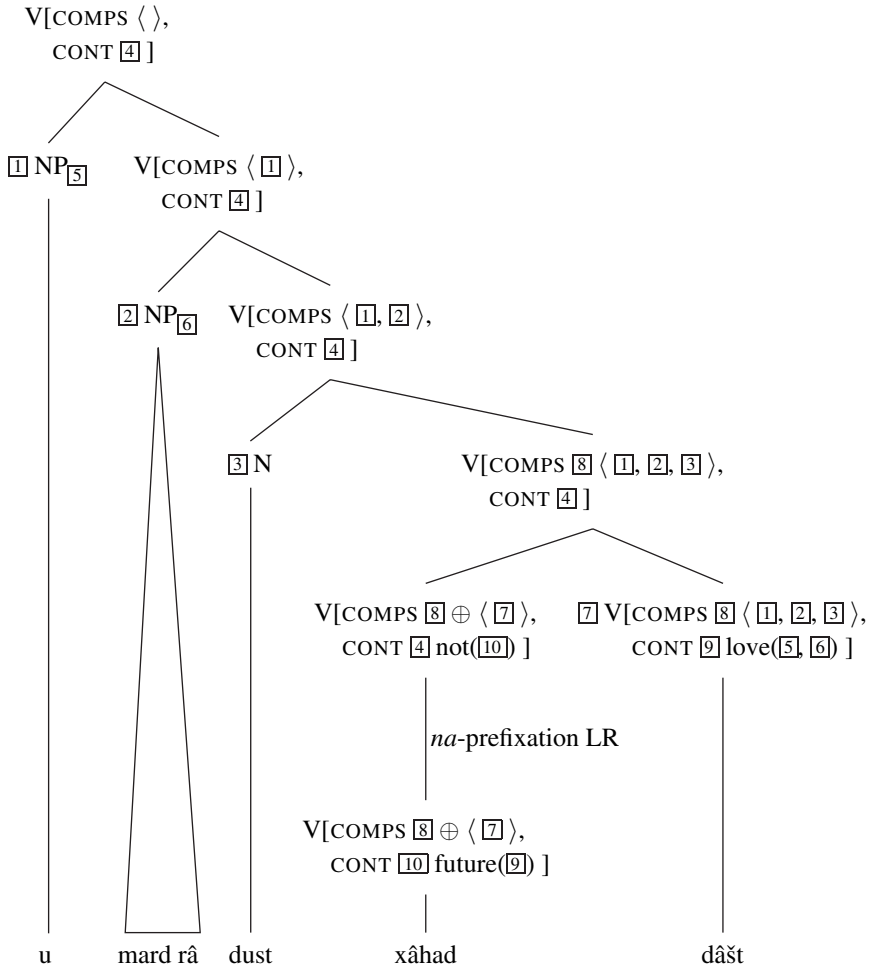


Figure 15  
 Analysis of *u mard-râ dust na-xâhad dâšt* ‘She/he will not love the man.’

5.6 Direct object clitics

The various clitic patterns can be accounted for in the following way. A clitic is attached to a verb that selects an argument that corresponds to this clitic. For instance, the verb *see* in (25) – repeated here as (53) – takes two arguments. The lexical clitic rule maps the verb *didam* with two arguments onto *didam*=*aš*, which does not select a direct object argument.

- (53) *did-am=aš*  
 saw-1SG=3SG  
 ‘I saw her/him/it.’

Persian also allows for clitic left dislocation as in (54):

- (54) Ali-râ did-am=eš  
 Ali-râ saw-1SG=3SG  
 ‘Ali, I saw him.’

Such cases of dislocation have been analyzed in HPSG grammars for Greek (Alexopoulou & Kolliakou 2002), Spanish (Bildhauer 2008), and Maltese (Müller 2009b). Due to space limitations I cannot present the details here.

In case of complex predicates, the clitic is realized either to the left of the verb as in (30a) or to the right of the verb as in (30b). The respective examples are repeated here as (55) for convenience:

- (55) (a) masxareh=aš kard-and  
 ridicule=3SG did-3PL  
 ‘They made fun of him.’  
 (b) masxareh kard-and=aš  
 ridicule did-3PL=3SG

While (55b) is analyzed by the rule discussed above, which attaches the clitic to the right of the light verb, in the analysis of (55a), I assume that *aš* is combined with the light verb by the Head Complement Schema. In the analysis suggested here, *kardand* is a light verb that selects *masxareh* and two NPs. One of the arguments is realized by the clitic.

Similarly, the Head Complement Schema licenses the clitic in cases like (56):

- (56) bâz=aš xâh-am kard  
 open=3SG FUT-1SG do.PST  
 ‘I will open it.’

Since the future auxiliary raises the arguments of the verb it embeds, the arguments of *kard* are accessible in the lexical entry of *xâham*. The technique used here is the same as that used by Miller & Sag (1997) and Monachesi (1998) to account for clitic climbing in French and Italian, respectively.

The lexical entry for the future auxiliary given in (49) requires the embedded verb to be LEX+. Since clitics are combined with their head verb via the Head Complement Schema, the LEX+ specification excludes clitics that are attached to the future auxiliary. However, such examples are possible in Classical/Literary Persian. The respective example in (37) is repeated here as (57):

- (57) bâz xâh-am=aš kard  
 open FUT-1SG=3SG do.PST  
 ‘I will open it.’

This order can be accounted for by relaxing the requirement that the embedded verbal element must be LEX+.



Interactions between negation and cliticization are accounted for since I assume that the negation prefix attaches to a verb without changing its valence properties. The clitic can be realized before or after any verb in whose valence list it appears.

### 5.7 *Compositional complex predicates*

So far I have mainly discussed idiomatic preverb-verb combinations. However, there is a large class of complex predicates involving a noun as preverb that are compositional and formed according to productive patterns. Karimi-Doostan (1997: 181) subdivides complex predicates with a nominal preverb into three sub-classes: those with non-predicative nouns (*guš kardan* ‘ear do’=‘to listen’, with verbal nouns (VN) (*ʔanjâm dâdan* ‘performing give’=‘to perform’), and with process nouns (*râhnamâʔi kardan* ‘advice do’=‘to advise’). Although I have worked out an analysis for all three cases, due to space limitations it cannot be presented here. It may suffice to say that I assume lexical items for the light verbs that select the nominal preverb. In the case of verbal nouns the arguments are selected for by the nominal preverb. I assume that the light verb attracts the arguments of the preverbal element (compare the argument-attraction analysis of future, above). The semantic content is taken over from the embedded verbal noun and combined with additional information regarding tense and aspect, which is contributed by the light verb. This analysis is similar to the argument-transfer analysis suggested by Grimshaw & Mester (1985), which is adopted by Mohammad & Karimi (1992) for Persian. The analysis sketched here differs from Grimshaw and Mester’s in its being rooted in the lexical entries for the light verbs. The way in which the arguments are attracted is specified in the lexical entry of the light verb and is not due to a rearrangement of arguments by a lexical rule or some other process. The analysis suggested here can account for the syntactic properties of Persian CPs while maintaining the assumption that all arguments are projected from the lexicon.

Instead of going into the details of the analysis of complex predicates with nominal preverbs, I will discuss causative and inchoative constructions with adjectival preverbs in more detail. As Vahedi-Langrudi (1996: section 4.1.5) and Karimi-Doostan (1997: section 3.3) point out, the combination of an adjective with *kardan* should be treated as a causative construction. In addition to the combination with *kardan*, there is an inchoative variant with *šodan*:

- (58) (a) *bâz kardan*  
 open make  
 ‘to open’ (transitive)  
 (b) *bâz šodan*  
 open become  
 ‘to open’ (intransitive)

Such combinations can be accounted for by the following lexical entries for *kardan* and *šodan*, respectively:

(59) Lexical entry for past stem of *kardan* as used in *bâz kardan*

PHON	⟨ <i>kard</i> ⟩
COMPS	⟨ NP <sub>[1]</sub> , [2], ADJ[SUBJ ⟨ [2] ⟩]:[3] ⟩
CONT	cause([1], become([3]))

This lexical entry selects an adjective which requires a subject ([2]). This subject is identified with the second element in the COMPS list of *kardan*, that is, it is raised to the object position of the verb. In addition to this argument, *kardan* selects a subject NP which is coindexed with the first argument of the *cause'* relation ([1]). The other argument of *cause'* is the *become'* relation, which in turn embeds the semantic contribution of the adjective ([3]).

The lexical entry for the inchoative variant differs only in not adding an additional argument and not contributing a *cause'* relation:

(60) Lexical entry for past stem of *šodan* as used in *bâz šodan*

PHON	⟨ <i>šod</i> ⟩
COMPS	⟨ [2], ADJ[SUBJ ⟨ [2] ⟩]:[3] ⟩
CONT	become([3])

It is important to note here that neither entry specifies the form of the adjective the verb has to be combined with. This is an important difference between these compositional cases and the cases like *dust dâštan* ‘friend have’ = ‘to like/love’) that were discussed in section 5.2.

The lexical entries for light verbs that are part of compositional complex predicate formations and for light verbs that form non-compositional complex predicates are organized in a type hierarchy that captures the commonalities between classes of Complex Predicate Constructions. In this way it is possible to capture the fact that even idiomatic complex predicate formations share properties with other idiomatic complex predicates that involve the same light verb or a light verb belonging to the same class. For instance the fact that complex predicates that are formed with *dâštan* are stative can be inherited from a general type that describes all stative light verbs. The type for the light verb that is used in *dust dâštan* is a further specification of a more general, stative *dâštan* type since the form of the preverb and the idiosyncratic semantic contribution are specified. See Family (2006) for a discussion of verbal islands in Persian and a suggestion for representing the semantic space of complex predicates in networks.

Construction grammarians often assume a continuum between lexicon and grammar. It is unclear what this means formally since every linguistic object is either a phrasal construction or a lexical item, but in the lexical approach presented here, the general idea is captured: there are lexical items

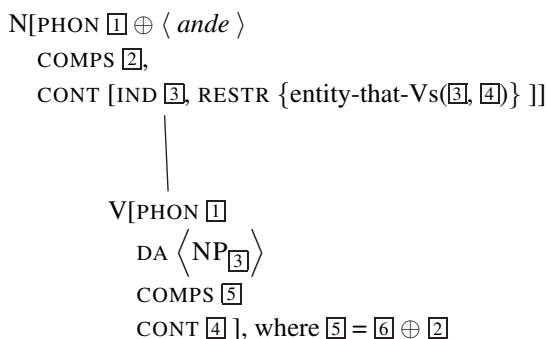


Figure 16  
Lexical rule for agentive nominalizations.

in which a lot of information regarding the arguments is specified (phonological information, selectional restrictions, pragmatic conditions) and there are others with fewer restrictions. Lexical items can be compared with regard to those values that are fixed, and hence it makes sense to talk about a continuum.

### 5.8 Nominalizations

The nominalization facts can be analyzed by letting the derivation affixes combine with the verbal stem, and carrying out the combination with the preverb after this combination. I assume the lexical rule in figure 16, which maps a verbal stem onto a nominal stem with appropriate semantics and adds the respective material to the PHON value of the output of the lexical rule.

The lexical rule applies to verbs that have an agent argument. I assume that lexical representations of verbs contain a pointer to their designated argument, if they have one. The designated argument is the subject of transitive and unergative verbs. Unaccusative verbs do not have a designated argument. The lexical rule requires the verb it applies to to have a designated argument and identifies the referential index of this designated argument (the index of the agent =  $\boxed{3}$ ) with the referential index of the resulting noun. The semantics of the verbal element ( $\boxed{4}$ ) is embedded in the restrictions of the index in the result of the lexical rule application. The valence list is split into two parts, one containing full phrases ( $\boxed{6}$ ) and one containing descriptions of signs that take part in complex formation ( $\boxed{2}$ ). The respective constraints on the members of  $\boxed{2}$  and  $\boxed{6}$  are not given in the figure. The realization of arguments is optional in general, but the complex-forming dependents have to be realized. This is captured by having  $\boxed{2}$  as the value of COMPS. Cases that license other arguments are not covered by the rule, but this is for expository reasons only.

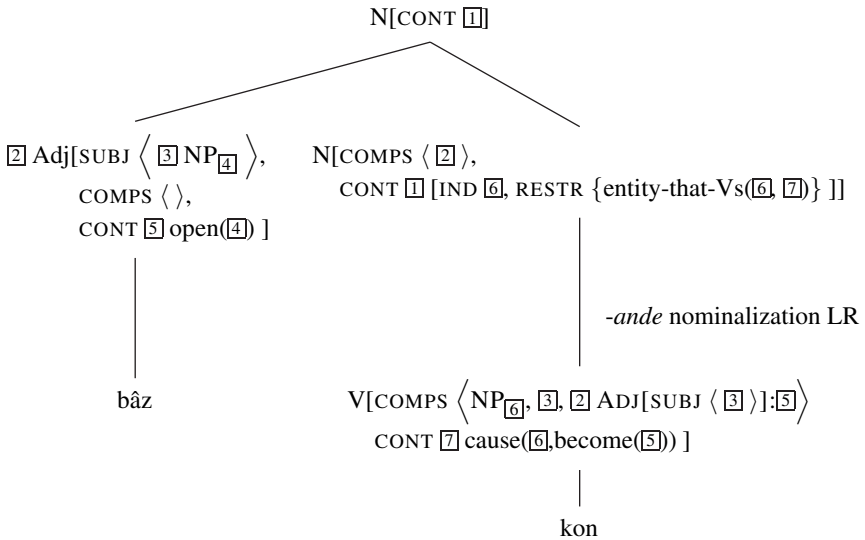


Figure 17  
Analysis of *bâz-kon-ande* ‘opener’.

If the verb does not have dependents that take part in complex formation, 2 is the empty list. Therefore the lexical rule applies both to simplex verbs and verbs that take part in complex predicate formation.

Figure 17 shows the analysis of the nominalization in (61).

- (61) *bâz-kon-ande-gân hanuz nay-âmade-and*  
 open-do-er-PL yet NEG-come-3PL  
 ‘The openers have not come yet.’

The lexical rule takes the semantics of the input stem (7) and embeds it under the *entity-that-Vs'* relation. The other argument of *entity-that-Vs'* (6) is linked to the subject (agent) of the embedded verb. The semantic contribution of the nominalization is the semantic index 6 together with the corresponding set of restrictions that contains the *entity-that-Vs'* relation. The arguments that take part in complex predicate formation have to be realized in the nominal environment as well. This is indicated by the 2 in the COMPS list of the derived noun. The other arguments are optional. In figure 17 they are not presented at the COMPS list of the noun. Since *konande* is the head in the structure for *bâz-konande*, the semantic content (1) is passed up to the mother node from *konande*.

The lexical rule for nominalization shares an important feature with the lexical rules for negation and for imperfective/subjunctive marking. The affixes are combined with the verbal element directly. The preverb is combined with the result of the lexical rule application. As was pointed out in Müller (2003), such a treatment of complex predicates also solves the alleged



- (64) (a) *didan-i*  
 see-worthy  
 ‘worthy to see’  
 (b) *pardâxt-kardan-i*  
 payment-do-worthy  
 ‘possible/worthy to pay’  
 (c) *darmân-kardan-i*  
 cure-do-possible  
 ‘possible to cure’

Vahedi-Langrudi concludes that this indicates that the affixations apply to the complex verbal heads in Persian. However, an alternative analysis which is in the spirit of the analysis presented above is possible for *-i* derivation. Müller (2003) suggests an analysis for similar derivations with the German suffix *-bar*: the suffix attaches to a stem that contains information about the particle and about the arguments that are contributed by the particle. That is, all information about the argument structure of the complete particle verb is already present in this lexical item in an under-specified form. *-bar* imposes the respective constraints, suppresses the subject of the verb and promotes the object to subject. Due to space limitations I will not provide the analysis of the Persian *-i* derivation here, but point to the formation of passive participles and *-bar* derivation that was discussed in Müller’s work.

### 5.9 Summary

In this section I have proposed a lexical treatment of complex predicates in Persian. The preverb is treated as a dependent of the base verb. This treatment can account for the separability of preverb and verb in syntax and morphology: the preverb can be fronted, the future auxiliary can intervene between preverb and verb, clitics can be placed between preverb and verb. Since the syntactic and semantic properties of the whole complex predicate are contained in the lexical item for the verbal stem already (although not fully specified in the case of productive complex predicate formations), an analysis is possible in which morphological processes that are sensitive to information provided by both parts of the complex predicate can affect the verbal stem. Such an analysis covers both the Persian cases and the German examples involving discontinuous derivation morphemes that attach to both sides of the verbal part of a complex predicate.

## 6. GOLDBERG’S ARGUMENTS AND COMPARISON WITH GOLDBERG’S PROPOSAL

Goldberg claims that an analysis like that which was developed in section 5 cannot account for the zero level properties of Persian complex predicates.

I will show in this section that this claim is not warranted, since there are many similar constructions in other languages which have been successfully dealt with by the Idiomatic Argument Analysis. In what follows I discuss Goldberg's arguments in detail.

### 6.1 *Separability*

Goldberg (2003: 141) claims that the Idiomatic Argument Analysis predicts that 'Persian CPs should be generally separable as are general DO + verb combinations in Persian. As we have seen, however, CPs are only separable in certain specific circumstances'. This is not the case, however; as was pointed out in the data discussion, the property of being realized together is a property that many idioms have despite their being clearly phrasal. Linearization constraints ensure that material that is part of an idiomatic expression is serialized together. For SOV languages like Persian, Dutch, or German this means that idiom parts are serialized clause-finally next to the verb. The same holds for complex predicates in Persian and particle verbs in German and Dutch.

Linearization is a complex matter and it is known from German that idiom parts may be extracted, but also rearranged locally under certain conditions. Since none of the constituent order rules for NPs in German is strict, researchers such as Uszkoreit (1987) suggest violable linearization rules. A grammar with such violable linearization rules can account for the examples in (11) – repeated here as (65) – and even explain the markedness of the first sentence.

- (65) (a) ??setâyeš Ali-râ kard-am  
 adoration Ali-RÂ did-1SG  
 'I adored Ali.'  
 (b) Ali-râ setâyeš kard-am  
 Ali-RÂ adoration did-1SG  
 'I adored Ali.'

Goldberg's account does not assign a structure to sentences like (65a) and hence cannot explain why they exist and why they are marked.

### 6.2 *Selection of certain properties*

Goldberg (2003: 141) further argues against the Idiomatic Argument Analysis by claiming that it makes necessary the selection of unusual properties. She writes:

the light verb would have to select, not only for the semantic type of its argument (which would be unremarkable), but also for its definiteness and specificity characteristics: the hosts must be both indefinite and non-specific. These characteristics usually mark the particular noun's role in

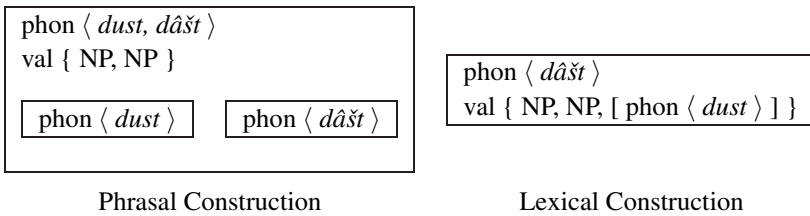


Figure 19

Comparison of the phrasal and the lexical approach.

discourse and are not specified by the verb. That is, we do not generally find unique stems in a language that are differentiated only by the definite/specificity characteristics of their arguments – such specifications are not typically part of a verb’s meaning.

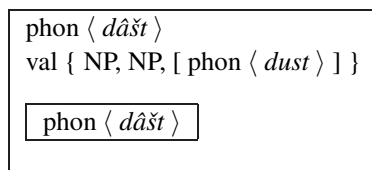
Goldberg’s argument is not valid, however, since the two approaches are entirely equivalent as regards the features of preverbs and verbal elements that have to be stated. In Goldberg’s analysis the constraints are stated in two boxes inside of a third box as in the case of *dust dâštan* ‘to love/like’, while in the Idiomatic Argument Analysis the same constraints would be stated inside of the special lexical entry for *dâštan* that selects *dust*. Figure 19 shows this graphically in a CxG-compatible form. The value of PHON is a list that contains information about the phonological realization of a linguistic object and the value of VAL is a set that contains all the arguments (Kay & Fillmore 1999). The constraints on the parts of the constructions are exactly the same in the two frameworks, but they are stated at different places.<sup>23</sup> Similar pictures can be drawn for compositional cases of complex predicates. So everything that Goldberg states in her phrasal constructions can be stated in the lexical construction. If she states restrictions on definiteness in the phrasal pattern, the same can be formulated as a restriction on the kind of argument that is selected. Neither way of formulating the restriction is more ad hoc or motivated than the other.

If one rejects the approach that assumes multiple related lexical items for light verbs, one can still assume an approach that is equivalent to the approach discussed in section 5. A set of related lexical rules would be applied to light verbs. This can be depicted as in figure 20. The light verb *dâšt* enters the lexical construction and licences an item that selects for the preverb.<sup>24</sup> Only this item provides the idiosyncratic meaning of *dust dâštan*. In a further

[23] See also Kay & Fillmore (1999: 19). The authors remark that they could have stated their phrasal *What’s X Doing Y* construction as a lexical entry.

[24] See also Kay (2005) for a lexical rule-based treatment of adjuncts in the framework of Construction Grammar.





*Figure 20*

Lexical approach with separate lexical entry for the light verb.

step the construction in figure 20 is combined with the preverb *dust*. Such a templatic approach would capture the proposal of Goldberg (2003) and Family (2006) that preverb and light verb enter into a certain configuration and by doing so a meaning emerges that differs from the meaning contributions of the parts.

If one compares this approach to the one that was suggested in section 5, it is clear that the facts are captured in both proposals. However, the approach just sketched uses one more lexical item: the one for the light verb. So instead of classifying lexical rules/templates like that in figure 20 in a hierarchy, one can classify lexical entries directly, as was suggested in section 5.

### 6.3 Learnability

Some proponents of Construction Grammar and Cognitive Grammar argue that the way children learn language is evidence for phrasal constructions since children acquire certain phrasal patterns first (Dąbrowska 2001: 85–86; Tomasello 2003: 107; 2006a, b; Goldberg, Casenhier & Sethuraman 2005). The same remark as in the previous section can be made in reply to this argument: the phrasal and the selectional approach are indistinguishable as far as learnability properties are concerned (see also Müller in preparation). In the first case the learner learns a certain pattern and a corresponding meaning; in the second case the learner learns valence classes or idioms that select certain other elements and that have a special meaning if they appear together with the selected element. The only thing that is required in the lexical approach in addition to what is required in the phrasal approach is knowledge about the fact that two elements may be combined at all, i.e. that there are functor argument structures. I consider this a very straightforward thing to acquire.

Tomasello (2005) argues that semantic and pragmatic information plays an important role in language acquisition but this does not differentiate between the phrasal and the lexical analysis if one assumes that syntactic, semantic and pragmatic information is contained in the representation of

linguistic objects. Tomasello's treatment of language acquisition is compatible with the lexical analysis of Persian complex predicates.

## 7. CONCLUSION

In this paper, I have discussed Persian complex predicates, which pose an interesting problem since they have both word-like and phrase-like properties. It has been shown that an analysis that relies on normal selection and embedding is more appropriate than an inheritance-based phrasal one. The inheritance-based analysis has to stipulate special future constructions for all non-compositional complex predicates. Therefore, it claims implicitly that all realizations of complex predicates with the future auxiliary have an idiosyncratic aspect, which is not true. The inheritance-based approach has problems accounting for the semantics of the future construction and negation since semantics requires embedding and to model such an embedding relation in inheritance hierarchies requires auxiliary features. If several embeddings interact, a complex system using several auxiliary features and complicated value identifications would be necessary. None of this is needed in the selection-based approach, which allows for semantic embedding in a straight-forward way.

This paper does not argue against Construction Grammar *per se*. In fact the analysis presented here is fully compatible with the main tenet of Construction Grammar: all linguistic objects are form-meaning pairs. The analysis addresses central issues in the treatment of complex predicates. All analyses must capture the fact that some complex predicates have idiosyncratic argument structures (see section 3.1.2). Goldberg is right to point out that approaches that establish a strong connection between form and meaning capture these facts directly. My point is that working in a non-transformational framework, all phenomena that interact with argument structure changes, rearrangement of constituents, and derivational morphology should be treated with base lexical entries + morphological rules + combinatorial syntactic schemata, rather than by the classification of phrasal patterns. Proposals in the framework of Construction Grammar that share this view can be found in Boas & Sag (in preparation) and Sag (2007b).

I do not want to end this paper without mentioning that there are phenomena for which phrasal constructions seem to be the only option if one does not want to stipulate empty heads. Sag (1997), Borsley (2006), Jacobs (2008), and Müller & Lipenkova (2009) discuss several cases.

The analysis that was outlined in this paper has been implemented in the TRALE system (Meurers, Penn & Richter 2002, Penn 2004, Müller 2007a) as part of a grammar fragment which uses a core grammar for German, Danish, Maltese, and Mandarin Chinese. The respective grammars can be downloaded at <http://hpsg.fu-berlin.de/Software/>.

## APPENDIX

## List of abbreviations used in example glosses

= marks clitic boundary	DUR = durative	PL = plural
1 = 1st person	EZ = Ezafe	PL = plural
2 = 2nd person	FUT = future	scaps>prfx = prefix
3 = 3rd person	INDEF = indefinite	PROG = progressive
ACC = accusative	IPFV = imperfective	PST = past
AUX = auxiliary	NEG = negation/negative	SG = singular
DAT = dative	NOM = nominative	

## REFERENCES

- Ackerman, Farrell & Gert Webelhuth. 1998. *A theory of predicates* (CSLI Lecture Notes 76). Stanford, CA: CSLI Publications.
- Alexopoulou, Theodora & Dimitra Kolliakou. 2002. On Linkhood, Topicalization and Clitic Left Dislocation. *Journal of Linguistics* 38.2, 193–245.
- Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes* (Linguistic Inquiry Monographs 22). Cambridge, MA: MIT Press.
- Artstein, Ron. 2005. Coordination of parts of words. *Lingua* 115, 359–393.
- Barjasteh, Darab. 1983. *Morphology, syntax and semantics of Persian compound verbs: A lexicalist approach*. Ph.D. dissertation, University of Illinois at Urbana–Champaign.
- Barwise, Jon & John Perry. 1983. *Situations and attitudes*. Cambridge, MA: MIT Press.
- Becker, Thomas. 1993. Back-formation, cross-formation, and ‘bracketing paradoxes’ in paradigmatic morphology. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology*, vol. 6, 1–25. Dordrecht: Foris.
- Behaghel, Otto. 1932. *Die deutsche Syntax: Eine geschichtliche Darstellung*, vol. 4: *Wortstellung: Periodenbau*. Heidelberg: Carl Winters.
- Bergen, Benjamin K. & Nancy Chang. 2005. Embodied Construction Grammar in simulation-based language understanding. In Jan-Ola Östman & Miriam Fried (eds.), *Construction Grammars: Cognitive grounding and theoretical extensions*, 147–190. Amsterdam & Philadelphia: John Benjamins.
- Bierwisch, Manfred. 1987. A structural paradox in lexical knowledge. In Elke van der Meer & J. Hoffmann (eds.), *Knowledge-aided information processing*, 141–172. Amsterdam: Elsevier.
- Bildhauer, Felix. 2008. *Representing information structure in an HPSG grammar of Spanish*. Ph.D. dissertation, Universität Bremen.
- Boas, Hans C. & Ivan A. Sag (eds.). In preparation. *Sign-based Construction Grammar*. Stanford, CA: CSLI Publications.
- Borsley, Robert D. 2006. Syntactic and lexical approaches to unbounded dependencies. *Essex Research Reports in Linguistics* 49, 31–57. [http://www.essex.ac.uk/linguistics/publications/errl/errl\\_49.pdf](http://www.essex.ac.uk/linguistics/publications/errl/errl_49.pdf) (14 October 2009).
- Bresnan, Joan & Sam A. Mchombo. 1995. The Lexical Integrity Principle: Evidence from Bantu. *Natural Language & Linguistic Theory* 13, 181–254.
- Bybee, Joan. 1985. *Morphology*. Amsterdam & Philadelphia: John Benjamins.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 4.3, 281–332. <http://lingo.stanford.edu/sag/papers/copestake.pdf> (11 October 2006).
- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Crysmann, Berthold. 2002. *Constraint-based coanalysis: Portuguese cliticisation and morphology–syntax interaction in HPSG* (Saarbrücken Dissertations in Computational Linguistics and Language Technology 15). Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz und Universität des Saarlandes.

- Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.
- Dabir-Moghaddam, Mohammad. 1997. Compound verbs in Persian. *Studies in the Linguistic Sciences* 27.2, 25–59.
- Dąbrowska, Ewa. 2001. From formula to schema: The acquisition of English questions. *Cognitive Linguistics* 11.1–2, 83–102.
- De Kuthy, Kordula. 2002. *Discontinuous NPs in German*. Stanford, CA: CSLI Publications.
- Dowty, David R. 1979. *Word meaning and Montague Grammar* (Synthese Language Library 7). Dordrecht: Reidel.
- Embick, David. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35.3, 355–392.
- Family, Neiloufar. 2006. *Exploration of semantic space: The case of light verb constructions in Persian*. Ph.D. dissertation, EHESS, Paris.
- Fanselow, Gisbert. 2001. Features,  $\theta$ -roles, and free constituent order. *Linguistic Inquiry* 32.3, 405–437.
- Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and idiomacity in grammatical constructions: The case of *let alone*. *Language* 64.3, 501–538.
- Folli, Raffaella, Heidi Harley & Simin Karimi. 2005. Determinants of event type in Persian complex predicates. *Lingua* 115, 1365–1401.
- Francis, Elaine J. & Laura A. Michaelis (eds.). 2003. *Mismatch: Form–function incongruity and the architecture of grammar* (CSLI Lecture Notes 163). Stanford, CA: CSLI Publications.
- Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12, 155–184.
- Geach, Peter Thomas. 1970. A program for syntax. *Synthese* 22, 3–17.
- Ghomeshi, Jila & Diane Massam. 1994. Lexical/syntactic relations without projection. *Linguistic Analysis* 23.3–4, 175–217.
- 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. 1995. *Constructions: A Construction Grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago: University of Chicago Press.
- Goldberg, Adele E. 1996. Words by default: Optimizing constraints and the Persian complex predicate. *Berkeley Linguistics Society (BLS)* 22, 132–146. <http://roa.rutgers.edu/files/415-0900/roa-415-goldberg-2.pdf> (18 October 2009).
- Goldberg, Adele E. 2003. Words by default: The Persian Complex Predicate Construction. In Francis & Michaelis (eds.), 117–146.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language* (Oxford Linguistics). Oxford: Oxford University Press.
- Goldberg, Adele E., Devin Casenhier & Nitya Sethuraman. 2005. The role of prediction in construction learning. *Journal of Child Language* 32, 407–426.
- Grimshaw, Jane & Ralf-Armin Mester. 1985. Complex verb formation in Eskimo. *Natural Language & Linguistic Theory* 3.1, 1–19.
- Gunji, Takao. 1986. Subcategorization and word order. In William J. Poser (ed.), *The Second International Workshop on Japanese Syntax*, 1–21, Stanford, CA: CSLI Publications.
- Haftka, Brigitta. 1981. Reihenfolgebeziehungen im Satz (Topologie). In Karl Erich Heidolph, Walter Flämig & Walter Motsch (eds.), *Grundzüge einer deutschen Grammatik*, 702–764. Berlin: Akademie.
- Haiman, John. 1985. *Natural Syntax: Iconicity and erosion* (Cambridge Studies in Linguistics 44). Cambridge: Cambridge University Press.
- Haspelmath, Martin. 2002. *Understanding morphology* (Understanding Language Series). London: Arnold.
- Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989. *Subcategorization and VP structure in German* (Aspects of German VP Structure, SfS-Report-01-93). Tübingen: Eberhard-Karls-Universität.
- Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In Nerbonne et al. (eds.), 11–38.
- Höhle, Tilman N. 1982. Über Komposition und Derivation: Zur Konstituentenstruktur von Wortbildungsprodukten im Deutschen. *Zeitschrift für Sprachwissenschaft* 1, 76–112.
- Jacobs, Joachim. 2008. Wozu Konstruktionen? *Linguistische Berichte* 213, 3–44.

- Kahnemuyipour, Arsalan. 2003. Syntactic categories and Persian stress. *Natural Language & Linguistic Theory* 21.2, 333–379.
- Karimi, Simin. 1997. Persian complex verbs: Idiomatic or compositional? *Lexicology* 3.2, 273–318.
- Karimi, Simin. 2005. *A Minimalist approach to scrambling: Evidence from Persian* (Studies in Generative Grammar 76). Berlin & New York: Mouton de Gruyter.
- Karimi-Doostan, Gholamhossein. 1997. *Light verb constructions in Persian and Kurdish*. Ph.D. dissertation, University of Essex.
- Kathol, Andreas. 1994. Passives without lexical rules. In Nerbonne et al. (eds.), 237–272.
- Kathol, Andreas. 1995. *Linearization-based German syntax*. Ph.D. dissertation, The Ohio State University.
- Kathol, Andreas. 1999. Agreement and the syntax–morphology interface in HPSG. In Robert D. Levine & Georgia M. Green (eds.), *Studies in contemporary phrase structure grammar*, 223–274. Cambridge: Cambridge University Press.
- Kathol, Andreas. 2000. *Linear Syntax*. New York & Oxford: Oxford University Press.
- Kay, Paul. 2005. Argument structure constructions and the argument–adjunct distinction. In Mirjam Fried & Hans C. Boas (eds.), *Grammatical constructions: Back to the roots* (Constructional Approaches to Language 4), 71–98. Amsterdam & Philadelphia: John Benjamins. <http://www.icsi.berkeley.edu/~kay/ASCs.pdf> (7 November 2006).
- Kay, Paul & Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The *What's X doing Y?* construction. *Language* 75.1, 1–33.
- Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8, 63–99.
- Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Max Niemeyer.
- Koenig, Jean-Pierre. 1999. *Lexical relations* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.
- Krenn, Brigitte & Gregor Erbach. 1994. Idioms and support verb constructions. In Nerbonne et al. (1994), 365–396.
- Krieger, Hans-Ulrich & John Nerbonne. 1993. Feature-based inheritance networks for computational lexicons. In Ted Briscoe, Ann Copestake & Valeria de Paiva (eds.), *Inheritance, defaults, and the lexicon*, 90–136. Cambridge: Cambridge University Press.
- Lüdeling, Anke. 1997. Strange resultatives in German: New evidence for a semantic treatment. In Ralph C. Blight & Michelle J. Moosally (eds.), *The 1997 Texas Linguistics Society Conference: The syntax and semantics of predication* (Texas Linguistic Forum 38), 223–233. Austin, TX: University of Texas, Department of Linguistics.
- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4.2, 201–225. <http://www.ling.upenn.edu/papers/v4.2-contents.html> (12 November 2004).
- Megerdooian, Karine. 2002. *Beyond words and phrases: A unified theory of predicate composition*. Ph.D. dissertation, University of Southern California. <http://www.zoorna.org/publications.html> (29 March 2007).
- Meurers, Walt Detmar. 2000. Lexical generalizations in the syntax of German non-finite constructions (Arbeitspapiere des SFB 340 No. 145). Tübingen: Eberhard-Karls-Universität. <http://www.ling.ohio-state.edu/~dm/papers/diss.html> (19 August 2002).
- Meurers, Walt Detmar, Gerald Penn & Fraqnk Richter. 2002. A web-based instructional platform for constraint-based grammar formalisms and parsing. In Dragomir Radev & Chris Brew (eds.), *Effective tools and methodologies for teaching NLP and CL* (Workshop held at 40th Annual Meeting of the Association for Computational Linguistics), 18–25. Philadelphia, PA. <http://www.ling.ohio-state.edu/~dm/papers/aclo2.html> (8 January 2004).
- Michaelis, Laura A. 2006. Construction Grammar. In Keith Brown (ed.), *The encyclopedia of language and linguistics*, 2nd edn., vol. 3, 73–84. Oxford: Elsevier. <http://spot.colorado.edu/~michaeli/publications.html> (17 January 2007).
- Michaelis, Laura A. & Josef Ruppenhofer. 2001. *Beyond alternations: A constructional model of the German applicative pattern* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.
- Miller, Philip H. & Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language & Linguistic Theory* 15.3, 573–639.

- Mohammad, Jan & Simin Karimi. 1992. Light verbs are taking over: Complex verbs in Persian. *The Western Conference on Linguistics (WECOL)*, 195–212.
- 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, CA: 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. <http://hpsg.fu-berlin.de/~stefan/Pub/complex.html> (29 October 2009).
- Müller, Stefan. 2003. The morphology of German particle verbs: Solving the bracketing paradox. *Journal of Linguistics* 39.2, 275–325. <http://hpsg.fu-berlin.de/~stefan/Pub/paradox.html> (29 October 2009).
- Müller, Stefan. 2005. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201, 3–39. <http://hpsg.fu-berlin.de/~stefan/Pub/satz-lb.html> (29 October 2009).
- Müller, Stefan. 2006. Phrasal or lexical constructions? *Language* 82(4), 850–883. <http://hpsg.fu-berlin.de/~stefan/Pub/phrasal.html> (29 October 2009).
- Müller, Stefan. 2007a. The Grammix CD Rom: A Software Collection for Developing Typed Feature Structure Grammars. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar engineering across frameworks 2007* (Studies in Computational Linguistics ONLINE). Stanford, CA: CSLI Publications.
- Müller, Stefan. 2007b. *Head-driven Phrase Structure Grammar: Eine Einführung* (Stauffenburg Einführungen 17), 1st edn. Tübingen: Stauffenburg. <http://hpsg.fu-berlin.de/~stefan/Pub/hpsg-lehrbuch.html> (29 October 2009).
- Müller, Stefan (ed.). 2009a. *16th International Conference on Head-driven Phrase Structure Grammar, University of Göttingen, Germany*. Stanford, CA, CSLI Publications.
- Müller, Stefan. 2009b. Towards an HPSG analysis of Maltese. In Bernard Comrie, Ray Fabri, Beth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese linguistics: 1st International Conference on Maltese Linguistics* (Bremen/Germany, 18–20 October 2007; Studies in Language Companion Series 113), 83–112. Amsterdam & Philadelphia: John Benjamins. <http://hpsg.fu-berlin.de/~stefan/Pub/maltese-sketch.html> (18 October 2009).
- Müller, Stefan. In preparation. *Grammatiktheorie: Von der Transformationsgrammatik zur beschränkungs-basierten Grammatik* (Stauffenburg Einführungen). Tübingen: Stauffenburg. <http://hpsg.fu-berlin.de/~stefan/Pub/grammatiktheorie.html> (14 November 2009).
- Müller, Stefan & Janna Lipenkova. 2009. Serial verb constructions in Mandarin Chinese. In Müller (2009a), 213–233. <http://hpsg.fu-berlin.de/~stefan/Pub/chinese-svc.html> (30 October 2009).
- Müller, Stefan & Pollet Samvelian. In preparation. *Persian in Head-driven Phrase Structure Grammar*.
- Nerbonne, John, Klaus Netter & Carl J. Pollard (eds.). 1994. *German in Head-driven Phrase Structure Grammar* (CSLI Lecture Notes 46). Stanford, CA: CSLI Publications.
- Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70.3, 491–538.
- Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology*. Ph.D. dissertation, University of California, Berkeley.
- Penn, Gerald. 2004. Balancing clarity and efficiency in Typed Feature Logic through delaying. 42nd Meeting of the Association for Computational Linguistics (ACL '04), main volume, 239–246. Barcelona.
- Pollard, Carl J. 1996. On head non-movement. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 279–305. Berlin & New York: Mouton de Gruyter. [Revised version of 1990 ms.]
- Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.
- Pollard, Carl J. & Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: University of Chicago Press.
- Pullum, Geoffrey K. 1977. Word order universals and grammatical relations. In Peter Cole & Jerrold M. Sadock (eds.), *Grammatical relations* (Syntax and Semantics 8), 249–277. New York: Academic Press.
- Richter, Frank & Manfred Sailer. 2009. Phraseological clauses as constructions in HPSG. In Müller (2009a), 297–317.



- Riehemann, Susanne Z. 1998. Type-based derivational morphology. *Journal of Comparative Germanic Linguistics* 2, 49–77. <http://doors.stanford.edu/~sr/morphology.ps> (5 April 2009).
- Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33.2, 431–484. <http://lingo.stanford.edu/sag/papers/rel-pap.pdf> (30 May 2004).
- Sag, Ivan A. 2007a. Remarks on locality. In Stefan Müller (ed.), *14th International Conference on Head-driven Phrase Structure Grammar*, 394–414. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/HPSG/8/> (29 July 2007).
- Sag, Ivan A. 2007b. Sign-based Construction Grammar: An informal synopsis. Ms., Stanford University. <http://lingo.stanford.edu/sag/papers/theo-syno.pdf> (23 October 2007).
- Samvelian, Pollet. 2007. A (phrasal) affix analysis of the Persian Ezafe. *Journal of Linguistics* 43, 605–645.
- Saussure, Ferdinand de. 1916. *Grundfragen der allgemeinen Sprachwissenschaft*. Berlin: Walter de Gruyter. [2nd edn. 1967]
- Stiebels, Barbara. 1996. *Lexikalische Argumente und Adjunkte: Zum semantischen Beitrag verbaler Präfixe und Partikeln* (Studia Grammatica 39). Berlin: Akademie.
- Stiebels, Barbara & Dieter Wunderlich. 1994. Morphology feeds syntax: The case of particle verbs. *Linguistics* 32.6, 913–968.
- Stump, Gregory T. 1991. A paradigm-based theory of morphosemantic mismatches. *Language* 67.4, 675–725.
- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, Michael. 2005. Beyond formalities: The case of language acquisition. *The Linguistic Review* 22, 183–197. [http://www.eva.mpg.de/psycho/pdf/Publications\\_2005\\_PDF/Beyond\\_formalities\\_05.pdf](http://www.eva.mpg.de/psycho/pdf/Publications_2005_PDF/Beyond_formalities_05.pdf) (24 September 2007).
- Tomasello, Michael. 2006a. Construction Grammar for kids. *Constructions*, special volume 1. <http://www.constructions-online.de/articles/specvol1/689> (11 January 2009).
- Tomasello, Michael. 2006b. Konstruktionsgrammatik und früher Erstspracherwerb. In Kerstin Fischer & Anatol Stefanowitsch (eds.), *Konstruktionsgrammatik: Von der Anwendung zur Theorie* (Stauffenburg Linguistik 40), 19–37. Tübingen: Stauffenburg.
- Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.
- Vahedi-Langrudi, Mohammad Mahdi. 1996. *The syntax, semantics and argument structure of complex predicates in Modern Farsi*. Ph.D. dissertation, University of Ottawa.
- Wiese, Richard. 1992. Prosodic phonology and its role in the processing of written language. In Günther Görz (ed.), *Konvens 92.1: Konferenz 'Verarbeitung natürlicher Sprache'*. Nürnberg 7.–9. Oktober 1992 (Informatik aktuell), 139–148. Berlin: Springer.
- Wiese, Richard. 1996. Phrasal compounds and the theory of word syntax. *Linguistic Inquiry* 27.1, 183–193.
- Zeller, Jochen. 1999. *Particle verbs, local domains, and a theory of lexical licensing*. Ph.D. dissertation, University of Frankfurt.
- Author's address* : Freie Universität Berlin, Institut für Deutsche und Niederländische Philologie, Deutsche Grammatik, Habelschwerdter Allee 45, D-14195 Berlin, Germany  
Stefan.Mueller@fu-berlin.de