Persian
in Head-Driven Phrase Structure Grammar

Stefan Müller
Freie Universität Berlin

Pollet Samvelian
Université Paris-Sorbonne & CNRS

Olivier Bonami
Université Paris-Sorbonne & CNRS

Friday 2nd November, 2012
Draft
comments welcome!
The Project

This book is part of a larger project, called CoreGram, with the goal to develop large scale computer processable grammar fragments of several languages that share a common core. Currently we work on the following languages:

- German (Müller, 2008, 2009b; Müller and Ørsnes, 2011)
- Danish (Ørsnes, 2009; Müller, 2009b; Müller and Ørsnes, 2011, To appear)
- Persian (Müller, 2010b; Müller and Ghayoomi, 2010; Müller, Samvelian and Bonami, In Preparation)
- Maltese (Müller, 2009a)
- Mandarin Chinese (Lipenkova, 2009; Müller and Lipenkova, 2009)
- Yiddish (Müller and Ørsnes, 2011)
- English
- Spanish
- French

For the implementation we use the TRALE system (Meurers, Penn and Richter, 2002; Penn, 2004), which allows for a rather direct encoding of HPSG analyses (Melnik, 2007). The grammars of German, Danish, Persian, Maltese, and Mandarin Chinese are of non-trivial size and can be downloaded at http://hpsg.fu-berlin.de/Projects/core.html. They are also part of the version of the Grammix CD-rom (Müller, 2007a) that is distributed with this book. The grammars of Yiddish and English are toy grammars that are used to verify cross-linguistic analyses of special phenomena and the work on Spanish and French is part of work in the Sonderforschungsbereich 632 which just started. See Bildhauer, 2008 for an implemented grammar of Spanish that will be converted into the format of the grammars mentioned above.

We believe that books are the best way to document such fragments since it is often not possible to construct a coherent view of one language from journal articles. The reason is that journal articles tend to need a long time from first submission to final publication and sometimes basic assumptions may have changed during the development of the linguistic theory in the meantime. The first book in this series was Müller, 2008, which describes a fragment of German that is implemented in the grammar BerliGram. Another book on the Danish Grammar developed in the DanGram project is in preparation (Müller and Ørsnes, To appear).

The situation in mainstream formal linguistics has often been criticized: basic assumptions are changed in high frequency, sometimes without sufficient motivation. Some concepts are not worked out in detail and formal underpinnings are unclear (see for instance Gazdar, Klein, Pullum and Sag, 1985, p.6; Pullum, 1985, 1989, 1991, p.48; Kornai and Pullum, 1990; Kuhns, 1986, p.550; Crocker and Lewin, 1992, p.508; Kolb

For a more detailed discussion of this point see Müller, 2010a, Chapter 3.7. As already mentioned, we work in the framework of HPSG, which is well-formalized (King, 1999; Pollard, 1999; Richter, 2004) and stable enough to develop larger fragments over a longer period of time. HPSG is a constraint-based theory which does not make any claims on the order of application of combinatorial processes. Theories in this framework are just statements about relations between linguistic objects or between properties of linguistic objects and hence compatible with psycholinguistic findings and processing models (Sag and Wasow, 2011).

As is argued in Müller, 2010a, Chapter 11.4, HPSG is compatible with UG-based models of language acquisition as for instance the one by Fodor (1998). See Fodor, 2001, p.385 for an explicit remark to that end. However, in recent years evidence has accumulated that arguments for innate language specific knowledge are very weak. For instance, Johnson (2004) showed that Gold’s proof that natural languages are not identifiable in the limit by positive data alone (Gold, 1967) is irrelevant for discussions of human language acquisition. Furthermore, there is evidence that the input that humans have is sufficiently rich to acquire structures which were thought by Chomsky (1971, p.29-33) and others to be inacquirable: Bod (2009) showed how syntactic structures could be derived from an unannotated corpus by Unsupervised Data-Oriented Parsing. He explained how Chomsky’s auxiliary inversion data can be captured even if the input does not contain the data that Chomsky claims to be necessary (see also Eisenberg, 1992 and Pullum and Scholz, 2002; Scholz and Pullum, 2002 for other Poverty of the Stimulus arguments). Input-based models of language acquisition in the spirit of Tomasello (2003) seem highly promising and in fact can explain language acquisition data better than previous UG-based models (Freudenthal et al., 2006, 2009). We argued in Müller, 2010a that the results from language acquisition research in the Construction Grammar framework can be carried over to HPSG, even in its lexical variants.1 If language acquisition is input-based and language-specific innate knowledge is minimal as assumed by Chomsky (1995); Hauser, Chomsky and Fitch (2002) or non-existing, this has important consequences for the construction of linguistic theories: Proposals that assume more than 400 morpho-syntactic categories that are all innate and that play a role in all languages of the world even though they are not directly observable in many languages (Cinque and Rizzi, 2010) have to be rejected right away. Furthermore, it cannot be argued for empty functional projections in language X on the basis of overt morphemes in language Y. This has been done for Topic Projections that are assumed for languages without topical morphemes on the basis of the existence of a topic morpheme in Japanese. Similarly, functional projections for object agreement have been proposed for languages like English and German on the basis of Basque data even though neither English nor German has object agreement. Since German children do not have any evidence from

1In fact we believe that a lexical treatment of argument structure is the only one that is compatible with the basic tenets of theories like Categorial Grammar (CG), Lexical Functional Grammar (LFG), CxG, and HPSG that adhere to lexical integrity (Bromen and Mchombo, 1995). For discussion see Müller, 2006, Müller, 2010a, Chapter 11.11, Müller, 2010b, and Müller, Submitted.

 language acquisition/ Universal Grammar (UG) / Poverty of the Stimulus Construction Grammar (CxG) Japanese English German Basque

language acquisition/ Universal Grammar (UG) / Government and Binding (GB) / Construction Grammar (CxG) Sign-Based Construction Grammar (CbG) Embodied Construction Grammar (CbG)
The Data

Acknowledgments

We thank Adele Goldberg for discussion.

We thank Mina Esmaili, Yasser Shakeri, and Mehran A. Taghvaipour for comments on the journal article that was the basis for the Chapter on Complex Predicates. In addition we want to thank Elham Alaei for discussion of data and Daniel Hole and Jacob Maché for comments on that paper. We thank Emily M. Bender and Felix Bildhauer for discussion. Special thanks go to Bob Borsley, Karine Megerdoomian, Ivan Sag, and several anonymous reviewers, whose comments on earlier versions of the paper improved it considerably.

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. We thank the respective institutions and the organizers of the HPSG conference for the invitation and all audiences for discussion.

We wish to thank the participants of the HPSG seminar at Paris Diderot University, as well as the anonymous reviewers and participants of the 2010 HPSG conference. Special thanks go to Olivier Bonami, Philip Miller, François Mouret, and Gert Webellith. This work is supported by the bilateral project “PerGram”, with funding from the ANR (France) and the DFGS (Germany) [grant no. MU 2822/3-1].

This book is typeset with XeLaTeX with the bidi package by Vafa Khalighi. We thank the developers of both BeLaTeX and XeLaTeX for their work and the members of the German German Language LaTEX Users Group Communication List and those replying at http://tex.stackexchange.com for many useful hints and suggestions. Special thanks go to Mostafa Nouri for his help with the Persian script. He even was so kind to offer setting up a virtual machine with all the components for typesetting Persian installed!

Berlin and Paris, Friday 2nd November, 2012

Stefan Müller, Pollet Samvelian and Olivier Bonami
2. A Brief Introduction to Head-Driven Phrase Structure Grammar

Head-Driven Phrase Structure Grammar (HPSG) was developed by Ivan Sag and Carl Pollard in the mid 80s. The main publications are Pollard and Sag, 1987, 1994. International conferences have been held since 1994 and there is a rich collection of publications regarding analyses of linguistic phenomena (in the area of phonology, morphology, syntax, semantics, and information structure), formal foundations of the framework, and computational issues like efficient parsing and generation. See http://hpsg.fu-berlin.de/HPSG-Bib/ for bibliographic data.

Since HPSG analyses are usually sufficiently formalized they can and have been implemented as computer processable grammars. This makes it possible to check the interactions of analyses with other phenomena and to use the linguistic knowledge in practical applications. See Bender et al., In Preparation for further details.

2.1. Formal Foundations

HPSG assumes feature structures as models of linguistic objects. Feature structures consist of feature value pairs. The values can be atomic or feature structures. Every feature structure is of a certain type. Types are ordered in hierarchies with the most general type at the top of the hierarchy and the most specific types at the bottom. Figure 2.1 shows an example hierarchy for the type case and its subtypes. Types in

```
  case
/\   /
nom gen dat acc  
```

Figure 2.1.: Subtypes of case in a grammar of German

A model of a linguistic object is maximally specific, that is, a noun or an attributive adjective in a model of an actual utterance has a case value that is nom, gen, dat, or acc. The linguist develops theories that describe possible feature structures. In contrast to feature structures, feature descriptions can be partial. For instance it is not necessary to specify a case value for the German word Frau (‘woman’) since Frau can be used in NPs of all four cases. (1) shows a simplified description of the nominal agreement information for the German noun Frau (‘woman’) (see Kathol, 1999 for details and Wechsler and Zlatić, 2003 for a comprehensive overview of agreement in HPSG). Frau has feminine gender, is compatible with all four cases, and is singular. The AVM has the type nom-agr. Types are written in italics. nom-agr is a complex type which introduces the features gen, case, and num. fem, case, sg are also types, but they are atomic.

\[
\begin{array}{ll}
\text{nom-agr} & \text{case nom} \\
\text{num} & \text{sg}
\end{array}
\]

\[
\begin{array}{ll}
\text{nom-agr} & \text{case nom} \lor \text{acc num} \\
\text{sg} & \text{pl}
\end{array}
\]

While structure sharing is the most important expressive means in HPSG there is one extension of the basic formalism that plays a crucial role in most HPSG analyses: relational constraints. Relational constraints are used to relate several values in a feature structure to each other. The relational constraint that is used most often in HPSG is append (‘⊕’). append is used to concatenate two lists. Schema 1, which will be discussed in Section 2.2.2, is an example for an application of such a constraint.

This brief sketch basically described all the formal tools that are used in HPSG. Of course a lot more could be and has been said about the properties of the formalisms, but
this introductionary section is not the place to discuss these issues in detail. However, it cannot be emphasized enough that it is important that the formal details are worked out and the interested reader is referred to the work of Shieber (1986), Pollard and Sag (1987, Chapter 2), Johnson (1988), Carpenter (1992), King (1994, 1999), Pollard (1999) and Richter (2004). The work of King, Pollard, and Richter reflects current assumptions, that is, the model theoretic view on grammar that is assumed nowadays.

Before I start to discuss several phenomena and their analyses in HPSG in the following sections I want to give an overview of the general feature geometry as it was developed in Pollard and Sag, 1994. (5) shows parts of the lexical item for Frau (‘woman’).

(5) SYNTH-SYNTHS

<table>
<thead>
<tr>
<th>PHONOLGY</th>
<th>(frau)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CATEGORY</td>
<td>HEAD</td>
</tr>
<tr>
<td></td>
<td>AGR</td>
</tr>
<tr>
<td></td>
<td>GEN: fem</td>
</tr>
<tr>
<td></td>
<td>CASE</td>
</tr>
<tr>
<td></td>
<td>NUM: sg</td>
</tr>
<tr>
<td></td>
<td>nom-agr</td>
</tr>
<tr>
<td>LOCAL</td>
<td>SPR</td>
</tr>
<tr>
<td></td>
<td>DET:AGR</td>
</tr>
<tr>
<td>CONTENT</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>INST X</td>
</tr>
<tr>
<td></td>
<td>frau</td>
</tr>
<tr>
<td>NONLOCAL</td>
<td>...</td>
</tr>
<tr>
<td>SYSTEM</td>
<td>word</td>
</tr>
</tbody>
</table>

The first feature value pair describes the phonological form of the word. The value of phon is a list of phonemes. For reasons of readability usually the orthographic form is given in HPSG papers and phonological structure is omitted, but see Bird and Klein, 1994 and Höhle, 1999 for analyses. The second feature is SYNTAX-SYNTHESES (SYNSEM) and its value is a description of all properties of a linguistic object that are syntactically and semantically relevant and can be selected by other heads. Information that is locally relevant (LOCAL) is distinguished from information that plays a role in non-local dependencies (NONLOCAL, see Section ?). Syntactic information is represented under CATEGORY (CAT) and semantic information under CONTENT (CONT). The example shows the HEAD value, which provides information about all aspects that are relevant for the external distribution of a maximal projection of a lexical head. In particular the part of speech information (noun) is represented under HEAD. The value of AGREEMENT (AGR) is the one given in (1). As well as information regarding the head features, valence information also belongs under CAT. The example shows the SPR feature, which is used for the selection of a specifier (see the next section for details on valence). The F is an example of structure sharing. It ensures that the specifier that is realized together with the noun has compatible agreement features.

This description shows the syntax-semantics (SYNSEM) associated with the HEAD of Peter sleeps (Figure 2.2). The feature value pair (SPR, COMPS) is empty since there is no other argument. The HEAD has no oblique arguments (GEN, OBJECT).
sleeps is realized as a sister of sleeps. Figure 2.3 shows a more complex example with a transitive verb. likes and Sandy form a VP (a verbal projection with an empty COMPS list) and this VP is combined with its subject to form a fully saturated verbal projection, that is, a clause.

2.2.2. Constituent Structure

As was explained in Section 2.1, HPSG exclusively uses feature structures with structure sharing and relational constraints for modeling linguistic objects. As a consequence of this the theory does not use phrase structure rules. Instead the dominance relation between linguistic objects is modeled with feature structures. Trees are used for visualization purposes only. The attribute value matrice that represents the dominance between linguistic objects is modeled with feature structures. Trees are used for visualization purposes only. The attribute value matrice that represents the dominance between linguistic objects is modeled with feature structures. Trees are used for visualization purposes only. The attribute value matrice that represents the dominance between linguistic objects is modeled with feature structures. Trees are used for visualization purposes only.

The value of PHON gives a list of phonological contributions of the daughter signs. The feature HEAD-DTR is appropriate for headed structures. Its value is the sign that contains the head of a complex expression (the verb in a VP, the VP in a clause). The value of NON-HEAD-DTRs is a list of all other daughters of a sign.

The following implication shows the constraints that hold for structures of type head-complement-phrase:

Schema 1 (Head-Complement-Schema (fixed order))

\[
\text{head-complement-phrase} \rightarrow \\
\text{SYNSEM}[\text{LOC}][\text{CAT}]\text{COMPS} \cdot \text{HEAD-DTR}[\text{SYNSEM}][\text{LOC}][\text{CAT}]\text{COMPS} \cdot \text{NON-HEAD-DTRs} \\
\text{SYNSEM} \cdot \text{HEAD-DTR} \cdot \text{SYNSEM} \\
\text{NON-HEAD-DTRs} \\
\text{SYNSEM} \\
\text{HEAD-DTR} \\
\text{SYNSEM} \\
\text{NON-HEAD-DTRs}
\]

This constraint splits the COMPS list of the head daughter into two parts: a list that contains exactly one element (\([\text{synsem}] \cdot \text{HEAD-DTR} \cdot \text{SYNSEM} \cdot \text{HEAD-DTR} \cdot \text{SYNSEM} \cdot \text{NON-HEAD-DTRs} \) and a remaining list (\([\text{synsem}] \cdot \text{HEAD-DTR} \cdot \text{SYNSEM} \cdot \text{NON-HEAD-DTRs} \) ). The first element of the COMPS list is identified with the SYNSEM value of the non-head daughter. It is therefore ensured that the description of the properties of the complement of a transitive verb like likes in Figure 2.3 is identified with the feature value bundle that corresponds to the properties of the object that is combined with the head (Sandy in the case of Figure 2.3). Since Schema 1 licenses structures with exactly one head daughter and exactly one non-head daughter, structures will be binary. This is not the only option for defining head complement structures. The constraints can be specified in a way that allows for the realization of any number of complements in one go. See for instance Pollard and Sag, 1994 for an analysis of English with a flat VP and Bouma and van Noord (1998) for an absolutely flat analysis of Dutch, including a flat verbal complex.

The Schema 1 licences the VP in Figure 2.3. The combination of the VP and its specifier is licenced by the HeadSpecifier-Schema.\(^1\)

Schema 2 (Specifier-Head-Schema)

\[
\text{specifier-head-phrase} \rightarrow \\
\text{SYNSEM}[\text{LOC}][\text{CAT}][\text{SYNSEM}][\text{LOC}][\text{CAT}][\text{SPECIFIER}][\text{NON-HEAD-DTRs}] \\
\text{SYNSEM} \\
\text{SPECIFIER} \\
\text{NON-HEAD-DTRs} \\
\text{SYNSEM} \\
\text{SPECIFIER} \\
\text{NON-HEAD-DTRs}
\]

This schema also licences the combination of nominal projections with a determiner.

\(^1\)Note that the non-head daughter is taken from the end of the SPR list, while the non-head daughter in head-complement phrases is taken from the beginning. For heads that have exactly one specifier this difference is irrelevant, but in the analysis of object shift and negation shift that is suggested by Müller and Östnøs (To appear), the authors assume multiple specifiers and the difference in order of combination will be relevant.
2.2. Constituent Order

In the simple NP example above the order of the elements is fixed: the head follows the non-head. However this is not always the case. For instance there are mixed languages like Persian that allow some heads to the left of their arguments and some heads to the right (Prepositional phrases are head initial and verb phrases are head final in Persian). For such reasons HPSG assumes a separation between immediate dominance (ID) constraints and linear precedence (LP) constraints as was common in GPSG (Gazdar et al., 1985). For instance, Schema 1 does not impose any order on the head and the non-head. This is taken care of by a set of separate constraints.

Heads that precede their complements can be marked as initial+. The following LP constraints ensure the right ordering of heads with respect to their complements:

(8) a. HEAD [initial+] < COMPLEMENT
b. COMPLEMENT < HEAD [initial-]

2.2.3. Constituent Order

In the simple NP example above the order of the elements is fixed: the head follows the non-head. However this is not always the case. For instance there are mixed languages like Persian that allow some heads to the left of their arguments and some heads to the right (Prepositional phrases are head initial and verb phrases are head final in Persian). For such reasons HPSG assumes a separation between immediate dominance (ID) constraints and linear precedence (LP) constraints as was common in GPSG (Gazdar et al., 1985). For instance, Schema 1 does not impose any order on the head and the non-head. This is taken care of by a set of separate constraints.

Heads that precede their complements can be marked as initial+. The following LP constraints ensure the right ordering of heads with respect to their complements:

(8) a. HEAD [initial+] < COMPLEMENT
b. COMPLEMENT < HEAD [initial-]

2.2.4. Free Constituent Order Languages

Schema 1 allows for the combination of a head with its complements in a fixed order (similar to what is known from Categorial Grammar). Taken together with the linearization constraint in (8a), this results in a fixed constituent order in which the verb precedes its complements and the complements are serialized according to their obliqueness. However there are languages with much freer constituent order than English. If one does not want to assume a base order from which other orders are derived by movement or equivalents to movement one has to find ways to relax the constraint on head complement structures. One way of doing this is to allow the non-head daughter to be an arbitrary element from the COMPS list of the head daughter. The respective modification of the schema is given as Schema 3:

Schema 3 (Head-Complement-Schema (free constituent order))

\[
\text{head-complement-phrase} \Rightarrow \left\{ \begin{array}{l} \text{synsem} = \text{loc}\mid \text{cat}|\text{comps} \oplus \oplus \oplus \\
\text{head-dtr} = \text{synsem} = \text{loc}|\text{cat}|\text{comps} \oplus \oplus \oplus \oplus \\
\text{non-head-dtrs} = \left\{ \begin{array}{l} \text{synsem} \oplus \oplus \oplus \oplus \\
\end{array} \right. \end{array} \right. 
\]

The COMPS list of the head daughter is split into three parts: a list of arbitrary length ([]), a list containing one element ([[]]) and another list of arbitrary length ([[]]). [], [], [] can be the empty list or contain one or more arguments.

For non-configurational languages it is assumed that the subject of finite verbs is treated like the other arguments, that is, it is mapped to COMPS instead of being mapped to SPR as in English. Having explained the difference in the HPSG analysis of configurational and non-configurational languages we can now give an example of an analysis of a language with rather free constituent order: Figures 2.5 and 2.6 show the analysis of the German sentences in (9):

Draft of Friday 2nd November, 2012, 21:32
its head. Such analyses were suggested very early in the history of HPSG by Gunji (1986) for Japanese. See also Hinrichs and Nakazawa (1989), Pollard (1996), and Engelkamp, Erbach and Uszkoreit (1992) for set-based approaches to constituent order in German.

A crucial difference between a set-based analysis and the list-based analysis advocated here is that the elements of the lists are ordered in order of obliqueness. This order is used in various subparts of the theory for instance for assignment of structural case and for expressing constraints on pronoun binding. So the obliqueness ordering has to be represented elsewhere in set-based approaches.

For authors who assume binary branching structures the difference between languages with fixed constituent order and languages with free constituent order lies in the value of \( \mathbf{a} \) and \( \mathbf{b} \) in Schema 3. If either \( \mathbf{a} = \mathbf{b} \) is the empty list one gets a fixed constituent order, with head complement combination either in order of obliqueness or in the reverse order of obliqueness.

To sum up, there are three approaches to free constituent order: Flat structures, linearization domains with discontinuous constituents, and the non-cancellation of syntactic and semantic properties of arguments.

### 2.2.5. Heads and Projection of Head Features

Section 2.1 introduced head features and Figure 2.3 shows that the information about part of speech of the head is present at every projection, but until now nothing has been said about head feature propagation. The identity of the head features of a head and of a mother node is taken care of by the following principle:

**Principle 1 (Head Feature Principle)** In a headed phrase, the head value of the mother and the head value of the head daughter are identical.

This can be formalized by the following implicational constraint:

\[
\text{headed-phrase} \Rightarrow \begin{align*}
\text{SYNSEM} & \rightarrow \text{LOCAL} \rightarrow \text{CAT} \rightarrow \text{HEAD} \rightarrow \mathbf{a} \\
\text{HEAD-DTR} & \rightarrow \text{SYNSEM} \rightarrow \text{LOCAL} \rightarrow \text{CAT} \rightarrow \text{HEAD} \rightarrow \mathbf{b}
\end{align*}
\]

The head daughter is the daughter that contains the syntactic head, that is, in the phrase *likes Sandy* in Figure 2.3 it is the lexical item *likes* and in the phrase *Kim likes Sandy* it is the constituent *likes Sandy*. The constraint is a constraint on structures of type headed-phrase. Types like head-complement-phrase and head-specifier-phrase are subtypes of headed-phrase and hence the constraint in (10) applies to them too.

### 2.3. Non-Cancellation of Valence Requirements

### 2.4. Semantics

The first publications on HPSG assumed Situation Semantics (Barwise and Perry, 1983) as the underlying semantic framework (Pollard and Sag, 1987, 1994). While there are also more recent publications in this tradition (Ginzburg and Sag, 2000), many current analyses use semantic formalisms that allow for the underspecification of scope constraints such as for instance Minimal Recursion Semantics (MRS, Copestake, Flickinger, Pollard and Sag, 2005) and Lexical Resource Semantics (LRS, Richter and Sailer, 2004).

#### 2.4.1. Minimal Recursion Semantics

(11) shows the examples for the semantic contribution of a noun and a verb in Minimal Recursion Semantics (MRS):

\[
\begin{align*}
\text{(11) a.} & \quad \text{dog} \quad \begin{bmatrix}
\text{IND} & \begin{bmatrix}
\text{PER} & \text{3} \\
\text{NUM} & \text{sg}
\end{bmatrix} \\
\text{RELS} & \begin{bmatrix}
\text{INST} & \text{dog}
\end{bmatrix}
\end{bmatrix} \\
\text{(11) b.} & \quad \text{chases} \quad \begin{bmatrix}
\text{IND} & \begin{bmatrix}
\text{EVENT} & \text{index}
\end{bmatrix} \\
\text{RELS} & \begin{bmatrix}
\text{AGENT} & \text{index} \\
\text{PATIENT} & \text{chase}
\end{bmatrix}
\end{bmatrix}
\end{align*}
\]

An MRS consists of an index, a list of relations, and a set of handle constraints, which will be introduced below. The index can be a referential index of a noun (11a) or an event variable (11b). In the examples above the lexical items contribute the *dog* relation and the *chase* relation. The relations can be modeled with feature structures by turning the semantic roles into features. The semantic index of nouns is basically a variable, but it comes with an annotation of person, number, and gender since this information is important for establishing correct pronoun bindings.

The arguments of each semantic relation (e.g. agent, patient) are linked to their index. An MRS consists of an index, a list of relations, and a set of handle constraints, which will be introduced below. The index can be a referential index of a noun (11a) or an event variable (11b). In the examples above the lexical items contribute the *dog* relation and the *chase* relation. The relations can be modeled with feature structures by turning the semantic roles into features. The semantic index of nouns is basically a variable, but it comes with an annotation of person, number, and gender since this information is important for establishing correct pronoun bindings.

The arguments of each semantic relation (e.g. agent, patient) are linked to their index. An MRS consists of an index, a list of relations, and a set of handle constraints, which will be introduced below. The index can be a referential index of a noun (11a) or an event variable (11b). In the examples above the lexical items contribute the *dog* relation and the *chase* relation. The relations can be modeled with feature structures by turning the semantic roles into features. The semantic index of nouns is basically a variable, but it comes with an annotation of person, number, and gender since this information is important for establishing correct pronoun bindings.
Generalizations over linking patterns can be captured elegantly in inheritance hierarchies (see Section 2.6 on inheritance hierarchies and Davis, 1996; Wechsler, 1991; Davis and Koenig, 2000 for further details on linking in HPSG).

2.4. Semantics

V[ spr ⟨⟩, comp ⟨⟩ ]

NP[ nom ]

V[ spr ⟨⟩, comp ⟨⟩ ]

NP[ acc ]

every dog

likes

a cat

Figure 2.7.: Analysis for Every dog chases a cat.

Before turning to the compositional analysis of (13a), I want to introduce some additional machinery that is needed for the underspecified representation of the two readings in (13b,c).

(13)

a. Every dog chased some cat.

b. ∀x(dog(x) → ∃y(cat(y) ∧ chase(x, y)))

c. ∃y(cat(y) ∧ ∀x(dog(x) → chase(x, y)))

Minimal Recursion Semantics assumes that every elementary predication comes with a label. Quantifiers are represented as three place relations that relate a variable and two so-called handles. The handles point to the restriction and the body of the quantifier, that is, to two labels of other relations. (14) shows a (simplified) MRS representation for (13a).

(14) ⟨ h0, { h1: every(x, h2, h3), h5: some(y, h6, h7), h2: dog(x), h6: cat(y), h4: chase(e, x, y) } ⟩

The tree-place representation is a syntactic convention. Formulas like those in (13) are equivalent to the results of the scope resolution process that is described below.

The MRS in (14) can best be depicted as in Figure 2.8. h0 stands for the top element. This is a handle that dominates all other handles in a dominance graph. The restriction of every points to dog and the restriction of some points to cat. The interesting thing is that the body of every and some is not fixed in (14). This is indicated by the dashed lines in Figure 2.8 in contrast to the straight lines connecting the restrictions of the quantifiers with elementary predications for dog and cat, respectively. There are two ways to plug an elementary predication into the open slots of the quantifiers:

(15)

a. Solution one: h0 = h1 and h3 = h5 and h7 = h4. 
(every dog has wide scope)

b. Solution two: h0 = h5 and h7 = h1 and h3 = h4. 
(some cat has wide scope)

The solutions are depicted as Figure 2.9 and Figure 2.10.

There are scope interactions that are more complicated than those we have been looking at so far. In order to be able to underspecify the two readings of (16) both slots of a quantifier have to stay open.

(16)

a. Every nephew of some famous politician runs.

b. every(x, some(y, famous(y) ∧ politician(y)), run(x))

c. some(y, famous(y) ∧ politician(y), every(x, nephew(x, y), run(x)))

Figure 2.9.: every(x, dog(x), some(y, cat(y), chase(x, y)))

Figure 2.8.: Dominance graph for Every dog chases some cat.
Figure 2.10: \( \text{some(y, cat(y), every(x, dog(x), chase(x, y)))} \)

In the analysis of example (13a), the handle of \textit{dog} was identified with the restriction of the quantifier. This would not work for (16a) since either \textit{some} or \textit{neither} can be the restriction of \textit{every}. Instead of direct specification so-called handle constraints are used (\( qeq \) oder \( \_eq \)). A \( qeq \) constraint relates an argument handle and a label: \( h = g 1 \) means that the handle is filled by the label directly or one or more quantifiers are inserted between \( h \) and \( l \). Taking this into account, we can now return to our original example. The correct MRS representation of (13a) is given in (17).

\[
(17) \quad \{ h_0, \{ h_1: \text{every}(x, h_2, h_3), h_4: \text{dog}(x), h_5: \text{chase}(e, x, y), h_6: \text{some}(y, h_7, h_8), h_9: \text{cat}(y) \}, \{ h_2 = q h_4, h_7 = q h_9 \} \}
\]

The handle constraints are associated with the lexical entries for the respective quantifiers. Figure 2.11 shows the analysis. For compositional cases as in Figure 2.11, the \textit{RELS} value of a sign is simply the concatenation of the \textit{RELS} values of the daughters. Similarly the \textit{HCONS} value is a concatenation of the \textit{HCONS} values of the daughters.

2.4.2. The Analysis of Non-Compositional Constructions

Copestake, Flickinger, Pollard and Sag, 2005 extended the basic analysis that concatenates \textit{RELS} and \textit{HCONS} to cases in which the meaning of an expression is more than the meaning that is contributed by the daughters in a certain structure. They use the feature \textit{C-CONT} for the representation of constructional content. While usually the semantic functor (the head in head argument combinations and the adjunct in head adjunct structures) determines the main semantic contribution of a phrase, the \textit{C-CONT} feature can be used to specify a new main semantic contribution. In addition relations and scope constraints may be introduced via \textit{C-CONT}. The feature geometry for \textit{C-CONT} is given in (18):

\[
\begin{align*}
\text{feature!C-CONT} & \quad \text{feature!HOOK} \\
\text{INDEX event-or-index} & \quad \text{LTOP handle} \\
\text{RELS list of relations} & \quad \text{HCONS list of handle constraints} \\
\text{e-cont} & \quad \text{c-cont}
\end{align*}
\]

The \textit{HOOK} provides the local top for the complete structure and a semantic index, that is a nominal index or an event variable. In compositional structures the \textit{HOOK} value is structure shared with the semantic contribution of the semantic functor and the \textit{RELS} list and the \textit{HCONS} list is the empty list. As an example for a non-compositional combination Copestake et al., 2005 discuss determinerless plural NPs in English. For the analysis of \textit{tired squirrels} they assume an analysis using a unary branching schema. Their analysis corresponds to the one given in (19).\footnote{We do not assume a unary branching schema for bare plurals but an empty determiner, since using an empty determiner captures the generalizations more directly: while the empty determiner is fully parallel to the overt ones, the unary branching schema is not parallel to the binary branching structures containing an overt determiner. See also Alqurashi and Bosley, 2012 for a similar point regarding relative clauses in Modern Standard Arabic with and without a complementizer.}

Figure 2.11: Analysis for \textit{Every dog chases a cat.}
2.4. Semantics

(19) \[
\begin{align*}
\text{synsem} & | \text{loc} | \text{cont} \\
\text{rels} & \oplus 0 \\
\text{hcons} & \oplus 1
\end{align*}
\]

\text{rels} \oplus 1

\text{hcons} \oplus 0

\text{c-cont} \oplus 0

\text{hook} \oplus 0

\begin{align*}
\text{arg0} & \oplus 0 \\
\text{arg1} & \oplus 1 \\
\text{arg2} & \oplus 0
\end{align*}

\text{arg0} \oplus 0

\text{arg1} \oplus 0

\text{arg2} \oplus 0

\begin{align*}
\text{arg0} & \oplus 1 \\
\text{arg1} & \oplus 0 \\
\text{arg2} & \oplus 0
\end{align*}

\text{arg0} \oplus 0

\text{arg1} \oplus 1

\text{arg2} \oplus 0

\begin{align*}
\text{lil} & \oplus 0 \\
\text{arg0} & \oplus 0 \\
\text{arg1} & \oplus 0
\end{align*}

\text{lil} \oplus 0

\text{arg0} \oplus 0

\text{arg1} \oplus 0

\begin{align*}
\text{IND} & \oplus 0 \\
\text{ltop} & \oplus 0
\end{align*}

\text{IND} \oplus 0

\text{ltop} \oplus 0

\text{ind} \oplus 0

\text{ltop} \oplus 0

(19)

The semantic content of the determiner is introduced constructionally in \text{c-cont}. It consist of the relation \text{udef-rel}, which is a placeholder for the quantifier that corresponds to \text{some} or \text{every} in the case of overt determiners. The \text{rels} and \text{hcons} values that are introduced constructionally \( \text{c-cont} \oplus 0 \) and \( \text{hcons} \oplus 0 \) are concatenated with the \text{rels} and \text{hcons} values of the daughters \( \text{c-cont} \oplus 0 \) and \( \text{hcons} \oplus 0 \).

The Semantics Principle can now be specified as follows:

**Principle 2 (Semantics Principle)** The main semantic contribution of a phrase is identical to the value of \text{c-cont} \oplus \text{hook}. The \text{rels} value is the concatenation of the \text{rels} value in \text{c-cont} and the concatenation of the \text{rels} values of the daughters. The \text{hcons} value is the concatenation of the \text{hcons} value in \text{c-cont} and the concatenation of the \text{hcons} values of the daughters.

### 2.4.3. Decomposition in Syntax vs. Underspecification

An interesting application of the underspecification of scope constraints is the treatment of the ambiguity of (20a).

(20) a. dass Max alle Fenster aufmachte

that Max all windows opened

‘that Max opened all windows’

b. \( \forall x (\text{window}(x) \rightarrow \text{CAUSE}(\text{max}, \text{open}(x))) \)

c. \( \text{CAUSE}(\text{max}, \forall x (\text{window}(x) \rightarrow \text{open}(x))) \)

The first reading corresponds to a situation in which all windows were closed and Max opens each window and the second reading corresponds to a situation in which some windows were open already and Max opened the remaining windows which results in a situation in which all windows are open.

Egg (1999) suggests specifying the meaning of \text{öffnen} (‘to open’) in an underspecified way. (21) gives an MRS version of his analysis:

(21) \( \{ \text{h0}, \{ \text{h1:CAUSE}(x, \text{h2}), \text{h3:open}(y) \} \}, \{ \text{h2} \equiv \text{h3} \} \) \)

The CAUSE operator embeds the \text{open} predicate and the embedded predicate and therefore admits the readings in (20b,c). The analysis also extends to the readings that can be observed for sentences with adverbs like \text{wieder} (‘again’). The sentence in (22) has three readings that originate from different scopings of \text{CAUSE}, \forall, and \text{wieder} (‘again’):

(22) a. dass Max alle Fenster \text{wieder} aufmachte

that Max all windows \text{again} opened

b. \( \text{CAUSE} > \forall > \text{again'} > \text{open'} \)

c. \( \forall > \text{CAUSE} > \text{again'} > \text{open'} \)

d. \( \forall > \text{again'} > \text{CAUSE} > \text{open'} \)

The first two readings are so-called repetitive readings and the third one is a restitutive reading. See Dowty, 1978, Section 5.6 on this phenomenon. Since only the relative scope of \text{CAUSE} and \text{open'} is fixed, other scope-taking elements can intervene.

With such a semantic representation the syntax-semantics interface can be set up as follows: the adverbial combines with \text{aufmachen} and the resulting phrase is combined with the object \text{alle Fenster} and the subject \text{Max}. The scoping of the universal quantifier and the adverbial \text{wieder} depends on the ordering of the elements, that is in (22a) only readings in which \( \forall \) outscopes \text{again'} are available. See Kiss, 2001 for more information of the treatment of quantifier scope in German in the framework of HPSG.

Egg (1999) suggests the underspecification analysis as an alternative to von Stechow’s analysis in the Minimalist Program (1996). Von Stechow assumes a decomposition in syntax in the style of Generative Semantics and relies on several empty heads and movement operations that are necessary to derive readings. As was pointed out by Jäger and Blümel (2003) the analysis does not get all attested readings. Apart from such empirical problems, the underspecification analysis has to be preferred for reasons of simplicity: the syntactic structures directly correspond to observable facts.

### 2.5. Lexical Rules

Since HPSG is a lexicalist theory, the lexicon plays an important role. The lexicon is not just a prison for the lawless as suggested by Di Sciullo and Williams (1987, p. 3), but is structured and lexical items are related to each other. One means of capturing generalizations is lexical rules. A lexical rule says if there is a lexical item with certain properties
then there is also another lexical item with certain other properties. An example for the application of lexical rules is morphology (Pollard and Sag, 1987, Chapter 8.2. Orgun, 1996, Riehmann, 1998, Ackerman and Webelluth, 1998, Kathol, 1999, Koenig, 1999). The HPSG lexicon (of inflecting languages) consists of roots that are related to stems or fully inflected words. The derivational or inflectional rules may influence part of speech (adjectival derivation) and/or valence (-able adjectives and passive). (23) is an example for a lexical rule. It was suggested by Kiss (1992) to account for the personal passive in German. The rule takes as input a verbal stem that governs both a nominative and an accusative. The nominative argument is not represented in the COMPS list of the output. The case of the object is changed from acc to nom. The remaining arguments (if there are any) are taken over from the input [D].

(23) Lexical rule for the personal passive following Kiss (1992):

\[
\begin{align*}
\text{PHON} & : \text{[D]} \\
\text{SYNSEM} & : \text{LOC|CAT} \\
\qquad \quad \| \text{CAT} \\
\qquad \quad \| \text{STEM} \\
\text{HEAD} & : \text{verb} \\
\text{SUBCAT} & : \langle NP[\text{nom}], \text{NP[acc]} \rangle \oplus \text{[D]} \\
\text{word} & : \text{[D]}
\end{align*}
\]

The stem is mapped to a word and the phonology of the input [D] is mapped to the passive form by a function f.

During the past decades there has been some discussion concerning the status of lexical rules. One way to formalize them is to fully integrate them into the formalism of typed feature structures. According to this view the input of the lexical rule is a daughter of the output (Krieger and Nerbonne, 1993a, Chapter 7.4.1; Copestake and Briscoe, 1992; Meurers, 1995, 2001; Riehmann, 1998). This is basically equivalent to a unary branching immediate dominance rule. (24) shows the lexical rule in (23) in a format that directly reflects this approach.

(24) Lexical rule for the personal passive (fully integrated into the formalism):

\[
\begin{align*}
\text{PHON} & : \text{f(D)} \\
\text{SYNSEM} & : \text{LOC|CAT} \\
\qquad \quad \| \text{CAT} \\
\qquad \quad \| \text{STEM} \\
\text{HEAD} & : \text{vform passive-part} \\
\text{SUBCAT} & : \langle NP[\text{nom}], \text{NP[acc]} \rangle \oplus \text{[D]} \\
\text{LEX-DTR} & : \text{SYNSEM} | \text{LOC|CAT} \\
\qquad \quad \| \text{STEM} \\
\quad \text{acc-passive-lexical-rule}
\end{align*}
\]

A further advantage of this notation is that lexical rules are constraints on typed feature structures and as such it is possible to integrate them into an inheritance hierarchy and to capture generalizations over various linguistic objects.

For instance it was argued by Höhle (1997) that complementizers and finite verbs form a natural class in German.

(25) a. dass Karl das Buch liest
    ‘that Karl reads the book’

b. Liest Karl das Buch?
    ‘Does Karl read the book?’

In head-movement-inspired approaches (see Borsley (1989) for a head-movement approach for English, Müller and Orsnes, To appear for a head-movement approach for Danish, and Kiss and Wesche, 1991; Kiss, 1995; Meurers, 2000; Müller, 2008 for head-movement approaches for German) the verb in (25b) is related to a lexical item for the verb as it occurs in (25a) by a lexical rule. The complementizer and the lexical rule are subtypes of a more general type capturing the commonalities of dass in (25a) and liest in (25b).

2.6. Generalizations

HPSG is a theory that places a lot of information in the lexicon. For instance lexical entries of verbs contain detailed descriptions of their arguments, they contain information on argument structure, the semantic contribution of the verb, information about semantic roles and so on. A good way to capture generalizations with respect to this lexical knowledge is to use type hierarchies with multiple inheritance (Pollard and Sag, 1987, Chapter 8.1). Sag (1997) argued for several different immediate-dominance schemata for variants of English relative clauses and modified the feature geometry of HPSG in a way that made it possible to capture the generalizations over the various schemata in an inheritance hierarchy. Figure 2.12 on the facing page gives an example of how (parts

---

For a more general passive rule that unifies the analyses of personal and impersonal passives see Müller, 2002, Chapter 3. This more general rule for the passive uses the distinction between structural and lexical case.
Figure 2.12: Part of an inheritance hierarchy that contains lexical entries and immediate dominance schemata.

of an inheritance hierarchy that includes both lexical and phrasal types may look. In Section 2.2.5 we discussed constraints on phrases of type headed-phrase. Since structures of the type head-complement-phrase are a subtype of headed-phrase, they inherit all the constraints from their supertype. Hence, head features at the mother node of a head complement phrase are identified with the head features of the head daughter. Similarly the constraint that there is an nominative and an accusative object is represented at the type transitive-verb. The type strict-transitive-verb adds the information that there is no further argument and the type ditransitive-verb adds the information about an additional dative argument.
A. List of Phrases Covered/Rejected by the Grammar

**idiom + light verb construction**

1. او مرد را دوست داشت.
he/she man DOM friend had
\'(He/she) loved the man.'

2. او علی را مرد داشت.*
he/she Ali DOM man had
\'(He/she) loved the man.'

**incorporation + light verb construction**

3. علی تلفن کرد.
Ali telephone did.
\'Ali called.'

**causative + light verb construction**

4. من رادیو بز کرد.
I radio open do.
\'I opened the radio.'

5. من رادیو بز xāham kard.
I radio open will do.
\'I will open a radio.'

6. من رادیو بز xāham kard.
I radio open will do.
\'I will open a radio.'

**light verb construction + coordination**

9. من رادیو بز xāham kardam.
I radio open and clean did
\'I opened and cleaned a radio.'

**idiom + light verb construction + negation**

10. او مرد را دوست نداشت.
he/she man DOM friend NEG-have.
\'(He/she) does not love the man.'

11. او مرد را دوست خواهد داشت.
he/she man DOM friend want have
\'(He/she) will love the man.'

**idiom + light verb construction + future**

12. او مرد را دوست نخواهد داشت.
he/she man DOM friend NEG-want have
\'(He/she) will not love the man.'

he/she man DOM friend want NEG-have
\'(He/she) will not love the man.'

* Draft of Friday 2nd November, 2012, 21:32
negation + passive

(14) مرمی در خیابان دیده شد.
Maryam in street seen become.
'Maryam was seen in the street.'

(15) مرمی در خیابان ندیده شد.
Maryam in street seen NEG-become.
'Maryam was not seen in the street.'

negation + copula

(16) مرمی غم گرفت.
Maryam sad become.
'Maryam became sad.'

cliticization + possessives

(19) دیدم ت.
Didam at.
saw-1-sg-2sg.
'I saw you'

(20) دیدم ش.
Didam aš.
saw-1-sg-3sg.
'I saw him'

(21) او روشن کرد.
U rošan aš.
He/she light DO.CL.3sg do
'He/she turned it on.'

(22) اوروشان کرد.
U rošan aš kard.
He/she light DO.CL.3sg do
'He/she turned it on.'

(23) من باید خواهش کرد.
Man būz xāham aš kard.
I open want DO.CL.3sg do
'I will open it.'

(24) من باید خواهش کرد.
Man būz aš xāham kard.
I open DO.CL.3sg want do
'I will open it.'

(25) من خواهش نوشته.
Man xāham aš nevešt.
I will DO.CL.3sg write
'I will write it.'

cliticization + Ezafe + possessives

(26) من مادر ش را دیدم.
Man mādar aš rā didam.
I mother-Poss-3.SG RA saw
'I saw his/her mother.'

(27) کتاب بزرگ تو.
Ketābe bozorg at
book+EZ big+2SG
'your big book'

(28) کتاب م
Ketāb am
book-my
'my book'

(29) کتاب من
Ketābe man
book+EZ I/me
'my book'
کتابم *
ketābe am
book + EZ - my
possessives + demonstrative determiner

این کتاب مریم
in ketābe maryam
this book of Maryam
inchoative + causative + light verb construction

رادیو باز شد.
Rādiyo bāz šod.
radio open became.
‘The radio opened.’

مادر بچه را آرام کرد.
mādar bače rā ārām kard.
mother child DOM silent make - Past
‘The mother silenced the child.’

بچه آرام شد.
bače ārām šod.
child silent become
‘The child became silent.’

بچه توسط مادرش آرام شد.
bače tavassote mādar aš ārām šod.
child by mother - his silent become
‘The child became silent by his mother.’

nominalization + light verb construction

من بازی کن را دیدم.
Man bāzī kon rā didam.
I play do RA saw
‘I saw the player.’

inchoative + verbal noun + light verb construction

علی ساسان را شکست داد.
Ali Sāsān rā šekast dād.
Ali Sasan - DOM defeat GIVE - Past
‘Ali defeated Sasan.’

ساسان شکست خورد.
Sāsān šekast xord.
Sasan defeat COLLIDE - Past
‘Sasan was defeated./ Sasan suffered defeat.’

verbal noun + inflection

* اهداء
ehdāhā
giving - Pl

علي کتاب را به ساسان اهداء کرد.
Ali ketāb rā be sāsān ehdā kard.
John book DOM to Sasan giving do - past
‘Ali gave the book to Sasan.’

verbal noun

این اهداء
* in ehdā
this giving

process noun + light verb construction

علی با ساسان حرف زد.
Ali ba Sāsān harf zad.
Ali with Sasan talk BEAT - Past
‘Ali talked to Sasan.’

حرفهای علی با ساسان
harfhaie Ali ba Sāsān
talks - EZ Ali with Sasan
‘Ali’s talks with Sasan’
**auxiliary placement**

(46) من کتاب را نوشته بودم
Man ketāb rā nevešte budam.
I book DOM written was-1st-Sg
'I had written the book.'

(47) من کتاب را بودم نوشته.
Man ketāb rā budam nevešte.
I book DOM was-1st-Sg written

(48) من کتاب را خواهم نوشت.
Man ketāb rā xāham nevešt.
I book DOM will-1st-Sg wrote
'I will write the book.'

(49) من کتاب را نوشته خواهم.
Man ketāb rā nevešt xāham.
I book DOM wrote will-1st-Sg
'I will write the book.'

(50) من خواهم کتاب را نوشت.
Man xāham ketāb rā nevešt.
I will-1st-Sg book DOM wrote
'I will write the book.'

(51) من این کار را انجام خواهم داد.
Man in kār rā anjām xāham dād.
I this job DOM performance will-1st-Sg gave
'I will do this job.'

(52) من این کار را انجام داد خواهم.
Man in kār rā anjām dāde xāham.
I this job DOM performance gave will-1st-Sg
'I will do this job.'

(53) من این کار را انجام داده بودم.
Man in kār rā anjām dāde budam.
I this job DOM performance given was-1st-Sg
'I had done this job.'

(54) من این کار را انجام بودم داده.
Man in kār rā anjām budam dāde.
I this job DOM performance was-1st-Sg given
'I had done this job.'

**auxiliary + clitic**

(55) سیب را خریده بودم
sīb rā xaride budam.
apple+rā bought was-1st-Sg
'I had bought the apple.'

(56) خریده بودم ش.
xaride budam aš.
bought was+1sgS+3sg
'I had bought it.'

(57) خریده یتش بودم
xaride yeš budam
bought+3sg was+1sgS
'I had bought it.'

(58) کتاب را خریدم خرید.
ketāb rā xāham xarid.
book+rā want+1sgS buy
'I will buy the book.'

(59) خریدم خرید.
xāham xarid aš.
want+1sgS+3sg buy
'I will buy it.'

(60) خریدم خرید.
xāham xarid aš.
want+1sgS buy+3sg
'I will buy it.'

**progressive**

**aspect + indicative + progressive + indefinite future**

(61) من روم
miravam.
Ind/Prog/Indef. Fut.-go-1sg
'I go. / I am going. / I will go.'

(62) من رفتم
miraftam.
Ind/Prog
'I used to go. / I was going.'

Draft of Friday 2nd November, 2012, 21:32
### present + continuous incompletive + complex predicate

<table>
<thead>
<tr>
<th>Number</th>
<th>Phrase</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>63</td>
<td>من دارم گوش می‌کنم</td>
<td>I happened to listen. / I was listening.</td>
</tr>
<tr>
<td>64</td>
<td>من گوش می‌کردم</td>
<td>I used to listen. / I was listening.</td>
</tr>
<tr>
<td>65</td>
<td>من گوش دارم</td>
<td>I am listening.</td>
</tr>
<tr>
<td>66</td>
<td>تو گوش داری می‌کنی</td>
<td>You are listening.</td>
</tr>
<tr>
<td>67</td>
<td>او گوش دارد</td>
<td>He/she is listening.</td>
</tr>
<tr>
<td>68</td>
<td>من داشتیم گوش می‌کردیم</td>
<td>I happened to listen. / I was listening.</td>
</tr>
<tr>
<td>69</td>
<td>تو داشتی میرافتی</td>
<td>You had listened. / You were listening.</td>
</tr>
<tr>
<td>70</td>
<td>او داشت میرافت</td>
<td>He/she had listened. / He/she were listening.</td>
</tr>
<tr>
<td>71</td>
<td>man دارم گوش می‌کنم</td>
<td>I have ear Ind/Prog/Indef. Fut.-do-1sg 'I am listening.'</td>
</tr>
<tr>
<td>72</td>
<td>man دارم گوش می‌کردم</td>
<td>I used to listen. / I was listening.</td>
</tr>
<tr>
<td>73</td>
<td>من گوش دارم</td>
<td>I am listening.</td>
</tr>
<tr>
<td>74</td>
<td>تو گوش داری می‌کنی</td>
<td>You are listening.</td>
</tr>
<tr>
<td>75</td>
<td>تو گوش داری می‌کنی</td>
<td>You are listening.</td>
</tr>
<tr>
<td>76</td>
<td>او گوش دارد</td>
<td>He/she is listening.</td>
</tr>
<tr>
<td>77</td>
<td>او گوش دارد</td>
<td>He/she is listening.</td>
</tr>
<tr>
<td>78</td>
<td>من داشتیم گوش می‌کردیم *</td>
<td>I happened to listen. / I was listening.</td>
</tr>
<tr>
<td>79</td>
<td>من داشتیم گوش می‌کردیم</td>
<td>I happened to listen. / I was listening.</td>
</tr>
</tbody>
</table>
(80) من گوش داشتم می‌کرد.
man guš dāštam mikardam.
'I was listening.'

(81) تو داشتی گوش می‌کرد.
to dašti guš mikardi.
you had ear Ind/Prog/Indef. Fut.-do-1sg
'I was listening.'

(82) او داشت گوش می‌کرد.
u dāšt guš mikard.
he/she had ear Ind/Prog/Indef. Fut.-do-1sg
'I was listening.'

(83) او داشت گوش می‌کرد.
u dāšt guš mikard.
he/she had ear Ind/Prog/Indef. Fut.-do-1sg
'I was listening.'

(84) امکانات گوش می‌کرد.
u guš dāšt mikard.
he/she had ear Ind/Prog/Indef. Fut.-do-1sg
'I was listening.'

definite future

(85) من گوش خواهید. می‌کرد.
man guš xāham raft.
'I will go.'

(86) تو گوش خواهید.
to xāham raft.
you futur-sg went.
'You will go.'

(87) او گوش خواهید.
u xāhad raft.
he/she futur-sg went.
'He/she will go.'

definite future + complex predicate

(88) من گوش خواهید کرد.
man guš xāham kard.
'I will do.'

A. List of Phrases Covered/Rejected by the Grammar

'I will listen.'

(89) تو گوش خواهید کرد.
to guš xāhī kard.
you ear futur-sg do
You will listen.'

(90) او گوش خواهید کرد.
u guš xāhad kard.
he/she ear futur-sg do
'He/she will listen.'

perfect

(91) مريم خندیده بود.
Maryam xandide bud.
'Maryam had laughed.'

(92) مريم کتاب را خوانده بود.
Maryam ketāb rā xānde bud.
'Maryam read the book.'

negation

(93) نمیرم را می‌کرد.
nemiravam.
NEG-IND-go-1sg
'I do not go.'

(94) نرم.
neravam
NEG-go-1sg

(95) نرم.
naravam
NEG-(SUBJUNCTIVE-)go-1sg

(96) نرم.
naraftam.
NEG-went-1sg
'I did not go.'
(97) نداشتیم نمی‌رفتم.
nusatīm nemirfātīm.
NEG-Progr NEG-IND-go-1sg
'I am not not going.'
(98) نداشتیم من نرفتم.
nusatīm minarfātīm.
NEG-Progr IND-go-1sg
(99) داشتیم نمیرافتم.
dastātīm nemirfātīm.
Progr NEG-IND-go-1sg

negation + complex predicate

(100) گوش نمی‌کنیم.
gūsh nemīkānim.
ear NEG-IND-do-1sg
'I do not listen.'
(101) گوش نکنیم.
gūsh nekānim
ear NEG-do-1sg
(102) گوش نکنیم.
gūsh nakānim.
ear NEG-(SUBJUNCTIVE-)do-1sg
'I did not listen.'
(103) گوش نکنیم.
gūsh nakānim.
ear NEG-did-1sg
'I did not listen.'

direct object marker

(104) تو را دیدم.
tū rā didām.
you+rā saw+1sgS
'I saw you.'

direct object marker + case

(105) علي فوت گرد.
ālī fūt kard.
Ali fawt kard.
Ali death did
'Ali died.'
(106) علي درا فوت گرد.
ālī rā fūt kard.
Ali DOM fawt kard.
Ali DOM death did
(107) علي فوت را کرد.
ālī fūt rā kard.
Ali DOM fawt DOM kard.
Ali DOM death DOM did

agreement

(108) من خندیدم.
mīn xandidam.
man xandidamS
'I laughed.'
(109) آنها خندیدند.
ānhā xandidand.
they laughed+3plS
'They laughed.'
(110) شما خندیدید.
šomā xandidid.
you laughed+2plS
'You laughed.'
(111) من افتادم.
mīn aftādam.
I fell+1sgS
'I fell.'
(112) کتاب م افتاده.
ketāb m aftādē.
book+1sg fell+1sgS
'My book fell.'
(113) کتاب م افتاد.
ketāb m aftād.
book+1sg fell+3sgS
'My book fell.'
کیف رادیدم
bag RA saw
'I saw the bag.'

کیف چرم رادیدم
leather-EZ RA saw
'I saw the leather bag.'

کیف چرم بزرگ رادیدم
big leather-EZ RA saw
'I saw a big leather bag.'

کیف چرم سوخته رادیدم
friend+1sg leather-EZ RA saw
'I saw my friend's leather bag.'

Draft of Friday 2nd November, 2012, 21:32
پک کتاب به ساسان داده شد.
ye ketāb be sāsān dāde šod.
a book to Sasan given was
‘A book was given to Sasan.’

passive + intransitive verb

مریم با سنگ به دیوار زد.
Maryam bā sang be divār zad.
Maryam with stone to wall hit
‘(Lit.) Maryam hit to the wall with a stone/stones.’

passive + complex predicate

مریم به امید تهیه زد.
Maryam be Omid tohmat zad.
Maryam to Omid slander hit
‘Maryam slandered Omid.’

adjective

علي بزرگ.
Ali bozorg.
Ali big

participle + passive + attributive

مرد دویده.
marde davide
man-EZ run.PAST

مرد در پسته.
marde dar baste
man-EZ door closed
‘a man who has closed doors’

مرد افتاده آمد.
marde sēfāde āmad.
man-EZ fall + -ed came
‘a fallen man came.’

Draft of Friday 2nd November, 2012, 21:32
It was in the cinema that I think she saw Kimiyā.

Yesterday it rained.

I know places that are worth visiting

I see a book (which is) for sale

These guests which seem to intend to leave.

I call and died.

Ali saw a man and laughed. / and he laughed.

Ali saw a man and I laughed.

Ali called and died.

Ali saw a man and laughed. / and he laughed.

Ali saw a man whom you saw laughed.

Ali saw a man whom you saw laughed.

A. List of Phrases Covered/Rejected by the Grammar
این مردی که شما دیدید خندید.

in mardi ke šomā didid xandid.

this man-RESTR COMP you-PL saw-2PL laughed

‘this man whom you saw laughed’

(163)

فلیسه که می دانم کیمیا دید که در مدت است.

filmī ke midunam Kimiyā did qangīn ast.

movie-RESTR COMP DUR-know-1sg that Kimiyā saw sad is

‘The movie whom I know that Kimiyā has seen is sad.’

(164)

نام جزئی + گفته جمعکننده

relative clause + resumptive pronoun

مردی که پیراهن سبز است

mardi ke pirāhān sabz hast

man-RESTR COMP shirt _ green be

مردی که پیراهن سبز است

mardi ke pirāhān sabz hast

man-RESTR COMP shirt _ green be

مردی که پیراهن سبز است

mardi ke pirāhān sabz hast

man-RESTR COMP shirt-EZ _ green be

Draft of Friday 2nd November, 2012, 21:32
Draft of Friday 2nd November, 2012, 21:32
سوالی که فکر می‌کنی مريم چه؟

با کی تو نکن که مريم حرف زد؟

آگر علی مريم را بپیسند عالی می‌شود.

agar Ali Maryam rā bebinad, āli nišavad.

‘If Ali sees Maryam, this would be great.’

Adjective + Negation

مرد نارام

marde nāāram

‘the impatient man’

Comparative

خانه بزرگ

xuneye bozorg

‘the big house’

خانه بزرگتر

xuneye bozorgtar

‘the bigger house’

خانه بزرگترین

xuneye bozorgtarine

‘this biggest house’

خانه بزرگترین

xuneye bozorgtarin

‘this biggest house’

‘the biggest houses’

Participle Adjectives

لیوان شکسته

livāne šekaste

‘the broken glass.’
yes/no questions

آیا علی را دیدی؟
āyā Ali rā didī?
INT Ali RS saw
‘Did you see Ali?’

noun to adjective derivation

کتاب ایرانی را دید.
letābe īrānī rā did.
book.EZ America.Adj RA saw
‘I saw the Iranian book / I saw the book of the Iranian’

کتاب آمریکایی را دید.
letābe āmrikāyi rā did.
book.EZ America.Adj RA saw
‘I saw the American book / I saw the book of the American’

این آمریکایی را دید.
in āmrikāyi rā did.
this American RA saw
‘I saw this American (person from America).’

اراده ایرانی را دید.
irānihā rā did.
Iranian.PL RA saw
‘I saw an Iranians.’
Bibliography


Verzeichnis der Sprachen

Basque, v
English, iv, v, 92
French, iv
German, iv, v
Japanese, v
Maltese, iv
Mandarin Chinese, iv
Modern Standard Arabic, 16
Persian, iv
Spanish, iv
Yiddish, iv

Sachregister

BerliGram, iv
Categorial Grammar (CG), v
Construction Grammar (CxG), v
Embodied, vi
Sign-Based, vi
CoreGram, iv
DanGram, iv
feature
c-cont, 15
hook, 16

Government and Binding (GB), vi
language acquisition, v–vi
Lexical Functional Grammar (LFG), v
Minimalism, vi
Poverty of the Stimulus, v
Universal Grammar (UG), v–vi