# Local Semantics in Head-Driven Phrase Structure Grammar Manfred Sailer\*

# **1** Introduction

Semantic research is generally divided into *lexical semantics* (LS) and *compositional semantics* (CS). LS is concerned with the relation between a semantic functor and its arguments, either in terms of *semantic selectional restrictions* or in terms of *linking*, i.e., the relation between syntactic complements and semantic argument slots. CS focuses on the way in which the semantic contributions of constituents in a sentence are combined to arrive at the interpretation of the sentence. The central notion here is the scope of quantifiers and other operators. This devision of labor in semantics has its parallel in syntax, which, for example, is evident in the A- vs. A-bar syntax of Government and Binding Theory. In the case of syntax, the modularization and the interaction of the two "kinds" of syntax have been studied fairly thoroughly. On the semantic side, however, the relation between the two kinds of semantics is still not so well understood.

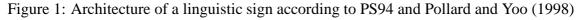
In this paper we will contribute to the study of the LS-CS interface by reviewing two empirical phenomena, linking and selectional restrictions. We will propose a distinction between local and non-local semantics which will be embedded in a general linguistic theory, *Head-Driven Phrase Structure Grammar* (HPSG, Pollard and Sag (1994)). We have chosen HPSG because it is a rigidly formalized linguistic framework (Richter, 2004), and for both LS and CS it offers a number of substantial proposals to build on. In recent years techniques of *underspecified semantics* have become popular within HPSG (Egg, 1998; Egg and Erk, 2002; Copestake et al., 2003; Richter and Sailer, 1999, 2004a). We will demonstrate that these techniques allow us to define a modular devision of the two kinds of semantics within a linguistic sign.

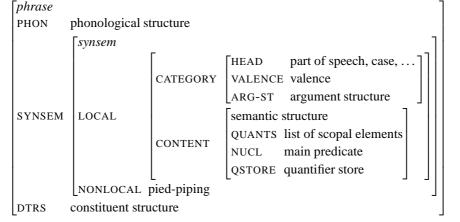
In the rest of the introduction we will characterize what we understand by a *local* phenomenon and present the structure of a linguistic sign as given in the standard form of HPSG. In Section 2 we will discuss two local semantic phenomena. In Section 3 we will present a concise introduction to the framework of *Lexical Resource Semantics* which can incorporate the LS-CS distinction. A conclusion will round off this paper in Section 4.

### 1.1 Local Phenomena

We will try to illustrate what we understand by a *local* phenomenon in contrast to *nonlocal* phenomena. *Local* phenomena are typically determined by lexical properties and concern the relation between a head and its dependents. In contrast to this, *nonlocal* phenomena are largely independent

<sup>\*</sup>I am grateful to Olivier Bonami, Frank Richter, Jan-Philipp Soehn, and to the CSSP reviewers for their comments. Thanks also to Guthrun Love for her help with the English.





of concrete lexical items, referring instead to structural properties. They may also go beyond direct head-dependent relations.

In syntax, categorial selection and case assignment are local phenomena. On the semantic level the corresponding phenomena are selection, in particular semantic selectional restrictions, and the assignment of thematic roles. On the syntax-semantics interface we find argument structure alternations such as dative shift, passive and such like and linking, i.e. the mapping between semantic arguments and syntactic complements. These phenomena are all local in the sense that one only needs to consider the projection of a head in order to formulate the regularities. On the other hand they typically involve a high degree of lexical idiosyncrasy.

Let us next turn to a number of nonlocal phenomena. In syntax extraction is by far the most extensively discussed nonlocal topic. Similarly pied-piping and, depending on the theory, scrambling fall in this category. Analogous nonlocal semantic phenomena are the scope of semantic operators (such as negation, quantifiers, or tense). Those phenomena are typically accounted for by general principles of the grammar. Often they apply to larger syntactic domains, in particular they may be "unbounded".

### **1.2** The architecture of HPSG

One of the major empirically motivated changes from the first presentation of HPSG in Pollard and Sag (1987) to recent versions of the theory, starting with Pollard and Sag (1994) (PS94), is the incorporation of the local-nonlocal distinction within the architecture of a linguistic sign. This has been quite successful for syntax, but less so for semantics. In Figure 1 we will outline the architecture of a linguistic sign of PS94, with slight modifications in the semantics.

HPSG signs comprise the phonological structure (as value of the PHON attribute), the constituent structure as the DAUGHTERS (DTRS) value and a so called *synsem* structure as its SYNSEM value. In the latter the NONLOCAL value may, among others, specify whether or not a sign contains a gap. In the LOCAL value we find the part-of-speech (within the HEAD value), and the syntactic valence.

There is also a CONTENT attribute, whose value contains the entire semantic structure of a sign. An HPSG-specific representation is very often chosen for the semantic structure. The CONTENT value of a verb contains a specification of the verb's semantic relation and of its arguments within the NUCL(EUS) value. The QUANT(IFIER)S list contains quantifiers which have scope over the nucleus. Nouns do not have the attributes NUCL and QUANTS; instead their CONTENT value contains an INDEX feature which expresses the referential index of the noun and a RESTRICTIONS set. The proposal also incorporates a Cooper store mechanism (Cooper, 1975, 1983), encoded with the QSTORE value.

In PS94 the attribute QSTORE was defined on the sort *sign*. Therefore the surface position of a quantifier determined its smallest possible scope. Consequently the *de dicto* reading could not be derived in sentences such as *A unicorn appears to be approaching*. (see PS94, p. 328). This empirical deficiency was solved in Pollard and Yoo (1998) by incorporating QSTORE inside the *local* structure. Building on this, Przepiórkowski (1998) argues for including QSTORE inside the CONTENT value. In Figure 1 we adopted this suggestion.

The argument structure (ARG-ST) of a sign contains the SYNSEM values of the signs it selects. This reveals two insights: Firstly, properties of the phonology or of the constituent structure of a selected element cannot be selected for. Secondly, since *synsem* contains both the syntactic category and the semantics, PS94 acknowledges that a selector can impose categorial as well as semantic restrictions on the selected elements.

The HPSG architecture of a linguistic sign assumes a sharp distinction between syntactic category (realized within *synsem* as the CAT(EGORY) value) and syntactic structure (within the DTRS value). In fact the syntactic phenomena characterized above as local are all treated in PS94 at the word level and concern relations within the SYNSEM value of a word. Unbounded dependencies, on the other hand, are treated by a global principle of the grammar. While this distinction is made clear in the syntax, all of the semantics are gathered within the LOCAL value.

In the following section we will look at local semantic phenomena and we will demonstrate that there is no empirical motivation for having quantifiers or other semantic operators as part of the LOCAL value, where they are visible for selectors. Consequently we will draw a line between local semantics and logical form which will be analogous to the division between syntactic category and constituent structure.

# **2** Local Semantics

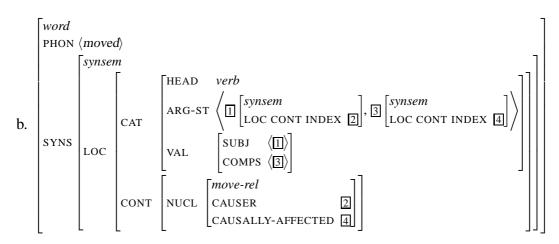
In this section we will discuss two local semantic phenomena, linking and semantic selectional restrictions. We will demonstrate that the architecture in Figure 1 is not restricted enough since neither of these phenomena manifest a need to refer to semantic operators or scope.

### 2.1 Linking

Linking is the mapping between semantic roles and syntactic complements. Our discussion will focus on linking constraints as formulated in Koenig and Davis (2003).<sup>1</sup> To illustrate the way in which linking is expressed within HPSG consider the example in (1a). In (b) we will describe the word *moved* as it occurs in (a).

(1) a. Pat moved the car.

<sup>&</sup>lt;sup>1</sup>See e.g. Davis and Koenig (2000) for an earlier version of their theory and Kordoni (2003) for an overview of the HPSG literature on linking.



The description in (1b) combines the architecture of a linguistic sign outlined in Figure 1 with the proposal of Koenig and Davis (2003). The description shows that the argument structure of the verb contains two *synsem* objects. The entire semantic contribution of the two arguments can be found in these *synsem* objects. However, only the INDEX information is relevant for linking. In Koenig and Davis (2003) linking constraints are expressed as implicational constraints of the form in (2).

(2) Linking constraint (adapted from Koenig and Davis (2003)):

$$\begin{bmatrix} word \\ s \ L \begin{bmatrix} CAT \left[ ARG-ST \left\langle NP, \dots \right\rangle \right] \\ CONT \left[ NUCLEUS \ cause-rel \right] \end{bmatrix} \end{bmatrix} \Rightarrow \begin{bmatrix} s \ L \begin{bmatrix} CAT \left[ ARG-ST \left\langle \left[ L \ CONT \ INDEX \ \square \right], \dots \right\rangle \right] \\ CONT \left[ NUCLEUS \ CAUSER \ \square \right] \end{bmatrix} \end{bmatrix}$$

(where *move-rel* is a subsort of *cause-rel*)

Stated informally, this constraint expresses that if a word has an NP as its first element on the ARG-ST list and introduces a semantic constant of sort *cause-rel*, then the index of the first syntactic argument and the value of the CAUSER thematic role are identical.

With this linking constraint, the identity between the INDEX value of the first element in the ARG-ST list of *move* in (1b) and the CAUSER value need no longer be stipulated, as it follows directly from the linking theory. Analogous linking constraints will ensure the identity between the second complement and the CAUSALLY-AFFECTED value.

Linking constraints such as in (2) indicate that linking is perceived as a local phenomenon in the sense characterized above: Firstly, linking constraints are formulated for words. Secondly, linking involves only a head and its direct dependents. Thirdly, the kind of information used in linking constraints are the semantic constant contributed by the head, the thematic roles which are defined for this constant, the syntactic category of the selected elements and their indices.

It is reasonable to assume that this locality applies to linking in general. Nonetheless, the architecture of linguistic signs as outlined in Figure 1 would also allow for linking constraints which refer to the particular quantificational nature of the complement. In (3) we will state the antecedent of a hypothetical linking constraint. This constraint would determine the linking of a *cause-rel* predicate in the case in which its first syntactic argument contains an unretrieved universal quantifier (i.e., has a *forall* object in its QSTORE).

(3) Hypothetical linking constraint:

$$\begin{bmatrix} \text{SYNS LOC} \begin{bmatrix} \text{CAT} & [\text{ARG-ST} \left< [\text{LOC CONT QSTORE} \left< \dots, [\text{forall}], \dots \right> ], \dots \right> \end{bmatrix} \end{bmatrix} \Rightarrow \dots \\ \begin{bmatrix} \text{CONT} & [\text{NUCLEUS cause-rel}] \end{bmatrix} \end{bmatrix}$$

(if the first argument contains an unretrieved universal quantifier, then ...)

We require that an adequate structure of signs should exclude the formulation of this kind of linking constraint. The cause of the problem with the current HPSG architecture lies in the absence of a strict separation between LS and CS.

### 2.2 Semantic Restrictions

While linking is a widely discussed topic within HPSG, semantic restrictions are largely ignored (with the exception of Androutsopoulos and Dale (2000)). Semantic restrictions have been discussed in comparison to categorial selection in Chomsky (1965), yet their status in grammar remains unclear. We cannot develop a theory of semantic restrictions here, but we will demonstrate that they are a local phenomenon, and do not refer to clausal semantic properties. We will address two kinds of semantic restrictions based on a distinction exemplified in Lang (1994): *sortal* and *selectional restrictions*.

#### 2.2.1 Sortal Restrictions

It has been established that sortal differences are important for grammar (see Dölling (1994), Chierchia (1998), Krifka (2003) among others). In this subsection we will consider primarily the analysis in Krifka (2003). Krifka assumes a semantic ontology which contains kinds, groups, individuals and numbers. These are encoded as semantic types. In a typed semantic representation language determiners and predicates can impose type requirements on their semantic arguments. This can account for the fact that a bare singular noun cannot occur as an argument of a verb which requires a kind (see (4)).<sup>2</sup>

(4) \* Dodo is extinct.

At first glance the data in (5a) seem to suggest that the predicate *be extinct* can restrict the quantificational status of its complement. However, adopting the richer semantic ontology, the contrast follows from the fact that *every dodo* expresses a quantification over individuals. In the web example (5b) the quantification is over kinds, and consequently the universally quantified NP is compatible with the type requirements of the verb.

- (5) a. The dodo/ \*Every dodo is extinct.
  - b. Wenn noch vor zehn Jahren jede Art des Positivismus als ausgestorben ... galt,
    'While 10 years ago every kind of positivism was still considered ... extinct, ...'

A predicate does not restrict the quantificational aspects of its arguments, but the type of its argument can be restricted. Quantifiers may have the effect of type shifting which accounts for the apparent sensitivity to particular determiners in (5a).<sup>3</sup>

<sup>&</sup>lt;sup>2</sup>The dodo was a flightless bird of Mauritius, extinct in the 17th century.

<sup>&</sup>lt;sup>3</sup>Analogously we expect that the inherently distributive or collective nature of predicates such as *die* and *besiege* can be captured respectively by the subtle sortal distinctions.

#### 2.2.2 Selectional Restrictions

Besides the sortal restrictions discussed in the previous subsection there are other more fine-grained semantic restrictions which a verb can impose on its complements. They are usually called *selectional restrictions*. In (6) an example is presented.

- (6) a. Hans pflückte eine Pusteblume. Hans picked a dandelion
  - b. ?? Hans pflückte ein Buch aus dem Regal. Hans picked a book from the shelf

The German verb *pflücken* and its English translation *pick* impose the same sortal restrictions on their argument. Nonetheless, *pflücken* is restricted to flowers and fruits. As illustrated in (b), this is not the case for English *pick*. Note also that selectional restrictions are independent of the occurrence of particular quantifiers or other semantic operators:

(7) Hans hat zwei/ alle Pusteblumen gepflückt/ pflücken wollen. Hans has two/ all dandelions picked/ pick want
'Hans picked/ wanted to pick two/ all dandelions.'

Within generative grammar the oddness of sentences such as (6b) is considered to follow from world knowledge rather than from the grammar (see for example Bennis and Hoekstra (1989, p. 23). The reason for this is that the context may improve the data:<sup>4</sup>

- (8) a. ?? Tom ate a keyboard.
  - b. Tom cannot eat a keyboard. (Androutsopoulos and Dale, 2000, p. 15)

It should be noted that the repair effect of the context in (8b) is systematic, while the selectional restrictions are idiosyncratic. I.e. it is an idiosyncratic property of *eat* to be compatible only with food, but it is a general property of the negation to allow for the violation of selectional restrictions. This shows that selectional restrictions qualify as a local semantic phenomenon: They are idiosyncratic properties of a lexical item and involve the semantic properties of a head and its dependents, but they are indifferent with respect to semantic operators.

There is a difference between sortal restrictions and selectional restrictions. Chomsky (1965) already distinguishes between two kinds of semantic features, one group being of relevance to the grammar, the other being more pragmatic in nature. We follow Lang (1994) in defining this distinction in terms of sortal versus selectional restrictions. Selectional restrictions are more subtle — for example, they allow us to distinguish between flowers and books, and they can be violated more readily.

To our knowledge Androutsopoulos and Dale (2000) is the only study which proposes an account of selectional restrictions within HPSG. The authors outline two different possible analyses, depending on whether the phenomenon is treated as primarily pragmatic or primarily semantic. For both analyses the verb needs only to have access to the INDEX value of its complements. In the first case it adds a restriction of this index to the context. In the alternative analysis they assume a complex

<sup>&</sup>lt;sup>4</sup>Classical examples for this argument can be found in Chomsky (1965, p. 158).

	[phrase		1				
	PHON	phonological structure					
		[synsem ]					
	SYNSEM	LOCAL	CATEGORY local syntactic structure				
			CONTENT [local semantic structure]				
		NONLOCAL pied-piping					
	LOGICAL-FORM [logical expressions associated with the sign]						
	DTRS constituent structure						

Figure 2: The architecture of semantics

hierarchy below the sort *index*. The verb *eat* then requires a complement whose INDEX value is a subsort of *edible*.

The HPSG architecture in Figure 1 makes the index of the complements available to the verb. However, the rest of the semantic contribution of the complements is also accessible. Thus there could in principle be a verb that can only take a universally quantified subject. Clearly such a selectional restriction is as implausible as the hypothetical linking constraint in (3) and should therefore be excluded by the structure of linguistic entities.

In this section we have looked at local semantic phenomena. We have demonstrated that the semantic information referred to in the description of these phenomena includes the basic semantic constant of a word, and its index. Yet, in the current architecture of semantics in HPSG, all of the semantics of a word are available for imposing lexical restrictions on the relation between a head and its dependents.

### **3** Lexical Resource Semantics

Lexical Resource Semantics (LRS) is an alternative system for combinatorial semantics in HPSG. Richter and Sailer (2004a) give a detailed presentation of the framework. LRS combines techniques of underspecified semantics (Reyle, 1993; Bos, 1996; Pinkal, 1996) with the properties of an HPSG grammar. LRS departs from the HPSG tradition in that it assumes a standard semantic representation language such as Ty2 (Gallin, 1975) as the logical form of a sentence. Expressions of this representation language are encoded as objects of a sort *meaningful-expression (me*, see Sailer, 2003). In LRS the semantic contribution of a sign is not considered a single *content* object, it is rather conceived of as a list of subexpressions of the final logical form. This kind of semantic representation is called *discontinuous* in Richter and Sailer (2004a).<sup>5</sup>

In LRS we can establish a distinction between local and non-local semantics which is analogous to the distinction between syntactic category and constituent structure. The resulting architecture is presented in Figure 2.

We assume two attributes for semantics: the local semantic representation appears as the CONTENT value within LOCAL. The clausal semantics, i.e., the logical form of a clause, constitutes the LOGICAL-FORM (LF) value. We will briefly present how nonlocal semantics is dealt with in LRS, and then explain our assumptions about local semantics.

<sup>&</sup>lt;sup>5</sup>This discontinuous approach proved successful in the analysis of a number of nonlocal semantic phenomena: German multiple interrogatives (Richter and Sailer, 2001), Polish negative concord (Richter and Sailer, 2004a,b), scope ambiguity in Dutch (Bouma, 2003), and Afrikaans tense phenomena (Sailer, 2004).

### 3.1 Nonlocal Semantics

This subsection will be exclusively devoted to the LF value. We will go through a simple example which illustrates the combinatorial mechanism of LRS. The LF attribute in Figure 2 takes values of sort *lrs*. In (9) we will give the appropriateness conditions for this sort.

(9) Appropriateness conditions of the sort *lrs*:

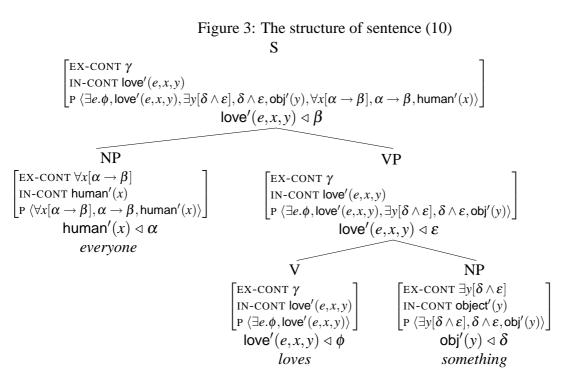
```
lrsEXTERNAL-CONTENTmeINTERNAL-CONTENTmePARTSlist(me)
```

The PARTS list contains all the subexpressions which are contributed by a sign. In an utterance, these subexpressions together constitute the EXTERNAL-CONTENT (EX-CONT) value, i.e. the overall logical form associated with that sign. The attribute INTERNAL- CONTENT (IN-CONT) is needed for the definition of the combinatorial principles of LRS. It specifies the scopally lowest expression in a head projection.

In the following we will illustrate how to derive the two readings of the sentence in (10).

(10) a. Everyone loves something.

```
b. \forall \exists-reading: \forall x[\text{human}'(x) \rightarrow \exists y[\text{object}'(y) \land \exists e[\text{love}'(e, x, y)]]]
\exists \forall-reading: \exists y[\text{object}'(y) \land \forall x[\text{human}'(x) \rightarrow \exists e[\text{love}'(e, x, y)]]]
```



In Figure 3 we summarized our analysis of sentence (10). The semantic contributions of the words are indicated on the leaves of the tree. The PARTS list (P) of the NP *everyone* contains the universal quantifier, the variable bound by this quantifier, the restriction to humans and the implication, i.e.,

the indication of how the restrictor and the nuclear scope should be connected in the interpretation. However the restrictor and the nuclear scope are not fully specified, which we mark by lower case Greek letters ( $\alpha$ ,  $\beta$ , ...). The lexical entry of *everyone* also specifies that the expression human'(x) must be part of the quantifier's restrictor.<sup>6</sup> We indicate this by the constraint human'(x)  $\triangleleft \alpha$  in the figure. The relation " $\triangleleft$ " encodes subexpressionhood. The semantic contribution of *something* is analogous to that of *everything*. The verb contributes the semantic constant love' together with the argument variables. Note that we assume an eventuality argument *e* for verbs. The verb also contributes the existential quantification over this variable.<sup>7</sup>

The LF value of a phrase is fully determined by the LF values of its daughters and the way in which the daughters are syntactically combined.<sup>8</sup> This is regulated in the SEMANTICS PRINCIPLE (SP). The SP states that the EX-CONT and the IN-CONT values of a phrase and its head daughter are identical. The PARTS list of the phrase consists of all the elements of the PARTS lists of the daughters. In addition the SP specifies further requirements depending on the syntactic structure. For our example, one such requirement is relevant: if the nonhead is a quantified NP. then the IN-CONT value of the head is a subexpression of the nuclear scope of the quantifier in the EX-CONT value of the nonhead.

The effect of the SP has already been integrated in the tree in Figure 3. Note that the expression love'(e, x, y) is required to be whithin the scope of both quantifiers. The relative scope of the two quantifiers is not constrained by the grammar. The EX-CONT value of the sentence must consist of all the elements of its PARTS list, and must respect the indicated subexpression constraints. This leaves two options for the EX-CONT value,  $\gamma$ : the two readings given in (10b).

The mechanism presented so far is very similar to other systems which build on techniques of underspecified semantics. Within HPSG, *Minimal Recursion Semantics* (MRS, Copestake et al., 2003) is particularly popular. It should be emphasized that the main argument of this paper applies to MRS just as well as to LRS. In fact the locality of semantic selection is sometimes mentioned in MRS publications. Nonetheless, no detailed argumentation has been presented so far, nor has this locality been reflected in the linguistic architecture. We have chosen LRS because it uses a standard semantic representation language, which allows us to integrate logical forms from the literature directly.

#### 3.2 Local Semantics

After this brief presentation of the combinatorial mechanisms of LRS we will indicate how local semantics can be integrated into the system. We will assume that the values of the attribute CONTENT in Figure 2 are objects of the sort *content*. In (11) we will specify the sort hierarchy and the appropriateness conditions below the sort *content*.

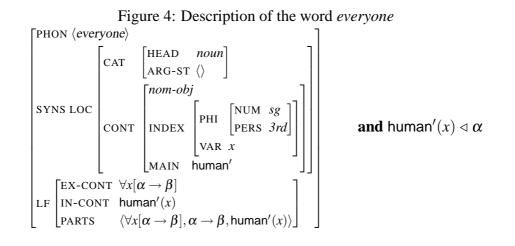
(11) The sort <i>content</i> :	content	
	INDEX	extended-index
	MAIN	me

Our CONTENT values are a considerably simplified compared to the structures in PS94. This is, of course, partly due to the fact that some of the semantic burden is transferred into the nonlocal

<sup>&</sup>lt;sup>6</sup>The PARTS lists are abbreviated in this paper for better readability. In fact, the PARTS list of *everyone* also contains the expressions x, human', human'(x).

<sup>&</sup>lt;sup>7</sup>For simplicity we assume the narrowest possible scope of  $\exists e$  and ignore its potential scopal interaction with other quantifiers.

<sup>&</sup>lt;sup>8</sup>The LF value is a semantic representation, not the semantic denotation of a sign. Thus, LRS obeys "systematicity" (Halvorsen, 1995), but is not strictly compositional.



semantics. All *content* objects have an INDEX and a MAIN attribute. The value of the MAIN attribute of a word is the major semantic constant contributed by this word. The attribute INDEX has values of the sort *extended-index*, which we will define in (12)

(12) The sort *extended-index*: *extended-index* PHI *index* VAR *me* 

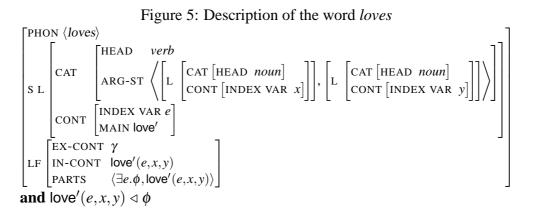
This new index has two attributes, PHI and VAR. The values of PHI are structured as prescribed by the sort *index* in PS94, i.e., they include the attributes PERSON, NUMBER and GENDER which encode the traditional " $\phi$ -features". The value of the attribute VAR contains an expression of the semantic representation language. In simple cases it is an individual variable. The VAR value corresponds intuitively to the referential semantic argument of the MAIN value. Thus, in the case of a quantified NP, the VAR value is the variable which is bound by the quantifier. For verbs the VAR value is an eventuality variable.

The PS94 theory draws heavily on the INDEX, in particular for Binding Theory, which we would like to preserve. However, in PS94 verbs do not have an index. In our approach, following Soehn (2003), we have an attribute INDEX defined for all parts of speech, including verbs. While an eventuality variable is needed in the semantics, it is not clear whether  $\phi$ -features are necessary. In order to establish this we could assume that the PHI value of verbs is of a new subsort of *index*, called *no-phi*. Alternatively we would have to make sure that the PHI values of verbs do not play any role in the grammar. In both cases the verb's PHI value would not be mentioned in the lexicon.

We will illustrate the interplay between the CONTENT value and the LF value with the description of two words in Figures 4 and 5. The quantified NP *everyone* in Figure 4 has a third person singular PHI value. Its VAR value (x) is the variable bound by the quantifier. The MAIN value expresses the main semantic constant contributed by the NP: the restriction to humans.

The INDEX PHI value of the verb in Figure 5 is not specified. Its VAR value is an event variable e and its MAIN value is the constant love'. Since the VAR value of the verb's complements are part of the *synsem* structures on the verb's ARG-ST list, they are accessible for the identification of the argument positions of love' on the PARTS list (love'(e, x, y)). The semantic typing of love' guarantees the correct semantic type for the VAR values, which are, in the present case, individuals.<sup>9</sup>

<sup>&</sup>lt;sup>9</sup>See section 3.2.1 for a refinement.



This brief illustration demonstrates that the information necessary for linking and sortal restrictions are present in the new architecture for semantics. Our major concern in Section 2 was that the traditional HPSG architecture does not adequately restrict the kinds of linking constraints or selectional restrictions which can be imposed by a head. In particular it was possible to write linking constraints which referred to the presence of quantifiers in the complements (see (3)). In the new architecture a constraint of this kind can no longer be formulated, since the quantificational impact of a complement is located entirely in its LF value, and thus is not accessible to the selecting head.

It is in line with the literature on linking and also with the theory of selectional restrictions in Androutsopoulos and Dale (2000) to assume that semantic constants are ordered hierarchically. In the HPSG encoding of Ty2 (Sailer, 2003), the sort *me* has a subsort *constant* which has a maximally specific subsort for each constant of Ty2. It is very natural to incorporate a more elaborated sort hierarchy below *constant*. To illustrate this, we will introduce a new subsort of *constant*, called *cause-rel*. All semantic constants whose first semantic argument can be interpreted as a causer will be subsorts of this new sort. Linking constraints can then refer to these more general supersorts in the usual way.

This sort hierarchy below *constant* is not only the basis for generalization on linking, it can also be seen as a conceptual organization of the constants. The first purpose is mainly fulfilled by the ordering of verbal predicates, the second by the ordering of nominal predicates. Thus there will be a sort *food* which will have *apple*, *chocolate*, etc as its subsorts. Such conceptual hierarchies are widely used in computational linguistic applications such as GermaNet (Kunze and Wagner, 2001). Since the semantic constant of a complement is now visible to a selecting head, selectional restrictions can be expressed. We leave it to further research to determine how these restrictions will be spelled out, i.e. whether in terms of the semantic or the pragmatic properties of the head.

After this general outline of local semantics in LRS, we will elaborate on some details in the following subsections. In Section 3.2.1 we will identify cases in which the INDEX VAR value of a syntactic dependent does not appear directly as a semantic argument of the head. Section 3.2.2 will be concerned with the distinction between MAIN and IN-CONT. Finally, in Section 3.2.3 we will investigate the potential objection that, contrary to our original goal, the proposed architecture allows heads to impose conditions on which quantifiers may appear in their argument positions.

#### Figure 6: Local semantics of the/every car

Γ	Γ	NUM	sg ]]]	
INDEX	PHI	GEN	neutr	
		PERS	3rd ]	
	VAR .	x <sub>e</sub>		
MAIN	$car'_{et}$		-	

#### 3.2.1 Complement INDEX versus Argument Type

In the simple example in (10) the verb *loves* occurs as outlined in Figure 5. Here, the INDEX VAR values of the syntactic arguments appear as semantic arguments of the MAIN constant love'. This is, however, not the case in general. Consider for example the logical forms of the following sentences one of which contains a definite NP, the other a universally quantified NP.<sup>10</sup>

(13) a. Mary likes the green car.

 $\exists e[\mathsf{like}'(e, m, \iota x[\mathsf{car}'(x) \land \mathsf{green}'(x)])]$ 

(where  $\iota x_{\tau}[\phi]$  denotes an individual *a* of type  $\tau$  such that  $[[\lambda x.\phi]](a) = 1$  if there is exactly one such individual, otherwise the denotation is undefined.)

b. Mary likes every green car.  $\forall x [[car'(x) \land green'(x)] \rightarrow \exists e [like'(e, m, x)]]$ 

For both direct object NPs we assume the same local semantic structure, as given in Figure 6. In the logical form the VAR value x is bound by an operator which is introduced by the determiner: the iota-operator in the case of *the*, and the universal quantifier in the case of *every*. Since the iota-operator conserves the semantic type, both x and  $tx[car'(x) \land green'(x)]$  are of type e, and, thus, compatible with the type requirements of the constant like'.

As a consequence, the lexical entry of a verb will require that the VAR values of its syntactic arguments occur *inside* the semantic argument slots of its MAIN constant. However, they do not need to be *identical* to these arguments slots. To illustrate this, consider the outline of the lexical entry of *like* in Figure 7.

In this lexical entry the argument slots of the MAIN value are not filled explicitly, instead of this, it is merely specified that some expressions of type e will appear here. The linking information is, however, preserved, since we state that the VAR value of the first syntactic argument must be a subexpression of  $\alpha$  and the VAR value of the second syntactic argument must be a subexpression of  $\beta$ .

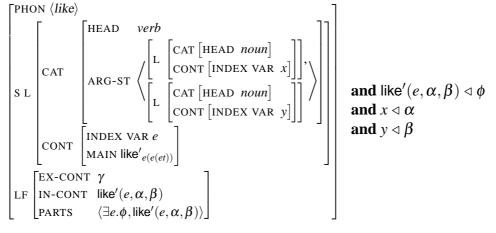
The case of definite NPs is the simplest instance of a mismatch between the VAR value of a syntactic argument and the corresponding semantic argument slot. The following two sentences, quoted from Krifka (2003), can be treated similarly.

(14) a. At the meeting, Martians presented themselves as almost extinct.

b. At the meeting, Martians claimed [to be almost extinct].

<sup>&</sup>lt;sup>10</sup>The semantics of the definite NP is a simplified version of the proposal in Krifka (2003).

Figure 7: Outline of the lexical entry of the verb like



The subject *Martians* binds respectively a reflexive pronoun or an unexpressed embedded subject. The standard HPSG analysis assumes INDEX identity between the subject *Martians* and the reflexive pronoun *themselves* in (14a), and between *Martians* and the unrealized subject of the infinite VP in (b). On the other hand, as Krifka (2003) argues, the sortal restrictions of the predicates are not compatible with each other, i.e., *present* and *claim* require an individual (or a group individual) as its first semantic argument. The predicate *be extinct*, however, requires a kind.

Since binding is expressed as the identity of INDEX values, we will have to demonstrate that, in our approach, the binder and the bindee can indeed have the same VAR values in our approach. The noun *Martian* has an individual variable as its VAR value. Following Krifka (2003) we assume that the plural operator can have the effect of creating a group individual or a kind. In both cases, the operator will bind the variable of the noun's VAR value. Thus, while *Martians* and *themselves* have different EX-CONT values, they can still have identical VAR values.

These examples served to illustrate the distinction between the VAR value of a syntactic argument and the corresponding semantic argument slots of predicates. The present discussion relies heavily on the possibility of integrating standard semantic representations into HPSG. For example, with the semantic representations used in PS94, it would not be obvious how the data should be represented.

#### 3.2.2 MAIN versus INTERNAL-CONTENT

The simple example in (10) has also glossed over another important distinction which we want to capture in LRS. In the case of *everyone* and *loves* mentioned above, the MAIN value was very similar to the IN-CONT value. The scope possibilities of opaque predicates such as German *fehlen* (*be missing*) indicate that this need not be the case. In (15) we will give an example with two possible readings. In (16) we will outline the lexical entry of the verb *fehlen*.

(15) Eine Schraube fehlt.

a screw is missing

*de re*-reading:  $\exists x [screw'(x) \land \exists e [be-missing'(e, \lambda P. P(x))]]$ *de dicto*-reading:  $\exists e [be-missing'(e, \lambda P. \exists x [screw'(x) \land P(x)])]$  (16) Outline of the lexical entry of *fehlen* (*be missing*):

$$\begin{bmatrix} PHON \langle fehlen \rangle \\ SYNS \begin{bmatrix} CAT \begin{bmatrix} HEAD \ verb \\ ARG-ST \langle \left[ LOC \begin{bmatrix} CAT \begin{bmatrix} HEAD \ noun \\ CONT \begin{bmatrix} INDEX \ VAR \ e \\ MAIN \ be-miss.' \end{bmatrix} \right] \end{pmatrix} \end{bmatrix} \end{bmatrix}$$
and be-miss.'(...)  $\triangleleft \phi$   
and ` $P(x) \triangleleft \delta$   
$$LF \begin{bmatrix} EX-CONT \ \gamma \\ IN-CONT \ 'P(x) \\ PARTS \ \langle \exists e.\phi, be-miss.'(e, `\lambda P.\delta), `\lambda P.\delta, `P(x) \rangle \end{bmatrix}$$

Note that the MAIN value of the verb *fehlen* is the constant be-missing' but this constant does not appear in the IN-CONT value. Instead, the scopally lowest subexpression contributed by the verb *fehlen* is the expression P(x). In accordance with what we have said above about the SP, when the verb combines with the subject *eine Schraube* (*a screw*) as in (15), the SP only requires that the verb's IN-CONT value be in the scope of the quantifier. Since this IN-CONT value is the expression P(x), this requirement is met in both readings. On the other hand, in the *de dicto* reading, the main semantic constant of the verb be-missing' is not in the scope of the quantifier.<sup>11</sup>

#### 3.2.3 Less Constrained than Intended?

It seems that we have achieved our goal of constructing a local semantics which does not give access to the quantificational behavior of the selected elements. In particular linking cannot made dependent on quantificational aspects of the syntactic arguments under the reasonable assumption that the antecedent of a linking constraint should only mention the SYNSEM value of a word.

It should be noted, however, that the EX-CONT value of a word may also impose conditions on the overall semantics. Since EX-CONT is part of LF, operators and quantifiers are accessible there. For example, if we specify the EX-CONT value in the lexical entry of a verb accordingly, the verb could enforce narrow scope for one of its arguments. A corresponding hypothetical lexical entry is outlined in (17).

(17) Parts of a hypothetical lexical entry:

```
\begin{bmatrix} word \\ \text{SYNS LOC CAT ARG-ST} \left\langle \dots \left[ \text{LOC CONT INDEX VAR } x \right] \dots \right\rangle \\ \text{LF} \left[ \text{EX-CONT } \dots \exists x [\dots] \right] \end{bmatrix}
```

The lexical entry in (17) specifies that the VAR value of a syntactic argument (x) must be bound by a certain quantifier within the EX-CONT value of the verb.

In this context it becomes relevant that the EX-CONT domain of a head is relatively restricted. For nouns it does not extend beyond the operator which binds the noun's VAR value. For verbs the

<sup>&</sup>lt;sup>11</sup>To deal with cases such as '*A unicorn appears to be approaching*' the SP in Richter and Sailer (2004a) specifies that the IN-CONT value of a raising verb is identical with the IN-CONT value of its infinitival complement. Thus we account for the narrow scope readings without a QSTORE mechanism or a QUANTS list in LOCAL. This principle also ensures that the IN-CONT values are shared in verbal complexes in German and other languages which are analyzed as instances of argument raising in HPSG (Hinrichs and Nakazawa, 1989).

EX-CONT is clause-bound. Syntactic arguments can in general take wider scope than the EX-CONT of their head. This is illustrated with the examples in (18).

- (18) a. [A representative from every city] was present.  $\forall y [\text{city}'(y) \land \exists x [\text{representative}'(x) \land \text{from}'(x, y) \land \text{be-present}'(x)]]$ 
  - b. Peter believed [that someone from Spain had called]  $\exists x [from-Spain'(x) \land believe'(p, `call'(x))]$

In the logical forms in (a) and (b) we have underlined the EX-CONT value of the noun *representative* and the verb *called* respectively. It can be seen that a quantifier introduced by a dependent of these words can outscope the underlined subexpressions of the logical form. In (a) there is a case of inverse scoping of the PP complement *from every city*, in (b) the *de re*-reading of an indefinite complement NP *someone from Spain* is given.

This indicates that the quantificational characteristics of a syntactic argument can only be restricted by a lexical specification as in (17) if the scope of this quantifier is also restricted. But, in this case, we are dealing with a genuine clausal semantic property.

At present it is unclear whether there are lexical entries which exploit the potential outlined in (17). The English *there*-construction is a possible candidate. It has been noted that "definite NPs" are excluded in sentences such as (19).<sup>12</sup>

(19) a. There are two/ some students in the park.

b. \* There are both/ the students in the park.

Zucchi (1995) argues that the presuppositions of "definite" NPs are not compatible with the felicity conditions of the construction. In a reply to this Keenan (2003) provides a characterization of the class of NPs which are permissible in *there*-sentences in terms of their semantic entailments. If Keenan's approach is correct, it seems that our claim from Section 2.2 that lexical heads are ignorant with respect to the quantificational nature of their dependents needs to be revised.

What is most crucial in the light of our discussion, however, is the fact that the NPs in *there*-sentences cannot have wide scope. This is reflected by the fact that (20a) can have both a *de re*- and a *de dicto*-reading, whereas (20b) can only have the *de dicto*-reading.

(20) a. Jane believes that a spy was in her office.

b. Jane believes that there was a spy in her office.

In the *de dicto*-reading the scope of the indefinite is within the EX-CONT value of the embedded verb. This, however, is exactly the domain which is made available for quantifier constraints in lexical entries.

We conclude that lexical entries of the form outlined in (17) might be needed. Our architecture of semantics embodies strong restrictions on which kinds of quantifier-sensitivity of a lexical head can be expressed: a lexical head can only restrict the quantificational aspects of one of its complements if: (i) the head also constrains the scope of these quantifiers, and (ii) this scope is narrow, i.e., within the head's EX-CONT. These restrictions seem to be empirically correct, and thus, provide further support for our proposal.

<sup>&</sup>lt;sup>12</sup>We are grateful to Olivier Bonami for pointing these cases out to us.

# 4 Conclusion

In this paper we made an attempt to establish a connection between LRS and the research on local semantic phenomena. Because of space limitations we could only present the main motivation and the technical realization of the proposal. We demonstrated that by using LRS it is possible to establish a distinction between local and combinatorial semantics, which is analogous to the distinction between syntactic category and constituent structure.

We motivated this split in semantics empirically. In particular, the formulation of linking constraints and the expression of semantic restrictions could be shown to be immune to CS properties of the selected elements. We made use of the modular organization of linguistic objects in HPSG to express this restriction in the architecture of a sign. The new architecture of local semantics has proved to be (i) more restrictive than the traditional HPSG proposal and (ii) still compatible with current analyses of LS and CS phenomena.

## References

- Androutsopoulos, Ion and Dale, Robert (2000). Selectional Restrictions in HPSG. In *Proceedings of COLING 2000*, Saarbrücken, pp. 15–20.
- Bennis, Hans and Hoekstra, Teun (1989). Generatieve Grammatica. Dordrecht: Foris.
- Bos, Johan (1996). Predicate Logic Unplugged. In P. Dekker and M. Stokhof (Eds.), *Proceedings* of the Tenth Amsterdam Colloquium, pp. 133–143. ILLC/Department of Philosophy, University of Amsterdam.
- Bouma, Gosse (2003). Verb Clusters and the Scope of Adjuncts in Dutch. In P. A. M. Seuren and G. Kempen (Eds.), *Verb Constructions in German and Dutch*. Amsterdam and Philadelphia: John Benjamins.
- Chierchia, Gennaro (1998). Reference to Kinds across Languages. *Natural Language Semantics* 6, 339–405.
- Chomsky, Noam (1965). Aspects of the Theory of Syntax. Cambridge: MIT Press.
- Cooper, Robin (1975). *Montague's Semantic Theory and Transformational Grammar*. Ph. D. thesis, University of Massachusetts, Amherst.
- Cooper, Robin (1983). Quantification and Syntactic Theory. Dordrecht: Reidel.
- Copestake, Ann, Flickinger, Dan, Pollard, Carl, and Sag, Ivan A. (2003, November). Minimal Recursion Semantics: An Introduction. Manuscript.
- Davis, Anthony and Koenig, Jean-Pierre (2000). Linking as constraints on word classes in a hierarchical lexicon. *Language* 76, 56–91.
- Dölling, Johannes (1994). Sortale Selektionsbeschränkungen und systematische Bedeutungsvariationen. In M. Schwarz (Ed.), Kognitive Semantik/ Cognitive Semantics. Ergebnisse, Probleme, Perspektiven, pp. 41–59. Tübingen: Narr.

- Egg, Markus (1998). Wh-questions in Underspecified Minimal Recursion Semantics. Journal of Semantics 15, 37–82.
- Egg, Markus and Erk, Katrin (2002). A Compositional Account of VP Ellipsis. In F. Van Eynde, L. Hellan, and D. Beermann (Eds.), *Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, pp. 162–179. Stanford: CSLI Publications.
- Gallin, Daniel (1975). Intensional and Higher-Order Modal Logic. Amsterdam: North-Holland.
- Halvorsen, Per-Kristian (1995). Situation Semantics and Semantic Interpretation in Constraint-Based Grammars. In M. Dalrymple, R. M. Kaplan, J. T. I. Maxwell, and A. Zaenen (Eds.), *Formal Issues* in Lexical-Functional Grammar, pp. 293–309. Chicago: CSLI and University of Chicago Press.
- Hinrichs, Erhard and Nakazawa, Tsuneko (1989). Flipped Out: Aux in German. In *Papers from the 25th Regional Meeting of the Chicago Linguistic Society*, Volume 25, Chicago, Illinois, pp. 193–202. CLS.
- Keenan, Edward L. (2003). The Definiteness Effect: Semantics or Pragmatics? *Natural Language Semantics* 11(2), 187–216.
- Koenig, Jean-Pierre and Davis, Anthony (2003). Semantically Transparent Linking in HPSG. In S. Müller (Ed.), *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, pp. 222–235. Stanford: CSLI Publications.
- Kordoni, Valia (2003). Morphosyntax-Semantics Interface in Lexicalist Theories. Course material for a course at ESSLLI 2003, Vienna.
- Krifka, Manfred (2003). Bare NPs: Kind-referring, Indefinites, Both, or Neither? In *Proceedings of SALT 13, 2003*. Also published in this volume.
- Kunze, Claudia and Wagner, Andreas (2001). Anwendungsperspektiven des GermaNet, eines lexikalisch-semantischen Netzes f
  ür das Deutsche. In I. Lemberg, B. Schröder, and A. Storrer (Eds.), *Chancen und Perspektiven computergest
  ützter Lexikographie*, pp. 229–246. T
  übingen: Niemeyer.
- Lang, Ewald (1994). Semantische vs. konzeptuelle Struktur: Unterscheidung und Überschneidung. In M. Schwarz (Ed.), Kognitive Semantik/ Cognitive Semantics. Ergebnisse, Probleme, Perspektiven, pp. 25–40. Tübingen: Narr.
- Pinkal, Manfred (1996). Radical Underspecification. In P. Dekker and M. Stokhof (Eds.), *Proceedings* of the Tenth Amsterdam Colloquium, pp. 587 606. ILLC/Department of Philosophy, University of Amsterdam.
- Pollard, Carl and Sag, Ivan (1987). *Information Based Syntax and Semantics. Vol.1: Fundamentals*. Stanford: CSLI Lecture Notes 13.
- Pollard, Carl and Sag, Ivan A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago and London: University of Chicago Press.
- Pollard, Carl J. and Yoo, Eun Jung (1998). A Unified Theory of Scope for Quantifiers and *wh*-Phrases. *Journal of Linguistics 34*, 415–445.

- Przepiórkowski, Adam (1998). 'A Unified Theory of Scope' Revisited. In *Proceedings of the FHCG-*98, 14–16 August 1998, Saarbrücken, pp. 185–195.
- Reyle, Uwe (1993). Dealing with Ambiguities by Underspecification: Construction, Representation and Deduction. *Journal of Semantics 10*(2), 123–179.
- Richter, Frank (2004). A Mathematical Formalism for Linguistic Theories with an Application in Head-Driven Phrase Structure Grammar. Phil. dissertation (2000), Universität Tübingen.
- Richter, Frank and Sailer, Manfred (1999). Underspecified Semantics in HPSG. In H. Bunt and R. Muskens (Eds.), *Computing Meaning*, Studies in Linguistics and Philosophy, pp. 95–112. Dordrecht: Kluwer Academic Publishers.
- Richter, Frank and Sailer, Manfred (2001). On the Left Periphery of German finite Sentences. In
  W. D. Meurers and T. Kiss (Eds.), *Constraint-Based Approaches to Germanic Syntax*, pp. 257–300. Stanford: CSLI Publications.
- Richter, Frank and Sailer, Manfred (2004a). Basic Concepts of Lexical Resource Semantics. to appear in the series of the Kurt Gödel Society, Vienna.
- Richter, Frank and Sailer, Manfred (2004b). Polish Negation and Lexical Resource Semantics. In L. S. Moss and R. T. Oehrle (Eds.), *Electronic Notes in Theoretical Computer Science*, Volume 53. Elsevier.
- Sailer, Manfred (2003). Combinatorial Semantics and Idiomatic Expressions in Head-Driven Phrase Structure Grammar. Phil. Dissertation (2000). Arbeitspapiere des SFB 340. 161, Universität Tübingen.
- Sailer, Manfred (2004). Past Tense Marking in Afrikaans. In *Proceedings of Sinn und Bedeutung 8.* to appear.
- Soehn, Jan-Philipp (2003). Von Geisterhand zu Potte gekommen. Eine HPSG-Analyse von PPs mit unikaler Komponente. Magisterarbeit, Universität Tübingen, Seminar für Sprachwissenschaft.
- Zucchi, Alessandro (1995). The Ingredients of Definiteness and the Definiteness Effect. *Natural Language Semantics 3*, 33–78.

Manfred Sailer SFB 441, Universität Tübingen mf@sfs.uni-tuebingen.de