How to Cope with Scrambling and Scope

Anette Frank Uwe Reyle
Institute for Computational Linguistics
University of Stuttgart
Azenbergstr. 12, 7000 Stuttgart 1
Tel: (0711) 121-1361 Fax: (0711) 121-1366
e-mail: uwe@adler.philosophie.uni-stuttgart.de

in: G. Görtz (Hrsg.), 1992:

KONVENS 92. 1. Konferenz "Verarbeitung natürlicher Sprache". Springer-Verlag.

Abstract

The paper presents an HPSG grammar for a fragment of German that deals with quantifier scope ambiguities triggered by scrambling and/or movement. The syntax-semantics interface we design states syntactic conditions on quantifier scoping according to the theory developed in [Frey] and constructs underspecified semantic representations (UDRSs) for scope ambiguities, for which inferences rules have been defined in [Reyle92a]. At both levels the processing of the information regarding scope ambiguity is fully incremental.

1 Introduction

Several authors[1] have put forward the idea that it is preferable to represent ambiguities in a single partial representation to which further constraints can be added monotonically to gain more information about the content of a sentence – rather than to build up a large number of alternative representations of the sentence which are then filtered by subsequent discourse and world knowledge. But as long as these partial representations are neither model-theoretically interpretable nor provided with a suitable proof theory the need to produce disambiguations remains. With it remain all the disadvantages partial interpretations were designed to eliminate. The system in [Reyle92a] breaks this deadlock. It provides a complete proof theory for structures that are partial with respect to quantifier scope.

In this paper we show how the semantics given in [Reyle92b] can be combined with an HPSG-style grammar. The basic idea of the combination being that syntax as well as semantics provide structures of equal right; that the principles internal to the syntactic and semantic level are motivated *only* by the syntactic and semantic theory, respectively; and that mutually constraining relations between syntax and semantics are governed by a separate set of principles that relate syntactic and semantic information appropriately. Thus the Semantics Principle of standard HPSG versions will be replaced by the following principle, which directly reflects the *monotonicity* underlying the interpretation process desigend in [Reyle92b]: At any stage of the derivation more details are added to the descriptions of the semantic relations between the various components of the sentence, i.e. the partial representation of any mother node is the union of the partial representations of its daughter nodes. The interface between syntactic and semantic structures is governed by principles implementing particular theories on syntactic/semantic restrictions. Here too, a

[1] See for example [Schubert/Pelletier], [Fenstad et. al.], [Nerbonne], [Alshawi], [Reyle92b].

particular perspicous way of formulating these principles is provided by the formalism of [Reyle92b]. In the present paper we will focus only on principles restricting scope ambiguities.[2] The underlying scope theory was developped originally by Frey in [Frey] for arguments of the verb and has then been extended to include adjuncts in [Frey/Tappe]. We give a brief overview of their theory in Section 2. Section 3 introduces to the formalism of [Reyle92b]. In Section 4 we introduce the Semantics Principle governing the construction of partial DRS's. We will state an HPSG grammar for a fragment of German that deals correctly with the scope principle of Frey and Tappe in Section 6. Although we adopt their theory in spirit, our analysis will not assume any traces in the parse. The grammar includes a precise statement the principles governing the interplay of scrambling and scope. An informal description of this interplay will be given in Section 5.

2 Syntactic constraints on Quantifier Scoping

Work by Frey and Tappe (see [Frey] and [Frey/Tappe]) has shown that in German the relation between the actual positions occupied by the quantificational argument phrases of the verb and their traces are instrumental in determining the possible scope relations between the arguments.[3] In (1) for example *mindestens einen Bewerber* may have wide scope over *fast jedem Mitarbeiter* because the former NP c-commands the latter; and *fast jedem Mitarbeiter* may have wide scope over *mindestens einen Bewerber* because it c-commands the trace of *mindestens einen Bewerber.*

(1) Mindestens einen Bewerber habe ich fast jedem Mitarbeiter vorgestellt.

(2) [[Mind. einen Bewerber]$_1$ habe [ich f. jed. Mitarbeiter t_1 vorgestellt]]

If on the other hand *mindestens einen Bewerber* is not moved into the "Vorfeld", then it cannot take wide scope over any of the other NP's. This is shown by the non-ambiguous sentence

(3) Ich habe fast jedem Mitarbeiter mindestens einen Bewerber vorgestellt.

Frey and Tappe assume that all the argument phrases of German verbs (including their subjects) are dominated by the verb's maximal projection, V^{max}. If the arguments have been moved from their so-called *base position* they leave traces that are coindexed with the moved arguments (compare t_1 in (2)). The movements that are relevant for the determination of scope ambiguities are, however, restricted to those occurring within – what is called – the local domain of the moved NP. This is exemplified by the non-ambiguity of examples like *Fast jeden Besucher meinte mindestens einer habe Maria gekannt*, in which the local domain of the NP *Fast jeden Besucher* is – roughly speaking – the complement structure of the matrix verb. In GB-terms the precise definition is as follows:

The *local domain* of an expression α is defined as the minimal complete functional complex, containing the licensing element of α as well as the lexically realized governor of α, where a complete functional complex is defined as the minimal maximal projection in which all Θ-roles are realized.

Given the notion of local domain we are able to state Frey's scope principle.

Syntactic Scope Principle

Suppose L_α is the local domain of an expression α. Then α *may have scope over* an expression β if either α or one of its traces c-commands β itself or one of β's traces.[4]

[2] For an implementation of binding restrictions see [Frank].

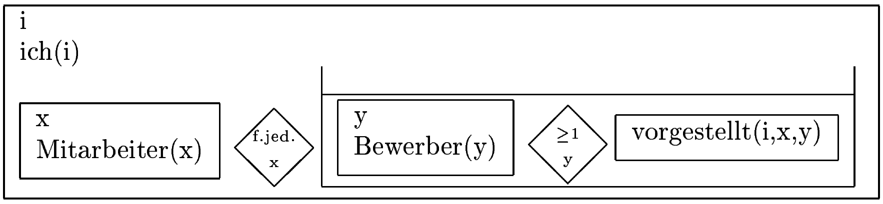
[3] By "quantificational" argument phrase we understand a real generalized quantifier. This means that indefinites are not quantificational and thus not subject to the restrictions discussed.

[4] We mentioned earlier that this principle may be applied also to adjuncts ([Frey/Tappe]). For reasons of space we cannot even touch the matter in this paper.

3 Underspecified DRS's

The easiest way to introduce U(nderspecified)D(iscourse)R(epresentation)S(tructures) is the following. Consider the DRS representation (4) of (3).

(4)



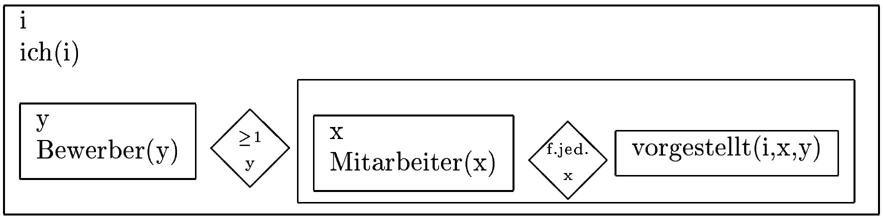
It contains two types of information. First, there is information about the hierarchical structure of the sub-DRS's, especially the information about nestedness of DRS's, or – as the term goes – about the *subordination* relation, <, between (sub-) DRS's. The second type of information relates to DRS's proper. It is of three types: universes of discourse referents, atomic conditions, and the generalized quantifier relations between (sub-) DRS's. Suppose, now, that each (sub-) DRS in (4) comes with a name l, then (4)'s information may equally well be represented by the following set of conditions.

(5)

$l_\top$:i		
$l_\top$:ich(i)	$l_1 \leq l_\top,\ l_\top \leq l_1$	(i.e. $l_1 = l_\top$)
l_1:l_{11}<f.jed,x>l_{12}	$l_{11} \leq l_1$	
l_{11}:x	$l_{12} \leq l_1$	
l_{11}:Mitarbeiter(x)	$l_2 \leq l_{12},\ l_{12} \leq l_2$	(i.e. $l_{12} = l_2$)
l_2:l_{21}<$\geq$ 1,x>l_{22}		
l_{21}:y	$l_{21} \leq l_2$	
l_{21}:Bewerber(y)	$l_{22} \leq l_2$	
l_3:vorgestellt(i,x,y)	$l_3 \leq l_{22},\ l_{22} \leq l_3$	(i.e. $l_3 = l_{22}$)

Thus, UDRS's are pairs consisting of conditions of the form l:τ together with an upper semilattice $< L, \leq >$ with one-element, $l_\top$. Note that the language of UDRS's uses *weak subordination*{/, $\leq$, instead of <. This has the advantage to be able to deal with scope disambiguation monotonically: Recall that (1) is ambiguos between (4) and (6).

(6)



What (4) and (6) have in common is given in (7), which is underspecified with respect to the scope relationship between the two quantified NP's.

$$
(7)\qquad
\begin{array}{ll}
l_\top\text{:i} & \\
l_\top\text{:ich(i)} & l_1 \leq l_\top \\
l_1\text{:}l_{11}\text{<f.jed,x>}l_{12} & l_{11} \leq l_1 \\
l_{11}\text{:x} & l_{12} \leq l_1 \\
l_{11}\text{:Mitarbeiter(x)} & \\
l_2\text{:}l_{21}\text{<}\geq\text{ 1,x>}l_{22} & l_2 \leq l_\top \\
l_{21}\text{:y} & l_{21} \leq l_2 \\
l_{21}\text{:Bewerber(y)} & l_{22} \leq l_2 \\
l_3\text{:vorgestellt(i,x,y)} & l_3 \leq l_{12},\ \ l_3 \leq l_{22}
\end{array}
$$

The construction of a meaning representation for a given sentence produces this kind of underspecification in a natural way. And it is also well suited to disambiguate such representations – once there is more information about the scope relationships of its parts. In case of (3) the syntactic scope principle tells us that the indirect object must have scope over the direct object. This information is expressed by a condition $l_2 \leq l_{12}$, which we may add to (7) in order to get the reading in (5). Thus the difference between the representations of the ambiguous sentence (1) and the non-ambiguous (3) manifests itself in the presence of the condition $l_2 \leq l_{12}$. There is no need to restructure (parts of) a semantic structure if more information about scope restriction has become available. This process of enrichment is characteristic to the construction of meaning representation: information from different sources (syntactic and semantic knowledge as well as knowledge about the world) may be incorporated into the structure by elaborating it in the sense just described.

4 The Semantics Principle

The main task of the construction of UDRS's consists in relating the labels of the information bits that are to be combined. What we mean is the following. Suppose a head complement structure. Then both of the daughters will have the description of a UDRS as value of their CONT feature. This description has the following form.

$$
\left[\text{CONT}\left[\begin{array}{l}\text{SUBORD}\{l \leq l', ...\} \\ \text{UDRS}\left[\begin{array}{l}\text{LS} = \textit{distinguished label} \\ \text{CONDS}\,\{\gamma_1, ...\}\end{array}\right]\end{array}\right]\right]
$$

The feature SUBORD contains the information about the partial order of labels and UDRS contains a set of labeled conditions which comes with a distinguished label LS. The task of combining the two CONT values is to give upper and lower bounds, with respect to $\leq$, for their distinguished labels. The identification of both of these bounds is subject to general principles. In the case of the head being verbal, e.g., the lower bound of the distinguished label l of the complement is given by the distinguished label l' of the head. This means that the description of the partial order that is attributed to the mother node will contain – beneath the subordination relations of its daughters – the condition $l' \leq l$. We will state this principle below. The form of the labelled conditions γ_i is determined by the lexical entries. Verbs will specify a relation together with its arguments.

$$
\left[\text{LOC}\left[\begin{array}{l}\text{CAT}\left[\text{SUBCAT} = < NP_{[nom,DREF=x]}, NP_{[acc,DREF=y]}, NP_{[dat,DREF=z]} >\right] \\ \text{CONT}\left[\begin{array}{l}\text{SUBORD} = \Gamma \\ \text{UDRS}\left[\begin{array}{l}\text{LS}\left[\text{L-MIN} = l\right] \\ \text{COND}\left\{\left[\begin{array}{l}\text{LABEL} = l \\ \text{REL} = \textit{vorstellen} \\ \text{ARG1} = x \\ \text{ARG2} = y \\ \text{ARG3} = z\end{array}\right]\right\}\end{array}\right]\end{array}\right]\end{array}\right]\right]
$$

The reason for the substructure LS[L-MIN $= l$] has to do with the representation of generalized quantifiers. Generalized quantifiers introduce two new labels identifying restrictor and scope as well as the type of quantification relation between the two. And for that reason there is no unique distinguished element for generalized quantifiers. The label for which the interpretation process has to identify an upper bound is the one to which the newly introduced labels are immediately subordinate, and the label for which a lower bound has to be identified is the label of the scope of the quantifier. This is the reason for the internal structure of the value of LS. We give the entry for *fast jeder*.

$$\left[\text{LOC}\left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{l}\text{HEAD} = \textit{quant}\\ \text{SUBCAT} =< [\text{LOC}\,[\text{CONT}\,[\text{UDRS}\,[\text{COND}\,\{[\text{LABEL} = l_{11}]\}\,]\,]\,]\,] >\end{array}\right]\\ \text{CONT} & \left[\begin{array}{l}\text{SUBORD} = \{l_1 \geq l_{11}, l_1 \geq l_{12}\}\\ \text{UDRS}\left[\begin{array}{ll}\text{LS} & \left[\begin{array}{l}\text{L-MAX} = l_1\\ \text{L-MIN} = l_{12}\end{array}\right]\\ \text{COND} & \left\{\left[\begin{array}{l}\text{LABEL} = l_1\\ \text{REL} = \textit{fast jeder}\\ \text{RES} = l_{11}\\ \text{SCOPE} = l_{12}\end{array}\right], \left[\begin{array}{l}\text{LABEL} = l_{11}\\ \text{DREF} = x\end{array}\right]\right\}\end{array}\right]\end{array}\right]\end{array}\right]\right]$$

As for all quantifiers a new discourse referent is introduced in the restrictor DRS, labelled l_{11}. Furthermore the feature SUBORD contains conditions saying that restrictor and scope are subordinate to the label l_1 of the entire condition. The entry for the indefinite determiner only introduces a new discourse referent. It does not distinguish between restrictor and scope.

$$\left[\text{LOC}\left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{l}\text{HEAD} = \textit{det}\\ \text{SUBCAT} =< [\text{LOC}\,[\text{CONT}\,[\text{UDRS}\,[\text{COND}\,\{[\text{LABEL} = l]\}\,]\,]\,]\,] >\end{array}\right]\\ \text{CONT} & \left[\begin{array}{l}\text{SUBORD} = \{\}\\ \text{UDRS}\left[\begin{array}{ll}\text{LS} & \left[\begin{array}{l}\text{L-MAX} = l\\ \text{L-MIN} = l\end{array}\right]\\ \text{COND} & \left\{\left[\begin{array}{l}\text{LABEL} = l\\ \text{DREF} = y\end{array}\right]\right\}\end{array}\right]\end{array}\right]\end{array}\right]\right]$$

Since we adopt a DP analysis, the SUBORD-conditions are stated in the entries of the determiners, and the entries for nouns are almost trivial.

$$\left[\text{LOC}\left[\text{CONT}\left[\begin{array}{l}\text{SUBORD} = \{\}\\ \text{UDRS}\left[\text{COND}\left\{\left[\begin{array}{l}\text{LABEL} = l\\ \text{REL} = \textit{Mitarbeiter}\end{array}\right]\right\}\right]\end{array}\right]\right]\right]$$

Having explained the roles of the features occurring within CONT we are in the position to formulate the basic components of our Semantics Principle. In this paper we will only consider verbal head complement structures. It will, however, become clear that only minor modifications are needed in order to apply the Semantics Principle to other configurations as well. We start with the case where the arguments of the verb show up in basic order (i.e. where no scope ambiguity will arise). In this case the subordination conditions stated in the lexicon or derived incrementally during the analysis will remain unchanged, i.e. the SUBORD value of the mother node is defined to contain the SUBORD value of both of its daughters.

Semantics Principle (I): CSP+BASIC (preliminary: final version in Section 6)

$$\left[\begin{array}{ll}\text{SYNSEM}\,[\text{LOC}\,[\text{CONT}\,[\text{SUBORD} = \ldots \cup \Gamma_1 \cup \Gamma_2\,]\,]\,] & \\ \text{DTRS} & \left[\begin{array}{l}\text{HEAD-DTR}\,[\text{SYNSEM}\,[\text{LOC}\,[\text{CONT}\,[\text{SUBORD} = \Gamma_1\,]\,]\,]\,]\\ \text{COMP-DTR}\,[\text{SYNSEM}\,[\text{LOC}\,[\text{CONT}\,[\text{SUBORD} = \Gamma_2\,]\,]\,]\,]\end{array}\right]\end{array}\right]$$

The next principle applies to COND and LS. The condition sets COND of the daughter UDRS's will simply be unified. But what shall we do with the distinguished labels of head and complement daughter? Recall that the lower bound relative to which L-MIN of the complement will be located in the partial order is given by the L-MIN of the head. Thus the latter gets inherited from daughter to mother in order to be available for the interpretation of possible other complements. For analoguous reasons the same holds for L-MAX.

Semantics Principle (II): Percolation of UDRS-Conditions and Upper/Lower Bounds

$$\left[\begin{array}{l} \text{SYNSEM}\left[\text{LOC}\left[\text{CONT}\left[\text{UDRS}\left[\begin{array}{l}\text{LS}\left[\begin{array}{l}\text{L-MAX} = \mathbf{l}_{max}\\ \text{L-MIN} = \mathbf{l}_{min}\end{array}\right]\\ \text{COND} = \Gamma_1 \cup \Gamma_2\end{array}\right]\right]\right]\right] \\ \text{DTRS}\left[\begin{array}{l}\text{HEAD-DTR}\left[\text{SYNSEM}\left[\text{LOC}\left[\text{CONT}\left[\text{UDRS}\left[\begin{array}{l}\text{LS}\left[\begin{array}{l}\text{L-MAX} = \mathbf{l}_{max}\\ \text{L-MIN} = \mathbf{l}_{min}\end{array}\right]\\ \text{COND} = \Gamma_1\end{array}\right]\right]\right]\right]\right]\\ \text{COMP-DTR}\left[\text{SYNSEM}\left[\text{LOC}\left[\text{CONT}\left[\text{UDRS}\left[\text{COND} = \Gamma_2\right]\right]\right]\right]\right]\end{array}\right]\end{array}\right]$$

The principle that restricts the possible target places of NP's downwards guarantees that the discourse referent x introduced by an NP that is subcategorized by some verb will bind x's occurrence in the (semantic) argument list of this verb. As was explained in Section 3, the feature L-MIN in the label structure serves this purpose.

Semantics Principle (III): Closed Formula Principle

$$\left[\begin{array}{l} \text{SYNSEM}\left[\text{LOC}\left[\text{CONT}\left[\text{SUBORD} = \ldots \cup \{l \geq l_{verb}\}\right]\right]\right]\\ \text{HEAD-DTR}\left[\text{SYNSEM}\left[\text{CONT}\left[\text{UDRS}\left[\text{LS}\left[\text{L-MIN} = l_{verb}\right]\right]\right]\right]\right]\\ \text{COMP-DTR}\left[\text{SYNSEM}\left[\text{CONT}\left[\text{UDRS}\left[\text{LS}\left[\text{L-MIN} = l\right]\right]\right]\right]\right]\end{array}\right]$$

Let us now consider the upper bound L-MAX. Recall what we said about the scope potential of indefinite NP's and genuine quantifiers. The former may take arbitrarily wide scope whereas the latter are allowed to take scope only over the elements that appear in their local domain. We will implement this restriction to the interpretation of real quantifiers by saying that the label L-MAX of the quantified NP must be subordinate to the label l_{max} which is associated with the upper bound of the local domain. The identification of L-MAX requires – as the definition in Section 2 clearly shows – a detailed discussion of the syntactic principles we are going to adopt. We will state these principles in Section 6. For the purpose of the present section it is sufficient to accept $\mathbf{l}_{max}$ as parameter which will be instantiated correctly by syntactic means. The Quantifier Scope Principle then states that for each complement that is a real quantifier the SUBORD value of the verb phrase will contain the further condition that the quantifier's maximal label in LS (L-MAX = l_{quant}) is subordinate to the label l_{max} indicating the upper limit of the local domain:

Semantics Principle (IV): Quantifier Scope Principle

$$\left[\begin{array}{l} \text{SYNSEM}\left[\text{LOC}\left[\text{CONT}\left[\text{SUBORD} = \ldots \cup \{l_{quant} \leq l_{max}\}\right]\right]\right]\\ \text{DTRS}\left[\text{C-DTR}\left[\text{SYNSEM}\left[\text{LOC}\left[\begin{array}{l}\text{CAT}\left[\text{HEAD} = \mathit{quant}\right]\\ \text{CONT}\left[\text{UDRS}\left[\text{LS}\left[\text{L-MAX} = l_{quant}\right]\right]\right]\end{array}\right]\right]\right]\right]\end{array}\right]$$

This finishes our list of general semantic principles. We now turn to the principle that governs the interaction of scrambling and scope.

5 Recognizing scope restrictions on-line

The UDRS's that can be built up according to the principles given in the last section depend on what we assume to be the value Γ of the SUBORD feature in the lexical entry of the verb. Suppose Γ is empty. Then the UDRS that is built up for a sentence like (1) would be the same as the one for (3), namely (7). I.e. all permutations of quantifiers are allowed. The realization of Frey's principle would in this case amount to monotonically add more and more restrictions to SUBORD. Although this approach is possible, we will not pursue it in this paper. (A brief discussion of this alternative is given in a footnote at the end of this section.) The other possibility is to encode in Γ the reading that Frey's principle predicts for the case all NP's occur in their base position. I.e. for *vorstellen* $\Gamma = \{ l_1 \geq l_2, l_1 \geq l_3, l_2 \geq l_3 \}$. In this case the algorithm starts with the default assumption that the NPs occur in their basic order, and then allows (parts of) this assumption to be defeated by scrambling (or movement). To pursue this approach we don't need to assume any traces in our syntactic analysis.

Frey's scope principle will be realized by comparing the actual order of verbal arguments in the sentence with their basic or 'normal' order, which can be identified in neutral intonation contexts.[5] This order is represented in the attribute BASIC, a list containing elements of type *synsem*. Due to the right-branching structure of the VP in German, the precedence relations holding among the elements of this list correspond to the c-command relations that hold among the verbal arguments if they appear in basic order. Thus in case the actual order is identical to the basic order we know that the scope relations are fixed and correspond exactly to this order. However, if at some stage of the derivation an argument a that is predicted by BASIC to appear in its base position does not occur, then we conclude that it has been moved and that it will no longer be restricted to have narrow scope with respect to the argument b that is actually processed. In this case we have to eliminate the condition(s) from SUBORD that assign a narrow scope relative to b. In general these conditions are computed by an operation of transitive closure over the precedence relation. (The formal definition of transitive closure will be given in Section 6). To see how this works in detail let us consider an example.

In (1), the value of BASIC is $< NP_{[nom]}, NP_{[dat]}, NP_{[acc]} >$. In the order of actual ocurrence the rightmost argument of the VP will be the NP marked dative (*fast jedem Mitarbeiter*). If it would also figure as rightmost element on the list of basic order, we could infer that it has narrow scope with respect to all other arguments that precede (i.e. c-command) it on the list. Since it does not, we know that the argument that is expected as rightmost element, ($NP_{[acc]}$), has been moved from its basic position (by scrambling or other means). With respect to scope relations, this tells us that the moved element will, from its landing position, c-command the NP marked dative[6] so that there will be – besides the narrow scope reading – a wide scope reading of the moved argument over the argument $NP_{[dat]}$. As a consequence, we have to modify the set of scope conditions we gained by the computation of transitive closure[7] over the list BASIC in the lexicon: the label associated with $NP_{[acc]}$ (l_3) will not be constrained any more to have narrow scope with respect to the semantic value of $NP_{[dat]}$ (l_2), so the condition $l_3 \leq l_2$ will be removed from the set of scope conditions.
Parallel to the treatment of SUBCAT, we now delete the argument that has been processed, ($NP_{[dat]}$), from the list of expected basic order, and proceed with the next element appearing in the sentence ($NP_{[nom]}$). Again we find that it does not figure as last element on the list, since the moved element still didn't show up. So it will have a wide scope reading over the nominative argument, too, and we remove the corresponding condition of narrow scope from the scope condition set. Again, the processed argument will be taken from the ordering list. Finally, the moved element ($NP_{[acc]}$) is analyzed in its landing site position. It figures as the last element on the ordering list, and no operation on the scope set is triggered. What we end up with, is the set of scoping conditions: $\{ l_2 \leq l_1 \}$, defining narrow scope of $NP_{[dat]}$ with respect to $NP_{[nom]}$.[8]

[5] See e.g. [Hoehle].

[6] We assume the principle of non-vacuous movement to hold.

[7] The reader convinces himself that it is necessary to formulate the principle w.r.t. the notion of transitive closure. The example to consider is: *Fast jedem Mitarbeiter hat mindestens einen Bewerber niemand vorgestellt.*

[8] The approach mentioned above, viz. to start with the value Γ of SUBORD being the empty set, is most easily realized, if we analyze scrambling by introducing traces for non-wh arguments. In this scenario, the NON-LOCAL attribute provides a SLASH attribute in the INHERITED feature. The order of elements on SUBCAT will be defined in the lexicon by unfying SUBCAT with the value of BASIC, which is now represented as a HEAD feature and thus gets projected by the Head

6 Syntactic Conditions on Scope

In Section 4 we introduced the Semantics Principle which defines the values of the COND attribute and the label structures of the partial DRS's as well as the projection of the SUBORD value. Since the scoping conditions are highly dependent on the underlying syntactic structure, we now have to state the syntactic conditions which constrain the subordination relations among the labels that identify the partial Sub-DRS's.

The main syntactic principles of HPSG presented in [Pollard] (Head-feature Principle, SUBCAT Principle etc.) are maintained. But we introduce more refined language-dependent Constituent Order Principles for verbal arguments in German. These operate on the order-defining attribute BASIC.

6.1 Constituent Order Principles for Verbal Arguments

Contrary to Pollard 1991, we will assume binary syntactic rules for German. Thus, the attribute COMP-DTR will only provide for an atomic object instead of representing a list of objects. The assumption of binary rules is crucial for our approach to the variation of word order and its consequences for scope relations.

Up to now, HPSG has not provided for a definition of the Constituent Order Principle that covers the range (and restrictions) of word order variation one finds in a scrambling language like German. As already mentioned, we define the basic order of verbal arguments in the lexicon. The attribute BASIC is used as controlling information structure that distinguishes between two constituent order principles for complements, COP-CH+V+BASIC, for the basic order, and COP-CH+V–BASIC, for non-basic order resulting from scrambling or topicalization of arguments. As shown in [Frank], licensing syntactic and semantic conditions on scrambling (such as definiteness, pronominal status, theme-rheme etc.) can be stated in COP-CH+V–BASIC, which constrains word order variations to wellformed sentences of German.

Due to the right-branching VP-structure, arguments that show up in the order defined in BASIC will fulfill the condition that they figure as last element on the actual value of the attribute BASIC. For ease of description, we state this condition by using a concatenation operator $\circ$ that defines $a \circ b \circ c$ as a the result of the concatenation of the lists a, b and c. If the BASIC value of the head daughter is partitioned into the (adjacent) sublists $\boxed{\text{basic}}$ and $\langle\boxed{\text{arg}}\rangle$, the phrase's value of BASIC is defined as the sublist $\boxed{\text{basic}}$, the list of remaining arguments that are expected in basic order.

Feature Principle. Note that this would mean that the order of elements on SUBCAT is no longer defined by the obliqueness hierarchy of grammatical functions. Since the traces are introduced as complement daughters, we would then state a single constituent order principle operating on SUBCAT, that enforces the actual complement being processed as complement daughter to appear as last element on the SUBCAT list. This restriction ensures that the argument(s) we find not being realized in basic order will be analyzed as traces and represented in the SLASH attribute. Accordingly we would state a filler rule that decharges SLASH if it can be structure shared with the TO-BIND value introduced by the value of a filler daughter. Our Complement Scope Principle would then (informally) state that for each argument a that is actually processed as a complement or filler daughter (i) it has to be identified with one of the arguments on the BASIC list of the HEAD and (ii) for each argument b on BASIC that follows it, and which is not contained in the actual SLASH value, the condition $l_a \geq l_b$ is added to the SUBORD value. As scrambling of multiple arguments over the subject as in (i) does not induce narrow scope of the object NP with respect to the indirect object, the introduction of scope restrictions by the Complement Scope Principle has to obey the general condition that SUBCAT may not be empty.

(i) *Fast jedem Mitarbeiter hat mindestens einen Bewerber niemand vorgestellt.*

We decided not to choose this alternative, in order to clearly separate pure subcategorization information, encoded in SUBCAT, from precedence conditions, subject to the Constituent Order Principles. In the grammar we sketch in the following section, the Constituent Order Principles operate on the order-defining attribute BASIC, while the SUBCAT-Principle is solely concerned with the processing of the subcategorization information proper.

COP-CH+V+BASIC: COP for Basic Order of Complements

COP-CH+V &

$$\begin{bmatrix} \text{SYNSEM} \left[\text{LOC} \left[\text{CAT} \left[\text{BASIC} = \boxed{\text{basic}} \right]\right]\right] \\ \text{DTRS} \begin{bmatrix} \text{HEAD-DTR} \left[\text{SYNSEM} \left[\text{LOC} \left[\text{CAT} \left[\text{BASIC} = \boxed{\text{basic}} \circ <\boxed{\text{arg}}> \right]\right]\right]\right] \\ \text{COMP-DTR} \left[\text{SYNSEM} \ \boxed{\text{arg}} \right] \end{bmatrix} \end{bmatrix}$$

The COP for non-basic order of arguments partitions the head daughter's value of BASIC in a way such that $<\boxed{\text{arg}}>$ precedes a non-empty list (nelist) of arguments ($\boxed{\text{non-basic}}$). As before, $\boxed{\text{arg}}$ is removed from BASIC to give the phrase's value of BASIC, which is defined as the concatenation $\boxed{\text{basic}} \circ \boxed{\text{non-basic}}$.

COP-CH+V-BASIC: COP for Non-Basic Order of Complements

COP-CH+V &

$$\begin{bmatrix} \text{SYNSEM} \left[\text{LOC} \left[\text{CAT} \left[\text{BASIC} = \boxed{\text{basic}} \circ \boxed{\text{non-basic}} \right]\right]\right] \\ \text{DTRS} \begin{bmatrix} \text{HEAD-DTR} \left[\text{SYNSEM} \left[\text{LOC} \left[\text{CAT} \left[\text{BASIC} = \boxed{\text{basic}} \circ <\boxed{\text{arg}}> \circ \boxed{\text{non-basic}}_{\text{nelist}} \right]\right]\right]\right] \\ \text{COMP-DTR} \left[\text{SYNSEM} \ \boxed{\text{arg}} \right] \end{bmatrix} \end{bmatrix}$$

6.2 Syntactic Conditions on Scope

Syntactic Constraints on the Complement Scope Principle

The differentiation between COPs for basic and non-basic order of arguments provides a means for defining the modification of scope relations among arguments that result from the change of syntactic c-command relations when scrambling or topicalization occurs.
In Section 4, we already mentioned that for arguments showing up in basic order, the SUBORD value of the mother is a superset of the SUBORD values of its daughters.

Semantics Principle (I): CSP+BASIC

$$\begin{bmatrix} \text{SYNSEM} \left[\text{LOC} \left[\text{CONT} \left[\text{SUBORD} = \ldots \cup \Gamma_1 \cup \Gamma_2 \right]\right]\right] \\ \text{DTRS} \begin{bmatrix} \text{HEAD-DTR} \left[\text{SYNSEM} \left[\text{LOC} \left[\text{CONT} \left[\text{SUBORD} = \Gamma_1 \right]\right]\right]\right] \\ \text{COMP-DTR} \left[\text{SYNSEM} \left[\text{LOC} \left[\text{CONT} \left[\text{SUBORD} = \Gamma_2 \right]\right]\right]\right] \end{bmatrix} \end{bmatrix}$$

For non-basic argument order, however, we concluded that the arguments being moved (represented in the sublist $\boxed{\text{non-basic}}$) are no longer constrained to have narrow scope with respect to the argument $\boxed{\text{arg}}$ currently processed. We therefore derive the transitive closure over the concatenated list $<\boxed{\text{arg}}> \circ \boxed{\text{non-basic}}$, which computes the set of subordination relations which no longer can be assumed to hold.[9] This set is deleted from the SUBORD set of the HEAD-DTR (Γ_1). The result is defined to be contained in the SUBORD value of the phrase's CONT.

---

[9] Transitive closure is defined as a function that takes a list of synsem objects and computes the set of conditions defining the subordination relations holding among the arguments:

$$\text{trans-closure}(<>) \rightarrow \{ \}$$

$$\text{trans-closure}(<_{synsem-object}>) \rightarrow \{ \}$$

$$\text{trans-closure}(< \left[\text{LOC} \left[\text{CONT} \left[\text{UDRS} \left[\text{LS} \left[\text{L-MAX} = \mathbf{l}_1 \right]\right]\right]\right]\right], \left[\text{LOC} \left[\text{CONT} \left[\text{UDRS} \left[\text{LS} \left[\text{L-MAX} = \mathbf{l}_2 \right]\right]\right]\right]\right] >) \rightarrow \{ \mathbf{l}_1 \leq \mathbf{l}_2 \}$$

$$\text{trans-closure}(< \left[\text{LOC} \left[\text{CONT} \left[\text{UDRS} \left[\text{LS} \left[\text{L-MAX} = \mathbf{l}_1 \right]\right]\right]\right]\right], \left[\text{LOC} \left[\text{CONT} \left[\text{UDRS} \left[\text{LS} \left[\text{L-MAX} = \mathbf{l}_2 \right]\right]\right]\right]\right], \left[\text{LOC} \left[\text{CONT} \left[\text{UDRS} \left[\text{LS} \left[\text{L-MAX} = \mathbf{l}_3 \right]\right]\right]\right]\right] >) \rightarrow \{ \mathbf{l}_1 \leq \mathbf{l}_2, \mathbf{l}_1 \leq \mathbf{l}_3, \mathbf{l}_2 \leq \mathbf{l}_3 \}$$

Semantics Principle (I): CSP–BASIC: COP-CH+V–BASIC &

$$\left[\begin{array}{l}\text{SYNSEM}\left[\text{LOC}\left[\text{CONT}\left[\text{SUBORD} = \ldots \cup \Gamma_2 \cup (\Gamma_1 \setminus \text{trans-closure}(<\boxed{\text{arg}}> \circ \boxed{\text{non-basic}}))\right]\right]\right] \\ \text{DTRS}\left[\begin{array}{l}\text{HEAD-DTR}\left[\text{SYNSEM}\left[\text{LOC}\left[\begin{array}{l}\text{CAT}\left[\text{BASIC} = \boxed{\text{basic}} \circ <\boxed{\text{arg}}> \circ \boxed{\text{non-basic}}\right] \\ \text{CONT}\left[\text{SUBORD} = \Gamma_1\right]\end{array}\right]\right]\right] \\ \text{COMP-DTR}\left[\text{SYNSEM}\left[\text{LOC}\left[\text{CONT}\left[\text{SUBORD} = \Gamma_2\right]\right]\right]\right]\end{array}\right]\end{array}\right]$$

Definition of Local Domain for the Quantifier Scope Principle

In Section 2, the local domain of an expression α has been defined as the minimal complete functional complex, containing the licensing element of α as well as the lexically realized governor of α, where a complete functional complex was defined as the minimal maximal projection in which all Θ-roles are realized. In our HPSG grammar this definition of the local domain for verbal arguments corresponds to the phrasal verb projection where all arguments have been realized, i.e. SUBCAT is saturated. We therefore instantiate the label l_{max} as the value of the feature L-MAX in the CONT attribute of the verb phrase, which the Quantifier Scope Principle stated in Section 4 defines as the upper local domain for quantified phrases.

$$\left[\text{LOC}\left[\begin{array}{l}\text{CAT}\left[\begin{array}{l}\text{HEAD} = verb \\ \text{LEX} = - \\ \text{SUBCAT} = <>\end{array}\right] \\ \text{CONT}\left[\text{UDRS}\left[\text{LS}\left[\text{L-MAX} = \mathbf{l}_{max}\right]\right]\right]\end{array}\right]\right]$$

7 Conclusion

We presented an HPSG grammar for German that defines a syntax-semantics interface for the construction of underspecified discourse representation structures. The properties of the principles governing the syntactic and semantic representations are stated individually in the form of general principles. The separation of syntactic and semantic principles enables us to clearly identify the interaction between the modules, i.e. the 'interface' between syntax and semantics. In the fragment we discussed, this interaction was formulated for the scoping properties of quantifiers, where syntactic constraints of word order restrict the set of possible readings.

We end by giving an informal description of all the Semantics Principles we have been discussing.

Semantics Principle

- **Percolation of Upper and Lower Bounds**
  In a headed structure, the value of the head-daughter's LS, the distinguished label of a UDRS, is token-identical to the value of the phrase's LS attribute.
- **Inheritance of UDRS Conditions**
  In a headed structure, the phrase's value of the set COND, representing atomic UDRS conditions, is the union of the COND values of the daughters.
- **Scoping Principles**
  In a head-complement structure, the phrase's value of SUBORD is defined as the union of the sets of conditions defined by the following clauses:
  1. **Complement Scope Principle**
     – If COP-CH+BASIC applies, the phrase's SUBORD value will contain the union of the SUBORD values of the daughters as a subset.

- If COP-CH-BASIC applies, the phrase's SUBORD value will contain the SUBORD value of the complement daughter and the subset that results from subtraction of scoping conditions from the SUBORD value of the head daughter (defined in CSP-BASIC).

2. **Quantifier Scope Principle**
If the head feature of the complement daughter is of type quant, the phrase's SUBORD value will contain the condition $l_{quant} \leq l_{max}$ where l_{quant} is the value of L-MAX of the quantified noun phrase, and l_{max} is the label L-MAX of the verb phrase's upper local domain.
3. **Closed Formula Principle**
The phrase's value of SUBORD will contain the condition $l_{verb} \leq l$ where l_{verb} is the value of L-MIN of the verbal projection and l is the value of L-MIN in the CONT attribute of the nominal complement.
4. No further conditions will be contained in the phrase's SUBORD value.

Literatur

[Alshawi] Alshawi, Hiyan (1990): "Resolving Quasi Logical Forms", in *Computational Linguistics*, Vol. 16, No. 3.

[Fenstad et. al.] Fenstad, J.E. / Halvorsen, P.-K. / van Benthem, J. (1987) *Situations, Language and Logic*, Reidel Publishing Company, 1987.

[Frank] Frank, Anette (1992): "A New HPSG-Analysis of Scrambling and Binding in German." ms. Stuttgart.

[Frey] Frey, Werner (1990): *Syntaktische Bedingungen für die Interpretation*, AIMS No. 01-90, Stuttgart.

[Frey/Tappe] Frey, Werner / Tappe Thilo (1992): *Grundlagen eines GB-Fragments für das Deutsche*, appears as *Arbeitspapier des Sonderforschungsbereichs 340*, Stuttgart.

[Hoehle] Hoehle, Tilman (1982): "Explikation fuer 'normale' Betonung und 'normale Wortstellung'. " in: W. Abraham (ed.): *Satzglieder im Deutschen. Vorschlaege zur syntaktischen, semantischen und pragmatischen Fundierung.* Tuebingen, Narr, p.75-153.

[Kamp/Reyle 91] Kamp, H. / Reyle, U. (1991): *From Discourse to Logic, Vol I*, to appear: Kluwer, Dordrecht (1990).

[Nerbonne] Nerbonne, J. (1992), "A Feature Base Syntax/Semantics Interface", ms. Saarbruecken.

[Pollard/Sag] Pollard, Carl / Sag, Ivan A (1987): *Information-Based Syntax and Semantics*, CSLI Lecture Notes Series 13.

[Pollard] Pollard, Carl (1991): *Topics in Constraint-Based Syntactic Theory.* ms.

[Reyle92a] Reyle, Uwe (1992): "Dealing with Ambiguities by Underspecification: A First Order Calculus for Unscoped Representations", in: *Proceedings of the Eighth Amsterdam Colloquium*, Amsterdam, 1992.

[Reyle92b] Reyle, Uwe (1992): "Dealing with Ambiguities by Underspecification: Construction, Representation and Deduction", to appear in: *Jounal of Semantics.*

[Schubert/Pelletier] Schubert, L.K. / Pelletier, F.J. (1982), "From English to Logic: Context-Free Computation of Conventional Logic Translations", in: *Journal of the Association for Computational Linguistics.*