Simple Trees with Complex Semantics: On Epistemic Modals and Strong Quantifiers

Frank Richter[1] and Manfred Sailer[2]

[1] Universität Tübingen , fr@sfs.uni-tuebingen.de
[2] Universität Göttingen, manfred.sailer@phil.uni-goettingen.de

Abstract. Lechner (2006) presents an analysis of the scopal interaction of English epistemic modal verbs with negated strong quantifiers in the framework of Transparent Logical Form, which constructs semantic representations as extensions of syntax trees. Noting conceptual and empirical shortcomings of this approach, we propose an analysis in Lexical Resource Semantics, a constraint-based semantics framework embedded in Head-driven Phrase Structure Grammar. Our analysis treats the original English data as well as semantically parallel constructions in German, although the syntax of the German constructions is substantively different from English. The LRS framework views syntax and semantics as independent modules connected by a complex interface employing modern underspecification techniques. It permits a surface-oriented syntactic analysis without a multitude of additional empty categories, which has clear advantages for the application of efficient parsing algorithms. At the same time, the LRS analysis elegantly captures crosslinguistic generalizations about the semantic phenomenon under investigation.

1 Introduction

In this paper we present empirical and conceptual arguments in favor of a syntax-semantics interface that links a surface-oriented syntax to a semantic representation in a well-established logical language, here Ty2 (Gallin 1975). This interface uses techniques of underspecified semantics to overcome mismatches between the syntactic structure and the semantic representation. We look at an especially intriguing case of such a mismatch, the scopal interaction of an epistemic modal verb with a negated strong quantifier.

We contrast our analysis with the one presented in (Lechner 2006) within the framework of Transparent Logical Form (Stechow 1993, Heim and Kratzer 1998). Transparent Logical Form is designed to derive scope ambiguities from differences in the syntactic structure: A sentence with two different scope readings must have two different syntactic representations. We show that such an architecture leads to undesirable consequences for both the syntactic and the semantic analysis. A problem for syntax occurs if two languages differ with respect to their (surface) syntactic properties but allow the same scopal readings in a particular constellation. In this situation, Transparent Logical Form is forced to treat the syntactic properties of the two languages more alike than they are. On the other hand, in some cases a restriction on the scoping possibilities depends on aspects of the semantic representation, but not on aspects of the syntax. If this is the case, Transparent Logical Form must turn semantic properties into syntactic features. The example that we are using in this paper will allow us to illustrate both types of problems.

Our own approach stands in the tradition of surface-oriented syntax. In such frameworks, the same syntactic structure can be associated with different scopal readings. There are several techniques to achieve this goal. One of the best-known is Cooper storage (Cooper 1975). Alternatively, type-shifting can be assumed (Hendriks 1993) to construct a semantic combinatorics which uses a categorial grammar set-up in parallel to a constituent-based syntactic parse. While these approaches assume a single syntactic structure, they use distinct semantic derivations for ambiguous sentences, and consequently, distinct semantic representations at the nodes in the syntactic tree. From a computational perspective, this is still a severe shortcoming since it multiplies the number of outputs for each sentence just as two distinct syntactic analyses do.

To overcome this problem systems of underspecified semantics were developed (Pinkal 1996). These systems combine a surface-oriented syntactic analysis and a reading-independent semantic construction, while the combinatorics still restricts the possible readings of a sentence in the appropriate way. Within Head-driven Phrase Structure Grammar (HPSG), the syntactic framework assumed in this paper, a number of variants of such approaches have been proposed, including (Egg 2004), (Richter and Sailer 2004), and (Copestake, Flickinger, Pollard and Sag 2005).

In the present paper we focus on *Lexical Resource Semantics* (LRS, (Richter and Sailer 2004)), demonstrating the way in which syntax and semantics are seen as modules which impose their own restrictions and which are linked to each other by constraints. As a consequence, LRS is immune to the two problems mentioned for Transparent Logical Form: First, even if two languages differ in their syntactic structure, we can derive identical readings. Second, we can express constraints that refer to semantic properties of constituents without having to encode them as syntactic properties.

2 Split Readings in LF Syntax

The sentences in (1) have a reading in which the epistemic modal intervenes between the negation and the universal quantifier. This reading constitutes a genuine challenge for semantic composition as the semantic contribution of the subject NP seems to be split by that of the modal. We will refer to this type of reading as the *split reading*.

(1) a. Not every boy can make the basketball team. $\neg \succ$ **CAN** $\succ \forall$
"It is not possible that every boy makes the basketball team" (Lechner 2006)

b. Nicht jeder kann gewinnen. $\neg \succ$ **CAN** $\succ \forall$
not everyone can win "It is not possible that everyone wins."

Lechner (2006) cites this reading as evidence of head movement at a syntactic level of LF. To derive the reading in the context of a theory of raising, he has to make four assumptions:

- L(i) strong NPs do not reconstruct under raising verbs,
- L(ii) NPs of the form *not NP* contain a semantically vacuous *not* but require to be contained in a NegP which contributes the negation,
- L(iii) the NegP is high in the tree but has a variable position,
- L(iv) an epistemic modal can move over the subject in the syntax.

This results in the structure in (2):

(2) The structure of (1-a) according to Lechner (2006):

$[_{NegP}$ NEG $[_{Neg^0}$ can_i $[_{TP}$ *not every boy*$_j$ $[t_i\ [_{VP}\ t_j$ make the team $]]]]]]$
$\lambda p.\neg p$ λp**CAN**(p) $\lambda P.\forall x($**boy**$(x) \rightarrow P(x)$ $\lambda x.$**make-team**(x)

Crucially, the assumptions L(i)–L(iv) do not follow from any other aspect of the grammar. They capture the empirical observations with the mechanisms of Lechner's framework. Henceforth we will refer to analyses couched in this type of framework as LF-Syntax (LFS) analyses. Their hallmark is the fact that they state semantic generalizations as conditions on appropriately extended syntax trees, or on their derivations. The advantage of the LFS analysis at hand is that the semantics can be read off the (LF-) syntactic structure directly. The price to pay is a highly complex syntactic derivation with many empty and functional categories for which there is no purely syntactic evidence and which make the approach unattractive for standard theories of computational syntax and semantics.

Newmeyer (2008) argues against this type of approach from a methodological perspective, showing that the inclusion of semantic distinctions in syntactic analyses makes it almost impossible to derive the basic syntactic generalizations about English modals: English modal verbs differ syntactically from non-modal verbs, but their syntactic behavior is independent of their epistemic or deontic interpretation, and of their scope.

Further problems for this approach arise when we take a different perspective and consider crosslinguistic data. Linguists generally agree that the syntactic constraints which govern the placement of finite verbs in German and English are fundamentally different. German does not exhibit syntactic distinctions between modal or auxiliary verbs and other finite verb forms. Instead, fronting of the finite verb to second position behind a sentence-initial constituent uniformly applies to all finite verb forms in V2 root clauses. However, the possible readings of the modals in (1-a) and (1-b) are the same. It seems impossible that this could be due to a syntactic parallelism related to L(iv), given that the syntax of the finite (modal) verb actually differs in the two languages. We take the clear semantic parallelism as further evidence that the split reading in (1) is not due to the availability of syntactic head movement.

A further argument against a syntactic theory of split readings with modal verbs comes from the behavior of strong quantifiers in argument positions of other verbs. L(i) describes the fact that universal quantifiers as subjects of raising verbs do not have de dicto readings in English (3-a). (3-b) is evidence that other quantifiers are not subject to this restriction. However, the fact that L(i) is a syntactic condition on raised NPs misses a generalization: Zimmermann (1993) shows that strong quantifiers also lack de dicto readings in the direct argument position of opaque verbs such as *seek*, as seen in (3-c). This suggests that an adequate formulation of L(i) should not be restricted to raised NPs and should not be syntactic in nature.

(3) a. Every student seems to have passed the test. (only de re)
b. A student seems to have passed the test. (de re/de dicto)
c. John seeks every unicorn. (only de re)
$\forall x(\mathbf{unicorn}(x) \rightarrow \mathbf{seek}(w', \mathbf{john}, \lambda w'' \lambda P.P(w'', x)))$
$\mathbf{seek}(w', \mathbf{john}, \lambda w'' \lambda P.\forall x(\mathbf{unicorn}(x) \rightarrow P(w'', x)))$

Although epistemic modals are raising verbs, they are exceptional in that they may take scope over strong quantifiers in subject position. To account for this apparent violation of constraint L(i), Lechner (2006) introduces the head-movement option L(iv) for epistemic modals to derive (1-a). But then the syntactic generalizations only capture part of a semantic phenomenon that, in our view, should receive a uniform account which has nothing to do with the syntactic position of the verb: If a strong quantifier occurs in a (surface) argument position of an opaque non-modal verb, it must take scope over the verb.

3 An LRS Analysis

We propose a syntax-semantics interface built on the syntactic structure (4-a) for all readings of sentence (1-a). The syntactic structure of the corresponding German sentence, which has the same readings, is shown in (4-b).

(4) a. [$_S$ [$_{NP}$ Not every boy] [$_{VP}$ can [$_{VP}$ make the team]]]
b. [$_{\overline{S}}$ [$_{NP}$ Nicht jeder]$_j$ [$_S$ kann$_i$ [$_{VP}$ t$_j$ gewinnen t$_i$]]]

A simpler syntax without functional categories such as (4-a) and, *mutatis mutandis*, the German structure (4-b), requires a more elaborate syntax-semantics interface, as also advocated by Culicover and Jackendoff (2005). Our surface-oriented syntax is provided by HPSG, while the interface to semantics is couched in *Lexical Resource Semantics* (LRS, (Richter and Sailer 2004)). There is a clear division of work: The syntactic structure reflects syntactic generalizations. Constraints on readings express the semantic generalizations at the interface. Using techniques of underspecified semantics, the interface conditions are formulated in terms of scope specifications in lexical entries and general scope principles. There is no empty abstract negation NEG, and the words *not* and *nicht* themselves contribute semantic negation, in contrast to the assumptions L(ii) and L(iii) of the LFS analysis above.

The core idea of the LRS interface is that the semantic representations of sentences result from accumulating the meaning contributions and semantic constraints associated with lexical entries in accordance with general phrase-level semantic principles. Here we will associate complex constraints on semantic representations with lexical and phrasal signs. These constraints are not themselves semantic representations. They state which fragments of semantic representations occur in a sign and what is known about their mutual relationships. This means that these constraints *denote* semantic representations. An important principle about the interpretation of utterances ultimately requires that the semantic representations which are associated with an utterance (a) use all and only those logical symbols which are introduced in the LRS constraints of the signs contained in the utterance, and

(b) respect all the restrictions on the mutual relationship of their subterms which are either lexically introduced or imposed by phrase-level semantic principles. Our semantic representations are expressions of a higher-order logic (two-sorted type theory, Ty2), but the representations of utterances are not derived by the lambda calculus like in most systems with representations from a type-theoretic higher-order language. Instead, the logical representations of syntactic daughters are combined by unification.[3]

To derive the truth-conditions of utterances in a principled fashion, LRS distinguishes different aspects of the semantic representation(s) associated with each sign. Three aspects of the semantics of phrases and words will be important in our discussion: (a) the main content of a lexical item (written as an underlined formula, $\underline{\phi}$), (b) the internal content of a lexical item (put between curly braces, $\{\psi\}$) and (c) its external content (prefixed with '^', $^\wedge\chi$) to mark scope boundaries within a head projection. Their significance is as follows: All operators that combine with a lexical sign along its head projection take scope over the internal content of the lexical head. The external content marks the term which includes the operator with the widest scope at the highest syntactic projection of a sign. The main content represents the core semantic contribution of a sign, and may also be viewed as its lexical meaning. The main content as well as the external and internal content are inherited along head projections and are therefore locally available to constraints at each phrasal projection. This is crucial for the principles of the combinatoric system, since they essentially operate at the phrase level and establish relationships between the main, internal and external content of syntactic daughters. In doing so, they enforce mutual restrictions (such as possible subterm relationships or necessary identities between subterms) in accordance with the requirements of the syntax-semantics interface. We will introduce all relevant principles as our discussion proceeds. Since the subterm-relation is a particularly important type of constraint in many LRS principles, it receives a special notation in CLLRS. Subterm conditions often come together with *meta-variables*, notated as upper-case letters. Meta-variables stand for nodes in the term tree that are still underspecified. $A : [\phi_1, \ldots, \phi_n]$ means that A comprises the subexpressions ϕ_1 to ϕ_n.

3.1 Epistemic Modals in Split Readings in English

We now turn to the analysis of the English sentence (1-a). The LRS constraint associated with the infinitival VP *make the team* is shown in (5). The main content of the VP, which we underline, is the constant **make-team**. The internal content of the VP (in curly braces) consists of the main content, applied to all its arguments, i.e. it is the expression **make-team**(w', x). We also mark the external content, which the VP does not specify any further. Note that the internal content is a subexpression of the external content. This is not an accidental fact of this expression, but it is required by an independent LRS principle.

(5) LRS constraint of the VP *make the team*: $^\wedge A : [\{\underline{\textbf{make-team}}(w', x)\}]$

LRS inherits the simple lexical linking mechanism from traditional approaches to semantics in HPSG such as (Pollard and Sag 1994). A verb has access to the SYNSEM value of its syntactic arguments. Under an attribute CONTENT *synsem* structures contain lexical aspects of the semantic structure, the referential variable associated with the syntactic argument and its main content. In contrast, the external content, the internal content and other parts of the content of the selected element are not visible to the selector. The availability of the referential variable of the subject, here x, allows us to link it directly to the right argument slot of the functor **make-team**. Local access to the main content is also needed to impose selectional restrictions (Soehn 2006).

Modals contribute quantification over possible worlds. The restrictor specifies the type of modality in terms of an appropriate accessibility relation between possible worlds, **acc**:

[3] In feature logical specifications of LRS, it is not unification but equality of substructures of logical representations that is employed in the grammar principles. Below we will use the computational specification language of LRS, CLLRS (Penn and Richter 2005), since its linear notation is more readable and more compact than feature logic descriptions. Computations with CLLRS do employ unification, and we use the term in this computational sense as a convenient metaphor for explaining the effect of LRS constraints on constructing semantic representations.

(6) LRS constraint of the verb *can*: $^\wedge\lambda w.A : [\exists w'(\underline{\mathbf{acc}}(w, w') \wedge B : [w', \{C\}])]$

In the present analysis, we do not use an index (eventuality) variable for verbal projections as in previous LRS analyses. The linking between the modal verb and its VP complement is instead encoded by means of the VP's main content. We require that the main content of the complement VP be a subexpression of the modal's internal content C.

(7) Linking constraint in the lexical entry of *can*:
The main content of the VP complement is a subexpression of *can*'s internal content.

Our analysis of modal verbs is taken from (Richter and Sailer 2004, Sailer 2006). The key idea is that modal verbs identify their internal content with that of their complement VP. This permits both wide scope and narrow scope of the modal with respect to the subject and to complements of the infinitival VP. Consequently, our analysis of modals achieves the same effect as the optional syntactic head-movement rule of modals in the LFS analysis (L(iv)) with respect to subjects. In contrast to L(iv), however, it also accounts for the scoping possibilities of complements.

The INTERNAL CONTENT RAISING PRINCIPLE enforces the identity of the internal contents of a modal and its VP complement. It is independently motivated by the analysis of opaque verbs (see (14) below) and of neg raising constructions (Sailer 2006). Here we make a few minor changes to its exact formulation compared to previous versions since it was previously formulated in the context of a less elaborate lexical analysis of modals and with reference to an eventuality index of verbs. Our revised formulation is given in (8).

(8) INTERNAL CONTENT RAISING PRINCIPLE (ICRP):
In a head-complement structure,
if the main content of the head is not a subexpression of its internal content,
and the index or the main content of the complement is a subexpression of the head's internal content,
then the internal content of the head and internal content of the complement are identical.

The ICRP leads to unification of the meta-variable C with the internal content of the VP, which is the expression $\mathbf{make\text{-}team}(w', x)$ according to (5). This results in the new LRS constraint shown in (9):

(9) LRS constraint of the VP *can make the team*:
$^\wedge\lambda w.A : [\exists w'(\underline{\mathbf{acc}}(w, w') \wedge B : [w', \{\mathbf{make\text{-}team}(w', x)\}])]$

Next we consider the LRS constraint of the NP *not every boy*, (10). The universal quantifier marks the external content of the NP. The negation must take scope over the external content, but there may be semantic material intervening between the negation and the quantifier, as the meta-variable D indicates.

(10) LRS constraint of the NP *not every boy*: $\neg D : [^\wedge\forall x(\{\underline{\mathbf{boy}}(x)\} \rightarrow E : [x])]$

The combination of a verbal projection with a quantified argument is subject to a sub-clause of the LRS SEMANTICS PRINCIPLE. Since we only need this one sub-clause, we introduce it here as a separate principle. It requires that the argument take scope over the internal content of the VP. For ease of reference we will call it the QUANTIFIER-HEAD PRINCIPLE. It ensures that all quantified dependents of a head take scope over the head's internal content. An analogon of this principle belongs to most systems that work with semantic underspecification, including (Egg 2004) and MRS (Copestake et al. 2005).

(11) QUANTIFIER-HEAD PRINCIPLE (QHP):
In a head complement structure, if the nonhead is a quantifier, then the head's internal content is a subexpression of the nonhead's scope.

In (10) the scope of the quantified subject is marked by the meta-variable E. When the finite VP combines with the subject, two subexpression constraints apply: First, The QHP requires that the internal content of the VP, the expression $\textbf{make-team}(w', x)$, be a subexpression of E. Second, the principles guiding the accumulation of semantic constraint information ensure that all meaning contributions of the subject be part of the external content of the utterance. In our example, the external content of the utterance is identical to the external content of the clause.[4] This leads to the LRS constraint in (12). Note that the meta-variable A is constrained to contain the LRS constraint of the finite VP from (9) ($\exists w'(\ldots)$) and the LRS constraint of the subject from (10) ($\neg D : [\ldots]$)):

(12) LRS constraint of sentence (1-a):

$$^\wedge\lambda w.A : [\exists w'(\underline{\textbf{acc}}(w, w') \wedge B : [w', \{\textbf{make-team}(w', x)\}]), \neg D : [\forall x(\textbf{boy}(x) \rightarrow E : [\textbf{make-team}(w', x)])]]$$

In Figure 1 we summarize the LRS analysis of sentence (1-a). Into each node in the tree we write the LRS constraint on the semantic representation of the relevant node. At branching nodes we indicate the effect of the ICRP and the QHP.

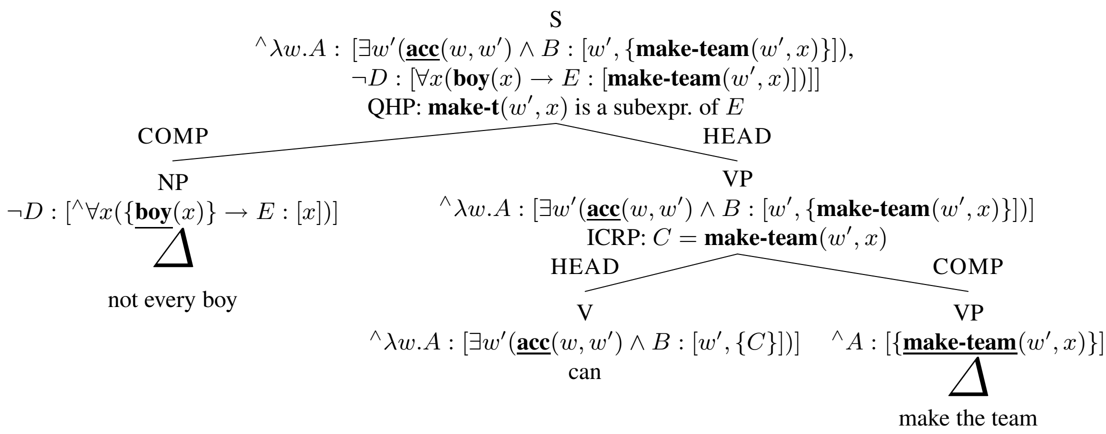

Fig. 1. Sketch of the analysis of sentence (1-a)

The LRS constraint (12) is resolved by assigning the meta-variables subexpressions from (12) which respect the given scoping relations. There are three possible resolutions for (12). In (13-a) we state the resolution that yields the split reading. The resulting expression is shown in (13-b).[5]

(13) a. $A = \neg D$, $B = \forall x(\textbf{boy}(x) \rightarrow E)$,
$D = \exists w'(\textbf{acc}(w, w') \wedge B)$, $E = \textbf{make-team}(w', x)$

b. Semantic representation resulting from the meta-variable assignment in (13-a):

$$\lambda w.\neg\exists w'(\textbf{acc}(w, w') \wedge \forall x(\textbf{boy}(x) \rightarrow \textbf{make-team}(w', x)))$$

With the analysis of the split reading in English we have achieved the main goal of revising the LFS analysis. The remaining question now concerns the predictions of the two analyses with respect to related syntactic constructions such as raising, and related semantic phenomena such as the behavior of strong quantifiers as arguments of opaque verbs.

[4] See Section 4 for examples with more than one clause.

[5] The other possible resolutions provide two more readings. These are in fact available but will be ignored since they are not immediately relevant for our discussion of the two alternative syntax-semantics interfaces.

(i) a. $\lambda w.\neg\forall x(\textbf{boy}(x) \rightarrow \exists w'(\textbf{acc}(w, w') \wedge \textbf{make-team}(w', x)))$
(It is not the case that every boy can make the team.)

b. $\lambda w.\exists w'(\textbf{acc}(w, w') \wedge \neg\forall x(\textbf{boy}(x) \rightarrow \textbf{make-team}(w', x)))$
(It is possible that not every boy makes the team.)

To address these issues, we have to take another look at the LFS rule L(i). L(i) states that strong quantifiers do not reconstruct below raising verbs. In our critique of the LFS approach we suggested that the empirical scope of this rule should be extended to take into account the more comprehensive observation that a strong quantifier in an argument position of a non-modal verb must take scope over that verb. In other words, in this syntactic constellation we do not find de dicto readings of strong quantifiers.

To see how LRS captures this generalization, we need to start with the relevant parts of the lexical entries of raising verbs and opaque verbs. The lexical specification of the verbs *seem* and *seek* according to the analysis of opaque verbs in (Richter and Sailer 2004) looks as in (14).

(14) a. LRS constraint of the verb *seem*: $^\wedge\lambda w.A : [\underline{\textbf{seem}}(w, \lambda w'.B[w', \{C\}])]$
b. LRS constraint of the verb *seek*: $^\wedge\lambda w.A : [\underline{\textbf{seek}}(w, x, \lambda w'.B[w', \{C\}])]$

The full specifications of the lexical entry of *seem* and *seek* include the appropriate argument linking. In the case of *seek* the index of the direct object NP must occur within the internal content of the verb. For *seem*, the main content of the VP complement must be a subexpression of the raising verb's internal content. As a direct consequence, the ICRP will identify the internal content of the opaque verbs with the internal content of their complements. This is the same effect that we already saw with modal verbs.

At this point we would predict that both a de re and a de dicto reading is available for all sentences in (3), contrary to fact: Whereas the scope of strong quantifiers relative to epistemic modals is not fixed, they outscope other opaque verbs. The difference between epistemic modals and other verbs follows from the principle in (15), which restricts the scope of strong quantifiers that bind variables in the syntactic arguments of verbal functors.

(15) STRONG QUANTIFIER RESTRICTION (SQR):
For each verb v and each NP n that is selected by v: n's index value may not be bound by a strong quantifier inside an argument position of v's main content.

Note that (15) is a typical interface principle since it refers to syntactic and semantic properties of the lexical items. We emphasize that the SQR is a lexical principle that constrains the syntax-semantics interface in verbs.

The split reading of (1-a) obeys the SQR since the main content of the modal is the accessibility relation (**acc**) whose arguments are worlds. The type system already excludes the occurrence of the strong quantifier in an argument position of **acc**. The observed scope ambiguity is due to the interaction of the strong quantifier with the existential quantifier that comes with the modal.

With non-modal raising verbs the semantic representations are crucially different. (16) shows the de re and the unavailable de dicto reading of a strong quantifier with *seem*. The main content of *seem* is **seem**. In (16-b) the strong quantifier appears inside an argument position of **seem**, violating the SQR (15).

(16) Everyone seems to sleep.
a. $\lambda w.\forall x(\textbf{human}(w, x) \rightarrow \underline{\textbf{seem}}(w, \lambda w'.\textbf{sleep}(w', x))$ (de re)
b. $\#\lambda w.\underline{\textbf{seem}}(w, \lambda w'.\forall x(\textbf{human}(x) \rightarrow \textbf{sleep}(w', x))$ (de dicto)

The SQR in (15) is a stipulation at the syntax-semantics interface, but so is the LFS restriction L(i) on the reconstruction of strong quantifiers after syntactic raising. In contrast to the LFS constraint, the SQR also accounts for obligatory de re readings of strong quantifiers in the object position of verbs like *seek*. To see how, consider the two potential readings indicated in (3-c). In the hypothetical de dicto reading the index of the direct object is bound by a strong quantifier inside an argument slot of the verb's main content, the constant **seek**. The SQR correctly rules out this reading, whereas the LFS analysis has nothing to say about it.

3.2 Epistemic Modals in Split Readings in German

The LRS interface forms a layer of constraints between syntax and semantics which shields semantic generalizations from the particulars of syntactic structure and *vice versa*. The advantages of this modular architecture become clear when we now consider the analysis of the split reading in German. Recall from (1) that corresponding finite sentences in English and German with negated strong quantifiers as subjects of epistemic modals exhibit parallel semantic behavior despite the differences in the syntax of the finite verb. In this section we will demonstrate that all substantial differences between the HPSG-based LRS analyses of the corresponding constructions in English and German are located in syntax. The semantic principles remain untouched, as one should expect in a situation with identical semantic generalizations. Most of this section will therefore consist of an explanation of the syntactic analysis and of how the new syntactic structures are hooked up to the interface. The interface neutralizes the syntactic differences, and the semantic constraints that govern the relationship between epistemic modals and strong quantifiers stay firmly in place.

Figure 2 reveals details of our HPSG analysis of sentence (1-b). We choose a syntactic representation with a trace, t_i, of the finite verb *kann*$_i$ in verb second position, because this analysis of German sentence structure is most similar to other frameworks. The constituent in the vorfeld, *nicht jeder*$_j$, is extracted from the mittelfeld using the most conservative HPSG analysis of unbounded dependency constructions, which introduces a trace at the bottom of the unbounded dependency.[6]

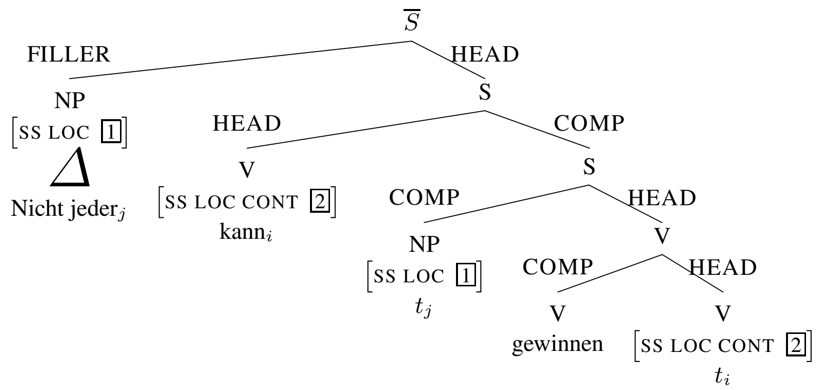

Fig. 2. Syntactic analysis of sentence (1-b).

From the bottom of the tree up to the lower S node, the syntactic structure consists of head-complement combinations like in English. In contrast to English, however, the head of the finite verbal projection is a trace, and so is the NP subject. The verb trace is related to the overt modal in verb second position, indicated here by the identity of subscripts, and the NP trace is related by the identity of LOC values to the topicalized filler constituent.

Before we turn to the traces, let us briefly review those parts of the structure that are already familiar from English. There are no differences between the LRS constraints associated with the non-finite verb *gewinnen* (win) in (17) and the negated strong quantifier *nicht jeder* (not everyone) in (18) compared to their English counterparts. The former correspond to the constraints associated with *make the team* and the latter to the constraints of *not every boy*.

(17) LRS constraint of the verb *gewinnen* (win): $\hat{}A : [\{\underline{\mathbf{win}}(w', x)\}]$

[6] The choice of analysis for German sentence structure has only minor consequences for the analysis of the phenomenon under investigation. A linearization-based account of verb second would eliminate the verb trace, and a traceless theory of complement extraction would eliminate the subject trace by locating the source of the unbounded dependency in the verb. Both changes would in fact make the connection from syntax to the LRS interface more similar to the analysis of English. The syntax of Figure 2 is the most interesting choice in the sense that it brings out best the flexibility of our syntax-semantics interface.

(18) LRS constraint of the NP *nicht jeder*: $\neg D : [{}^\wedge\forall x(\{\underline{\textbf{person}}(x)\} \rightarrow E : [x])]$

In the syntactic analysis of the relationship between the finite verb in verb second and its trace we follow Müller (2007), who assumes that they are ultimately related by lexical rules. We derive the trace t_i and the verb second form of *können* from the same underlying base form.[7] As the surface position of the finite verb has no influence on scope in German, it is generally thought that the verb trace has the same semantics as the verb when it appears in final position. We follow this line of reasoning and assume that the LRS constraint of the modal verb *kann* (can), given in (19), is identical to that of the verb trace.[8]

(19) LRS constraints of the lexeme *kann* (can) and its verb trace t_i:

$${}^\wedge\lambda w.A : [\exists w'(\underline{\textbf{acc}}(w,w') \wedge B : [w', \{C\}])]$$

When the verb trace combines with the non-finite verb the ICRP applies. The resulting LRS constraint of the verbal complex corresponds to the one we saw for the VP *can make the team* in (9).

(20) LRS constraint of *gewinnen* t_i: ${}^\wedge\lambda w.A : [\exists w'(\underline{\textbf{acc}}(w,w') \wedge B : [w', \{\textbf{win}(w',x)\}])]$

The subject trace shares its entire LOCAL value with the filler phrase. Since the *local* structure comprises the CONTENT value, the subject trace shares the referential variable and the main content, both located under CONTENT, with the overt constituent. In contrast to the verb trace, the subject trace does not contribute the same LRS constraints as the fronted constituent. It only contributes those parts that are identical with the filler, the referential variable x and the main content.

(21) LRS constraint of the subject trace t_j: ${}^\wedge F : [\{x\}, \underline{\textbf{person}}]$

When the subject trace combines with the VP *gewinnen* t_i, its LRS constraint is added to the constraint from the verbal projection:

(22) LRS constraint of t_j *gewinnen* t_i: ${}^\wedge\lambda w.A : [\exists w'(\underline{\textbf{acc}}(w,w') \wedge B : [w', \{\textbf{win}(w',x)\}]), x, \textbf{person}]$

Since the verb trace is associated with the same LRS constraint as the finite verb, the latter does not contribute new restrictions. When *kann* combines with its complement in a head-complement structure, its restrictions on external content, internal content and main value unify with those of the complement. Since they are necessarily a subset of the restrictions accumulated in the complement, the LRS constraint at the mother node is identical to the one at the complement.

(23) LRS constraint of $kann_i$ t_j *gewinnen* t_i :

$${}^\wedge\lambda w.A : [\exists w'(\underline{\textbf{acc}}(w,w') \wedge B : [w', \{\textbf{win}(w',x)\}]), x, \textbf{person}]$$

We already saw the LRS constraint of the subject NP in (18). It combines with the verbal projection in a head-filler structure according to standard HPSG principles. While verb second movement does not have an effect on the truth conditions of a sentence, the fronting of an NP is not semantically neutral. Fronting fixes the maximal scope of the quantifier: A fronted quantifier cannot take scope in a higher clause. Head-filler structures obviously introduce restrictions on semantic composition that are different from head-complement clauses, and we need a separate sub-clause of the SEMANTICS PRINCIPLE to treat them:

---

[7] See (Soehn 2006, pp. 188–193) for a discussion of how the right correspondence of trace and verb is guaranteed and for details of the LRS analysis of verb second in German with lexical rules and a verb trace.

[8] In (Soehn 2006) the relationship of the LRS constraint of the finite verb and the trace is mediated by phrase-level LRS principles which lead to the same result that we specify lexically. Our lexicalized alternative might in fact be computationally more attractive.

(24) FILLER SCOPE PRINCIPLE (FSP):
In a head-filler structure, the external content of the filler must be a subexpression of the external content of the head.

Applying the FSP, we obtain the LRS constraint in (25):

(25) LRS constraint of sentence (1-b): $^\wedge\lambda w.A : [\exists w'(\underline{\mathbf{acc}}(w,w') \wedge B : [w', \{\mathbf{win}(w',x)\}]),$
$x, \mathbf{person},$
$\neg D : [\forall x(\mathbf{person}(x) \rightarrow E : [x])]]$

Figure 3 summarizes our analysis. The LRS constraint at the top node has exactly the same resolutions as that for the corresponding English example in Figure 1. In particular the resolution seen in (13-a) yields the split reading.

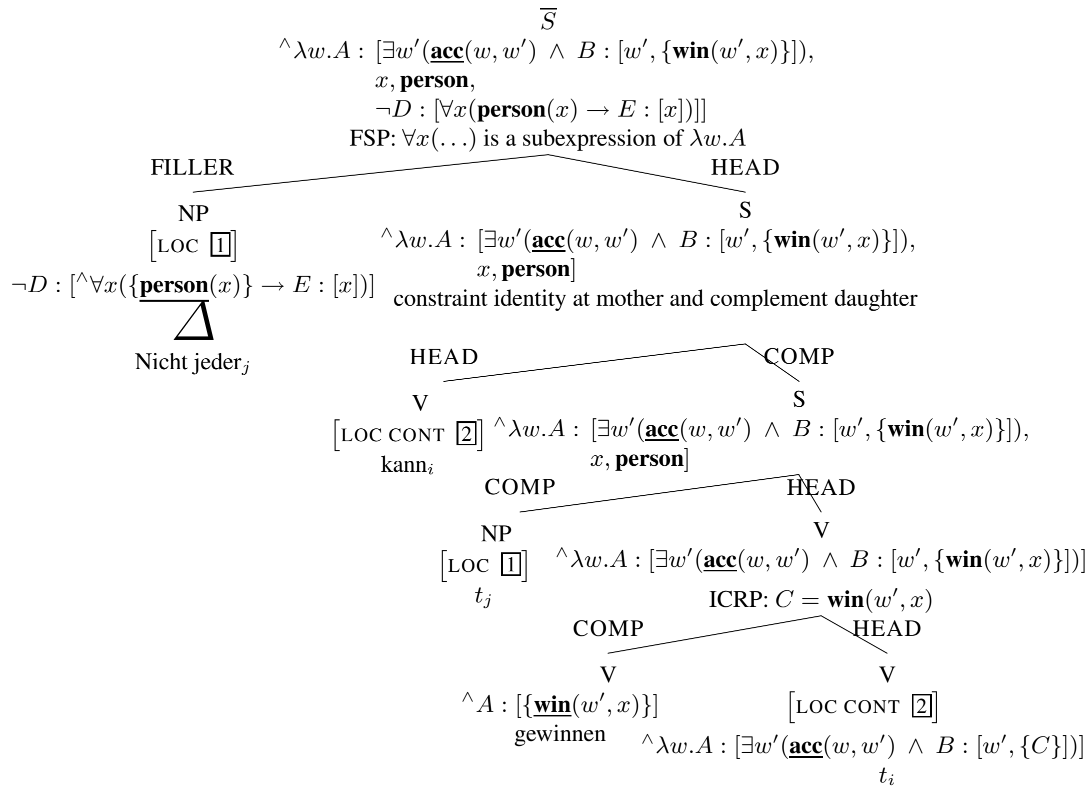

Fig. 3. Sketch of the LRS analysis of sentence (1-b).

Note that, everything else being equal, the SQR, which distinguishes epistemic modals from other raising verbs and opaque verbs such as *suchen* (seek), will apply exactly as it did in English.

4 Further Remarks on Quantification Domains in LRS

In the previous sections we introduced three constraints that restrict the domain of quantification: (i) The QHP in (11) ensured that the scope of a quantifier includes at least the internal content of the head. (ii) The SQR in (15) formulated the idea of the minimal scope of a quantifier and distinguished between the scoping options of strong and weak quantifiers. (iii) The FSP in (24) imposed a restriction on the maximal scope of a quantifier: A moved quantifier may not take scope outside the clause of its surface position.

The QHP and the FSP are sensitive to particular syntactic constructions, the SQR is a lexical constraint on verbs. This expresses the insight that the scoping options of a quantifier may be contingent on its structural position. Possible scopings depend on whether or not a quantifier is extracted, on the type of the quantifier (strong or weak), and on the type of verb whose argument it is. As far as we can see, all three constraints are language-independent, and we take them to belong to the core system of combinatorial constraints in LRS.[9] The SQR was already discussed at the end of Section 3.1. Here we would like to elaborate on the differences between the QHP and the FSP.

In contrast to the QHP, the FSP does not establish a relation between the internal content of the head and the scope of a quantifier in the filler. However, as our German example illustrated, a fronted quantifier binds a variable x in the internal content of the head. In Figure 3 this implied that $\mathbf{win}(w', x)$ was in the scope of the extracted universal quantifier. Let us now take a second look at the FSP on the basis of topicalization and wh-movement constructions in English.

The internal content of the utterance in the two complex sentences in (26) is the constant **claim** combined with its arguments. The two utterances share the same readings (27): In a Montagovian treatment of the indefinite NP *a man from Sweden*, there is a de re and a de dicto reading.

(26) a. Chris claims that Pat loves a man from Sweden.
b. A man from Sweden, Chris claims that Pat loves.

(27) a. de re: $\lambda w.\exists x(\mathbf{man\text{-}f\text{-}Sweden}(x) \wedge \mathbf{claim}(w, \mathbf{chris}, \lambda w'.\mathbf{love}(w', \mathbf{pat}, x)))$
b. de dicto: $\lambda w.\mathbf{claim}(w, \mathbf{chris}, \lambda w'.\exists x(\mathbf{man\text{-}f\text{-}Sweden}(x) \wedge \mathbf{love}(w', \mathbf{pat}, x)))$

In (26-a) the QHP requires that the internal content of the complement clause, $\mathbf{love}(w', \mathbf{pat}, x)$, be in the scope of the existential. This is the case both for the de re and the de dicto reading.[10] Both readings are also available if the quantifier is fronted such as in (26-b). In the de dicto reading, the internal content of the matrix verb is $\mathbf{claim}(w, \mathbf{chris}, \lambda w'.\exists x(\mathbf{man\text{-}from\text{-}Sweden}(x) \wedge \mathbf{love}(w', \mathbf{pat}, x)))$. It follows that the fronted quantifier is contained in the matrix verb's internal content.

On the other hand, when a quantifier is fronted within the complement clause, it cannot take scope beyond its surface position. (28) does not have a reading in which the moved existential takes scope over the matrix verb, and only the reading in (27-b) is available.

(28) Chris claims that a man from Sweden Pat loves.

The clauseboundedness of the scope of a syntactic filler was also observed with wh-phrases. Let us suppose a quantificational treatment from (Karttunen 1977). Karttunen's sentence (29) contains three wh-phrases. What is interesting in the context of our discussion are the scopal options of the two wh-phrases in the complement clause: The clause-initial wh-phrase *where* is fronted and in a filler-head construction. The wh-phrase *which book* remains in situ. The sentence has two readings, indicated by the two possible types of answers.

(29) Who remembers where Mary keeps which book?
a. Bill remembers where Mary keeps which book.
b. Joe remembers where Mary keeps Aspects and Max remembers where Mary keeps Syntactic Structures. (Karttunen 1977, p. 26)

In the first answer, (29-a), *which book* is interpreted as taking scope inside the complement clause. In the second answer, (29-b) the direct object wh-phrase takes scope in the matrix clause. The data are parallel to the scope options in sentence (26-a): An in-situ quantifier can take scope within its clause or outside of its clause. The complement clause in (29) contains a further wh-phrase, *where*. In both

[9] Language-specific constraints on the semantic combinatorics for modeling the typological variation of negative concord in several languages are discussed in (Richter and Sailer 2006).

[10] It is well-documented that an in-situ strong quantifier must take scope inside the clause in which it appears (Stechow 1993). There is no de re reading for (i). To impose this restriction, we could assume a principle such as (ii):

(i) Chris claims that Pat loves every Swedish dish.

(ii) For each clause c: Every strong quantifier that is contributed in c must be a subexpression of c's external content.

readings of the question, the scope of *where* is restricted to the complement clause. This is what we expect as a consequence of the FSP.

5 Conclusion

We argued that the syntactic structure of a sentence should not be made dependent upon the semantic interpretation of scopal elements such as modal verbs, and the semantic interpretation of a scope-taking expression should not necessarily have an effect on the syntactic representation. Syntactic representations and semantic representations should be separate modules which are connected by a flexible interface that enables the linguist to state linguistic generalizations in the appropriate part of the grammatical system. A flexible syntax-semantics interface, based on underspecification techniques, can capture the ambiguities of modal verbs and negated strong quantifiers in a principled fashion. LRS permits the formulation of the necessary scope constraints in a direct way. The overall architecture of the interface is not more complicated than the analysis based on LF syntax since our system is not bound to cast its conditions in terms of syntactic operations. A modular interface of the type we proposed offers a number of clear advantages, such as a significant independence from details of the syntactic representation in the formulation of semantic generalizations, and syntactic and semantic representations that are suitable for computational implementation.

References

Cooper, R.: 1975, *Montague's Semantic Theory and Transformational Grammar*, PhD thesis, University of Massachusetts, Amherst.

Copestake, A., Flickinger, D., Pollard, C. and Sag, I. A.: 2005, Minimal Recursion Semantics: An introduction, *Journal of Research on Language and Computation* **3**, 281–332.

Culicover, P. W. and Jackendoff, R.: 2005, *Simpler Syntax*, Oxford University Press.

Egg, M.: 2004, Mismatches at the syntax-semantics interface, *in* S. Müller (ed.), *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, CSLI Publications, Stanford, pp. 119–139.

Gallin, D.: 1975, *Intensional and Higher-Order Modal Logic*, North-Holland, Amsterdam.

Heim, I. and Kratzer, A.: 1998, *Semantics in Generative Grammar*, Blackwell.

Hendriks, H.: 1993, *Studied Flexibility*, ILLC Dissertation Series 1995-5, Institute for Logic, Language and Computation, Amsterdam.

Karttunen, L.: 1977, Syntax and semantics of questions, *Linguistics and Philosophy* **1**(1), 3–44.

Lechner, W.: 2006, An interpretive effect of head movement, *in* M. Frascarelli (ed.), *Phases of Interpretation*, Mouton de Gruyter, Berlin und New York.

Müller, S.: 2007, *Head-Driven Phrase Structure Grammar. Eine Einführung*, Stauffenburg, Tübingen.

Newmeyer, F. J.: 2008, Semantic evidence and syntactic theory, *in* N. Choeb and B. Zeisler (eds), *Pre-Proceedings of the International Conference on Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, Universität Tübingen, Tübingen, pp. 58–60.

Penn, G. and Richter, F.: 2005, The other syntax: Approaching natural language semantics through logical form composition, *in* H. Christiansen, P. R. Skadhauge and J. Villadsen (eds), *Constraint Solving and Language Processing. First International Workshop, CSLP 2004, Roskilde, Denmark, September 1-3, 2004, Revised Selected and Invited Papers*, Vol. 3438 of *Lecture Notes in Computer Science*, Springer, pp. 48–73.

Pinkal, M.: 1996, Radical underspecification, *in* P. Dekker and M. Stokhof (eds), *Proceedings of the Tenth Amsterdam Colloquium*, ILLC/Department of Philosophy, University of Amsterdam, pp. 587 – 606.

Pollard, C. and Sag, I. A.: 1994, *Head-Driven Phrase Structure Grammar*, University of Chicago Press, Chicago and London.

Richter, F. and Sailer, M.: 2004, Basic concepts of Lexical Resource Semantics, *in* A. Beckmann and N. Preining (eds), *ESSLLI 2003 – Course Material I*, Vol. 5 of *Collegium Logicum*, Kurt Gödel Society Wien, Vienna, pp. 87–143.

Richter, F. and Sailer, M.: 2006, Modeling typological markedness in semantics. The case of negative concord, *in* S. Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar, Varna 2006*, CSLI Publications, Stanford, pp. 305–325.

Sailer, M.: 2006, *Don't Believe* in underspecified semantics. Neg raising in Lexical Resource Semantics, *in* O. Bonami and P. Cabredo Hofherr (eds), *Empirical Issues in Formal Syntax and Semantics*, Vol. 6, CNRS, Paris, pp. 375–403. *http://www.cssp.cnrs.fr/eiss6/sailer-eiss6.pdf

Soehn, J.-P.: 2006, *Über Bärendienste und erstaunte Bauklötze. Idiome ohne freie Lesart in der HPSG*, Peter Lang, Frankfurt am Main. Ph.D. thesis, Friedrich-Schiller-Universität Jena.

Stechow, A. v.: 1993, Die Aufgaben der Syntax, *in* J. Jacobs, A. von Stechow, W. Sternefeld and T. Vennemann (eds), *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*, Vol. 1, Walter de Gruyter, Berlin and New York, pp. 1–88.

Zimmermann, T. E.: 1993, On the proper treatment of opacity in certain verbs, *Natural Language Semantics* **1**(2), 149–179.