Syntax für eine unterspezifizierte Semantik: PP-Anbindung in einem deutschen HPSG-Fragment

Frank Richter   Manfred Sailer
Universität Tübingen
SFB 340, B8

Zusammenfassung

Im vorliegenden Papier wird gezeigt, daß es durch eine starke Linearisierungskomponente der HPSG möglich ist, mehrdeutigen Sätzen mit PP-Adjunkten eine eindeutige syntaktische Struktur zuzuweisen. Die möglichen Lesarten, die von der tatsächlichen Wortreihenfolge abhängen können, werden in einer unterspezifizierten Semantik zusammengefaßt. Zu diesem Zweck integrieren wir eine Higher-Order Constraint Language wie in Bos (1995) definiert in HPSG. Im Unterschied zu anderen Ansätzen definieren wir die semantische Metasprache auf der Objektebene von HPSG.

Abstract

In this paper it is argued that using an HPSG with a strong linearization component it is possible to assign sentences with ambiguous PP-adjuncts a unique syntactic structure. The possible readings, which may depend on the actual word order, are given by an underspecified semantics. For this purpose a higher-order constraint language along the lines of Bos (1995) is integrated into HPSG. In contrast to other approaches, we define the semantic metalanguage on the object level of HPSG.

1 Einleitung

Traditionell wird semantisch ambigen Sätzen mit Adjunkt-PPs für jede Lesart eine eigene syntaktische Struktur zugeordnet. So hat Satz (1a) die Lesart (2a) und (1b) die Lesart (2b).

(1) a. daß Peter [den [Mann mit dem Fernglas]] sieht

b. daß Peter den Mann [[mit dem Fernglas] sieht]

(2) a. `sieht( Peter, mit-dem-Fernglas( den-Mann ))`

b. `mit-dem-Fernglas( sieht( Peter, den-Mann ))`

Im vorliegenden Papier wollen wir im Rahmen der HPSG einen alternativen Vorschlag skizzieren, in dem wir Sätzen mit Adjunkt-PPs genau eine syntaktische Struktur mit einer unterspezifizierten semantischen Repräsentation zuweisen. Einer unterspezifizierten semantischen Repräsentation wird ihrerseits eine Menge von Lesarten zugeordnet werden.

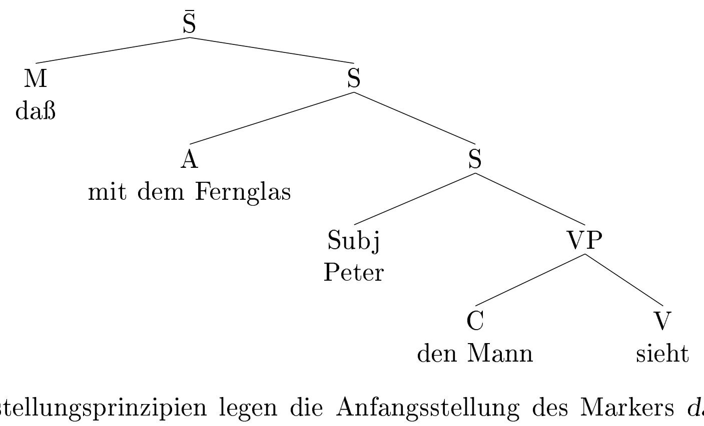
Dieser Vorschlag macht wesentlichen Gebrauch von den formalen Möglichkeiten von Linearisierungsgrammatiken (Reape (1994), Kathol (1995), Richter und Sailer (1995)), die eine wohldefinierte, von der syntaktischen Baumstruktur unabhängige Realisierung der phonologischen Wortabfolge der im Satz vorkommenden Konstituenten erlauben. Wir werden zeigen, wie im Rahmen der auf King (1989) zurückgehenden klassischen denotationellen Semantik für HPS-Grammatiken zur syntaktischen Struktur eine an Bos (1995) orientierte unterspezifizierte semantische Repräsentation definiert werden kann. Diese dient als Metasprache zur Spezifikation der Menge der Lesarten.

Der hier vorgestellte Ansatz erlaubt es, auf PP-Anbindung zurückzuführende Ambiguitäten beim syntaktischen Parsing zu ignorieren, da sie nicht länger mit verschiedenen syntaktischen Strukturen verknüpft sind. Wir werden außerdem zeigen, daß sich die bekannten Skopusambiguitäten bei quantifizierenden Ausdrücken in der HPSG mit den gleichen Techniken behandeln lassen, die wir für PP-Adjunkte definieren.

2 Syntax

Ein wichtiges Kennzeichen der Grammatikarchitektur von HPSG besteht darin, daß Syntax und Semantik eines sprachlichen Zeichens als Teile eines einzigen linguistischen Objektes repräsentiert werden. Das heißt, daß es keine voneinander verschiedenen syntaktischen und semantischen Repräsentationsebenen etwa in dem Sinne gibt, daß eine oder mehrere logische Formen aus einer gegebenen syntaktischen Struktur abgeleitet werden. Gibt es zu einer syntaktischen Struktur zum Beispiel zwei logische Formen, so bedeutet das für die HPSG, daß es im linguistischen Universum zwei voneinander verschiedene, disjunktiv beschriebene Objekte gibt, die sich alleine in dem Teil ihrer Struktur voneinander unterscheiden, welcher als semantische Repräsentation dient.
Betrachten wir den umgekehrten Fall, in dem mehrere syntaktische Stukturen dieselbe unterspezifizierte semantische Repräsentation haben sollen, so handelt es sich auf der Ebene der durch die grammatischen Prinzipien beschriebenen sprachlichen Zeichen wiederum um voneinander verschiedene Objekte, die sich in diesem Fall jedoch nur im syntaktischen Teil ihrer Struktur unterscheiden. Ziel einer unterpezifizierten semantischen Repräsentation ist es, disjunkte Objekte durch eine kompaktere Darstellung zu vermeiden, die nur bei Bedarf aufgelöst wird, und so einer kombinatorischen Explosion zu entgehen. Im Falle der PP-Anbindung, wo wir zwei traditionell verschiedenen syntaktischen Strukturen eine einzige semantische Repräsentation zuordnen wollen, ergeben sich in dem hier skizzierten Rahmen somit zwei Möglichkeiten. Entweder führt man eine Ebene der unterspezifizierten syntaktischen Repräsentation ein, welche disjunkte syntaktische Objekte in einer (neuen) Repräsentation zusammenfaßt. Dies ist die zugrundeliegende Idee bei unterspezifizierten TAGs (Kallmeyer 1996) und Baumgrammatiken (Schiehlen 1996). Oder man gibt die Idee auf, daß es sich überhaupt um verschiedene syntaktische Strukturen handelt. Diesen Weg beschreiten wir bei der Behandlung von Adjunkt-PPs.
Wir nehmen an, daß Sätze im Deutschen unabhängig von der Wortreihenfolge eine einheitliche syntaktische Baumstruktur aufweisen. Kopf der Struktur ist das finite Verb, das zunächst seine Komplemente realisiert bevor Adjunkte, insbesondere auch PP-Adjunkte, in den Baum eingeführt werden. Gegebenenfalls schließt ein Marker die Konstruktion ab.

(3) *Einheitliche syntaktische Struktur der Sätze (1a) und (1b):*



Wortstellungsprinzipien legen die Anfangsstellung des Markers *daß* und die Endstellung des finiten Verbs fest. Dagegen ist die Stellung der Mittelfeldkonstituenten im allgemeinen frei und nicht beschränkt, das heißt, Scrambling ist erlaubt. Andererseits kann die tatsächlich realisierte Wortreihenfolge die möglichen Lesarten des Satzes beschränken. Solche semantischen Prinzipien werden im Abschnitt 4.3 diskutiert.

3 Semantische Repräsentationen

Bislang haben wir uns noch nicht dazu geäußert, wie eine unterspezifizierte semantische Repräsentationen in die Grammatikarchitektur von HPSG eingeführt werden kann. Nach den Vorschlägen von King (1989) und King (1994), die als formale Explikation von (Pollard und Sag 1994) verstanden werden können, besteht in HPSG eine Grammatik aus einer Signatur und einer Theorie, welche eine Menge von Beschreibungen ist. Sie sind in der Regel als Implikationen formuliert.

Beispiele sind das Head Feature Principle und das Subcategorization Principle. Jedes Objekt des linguistischen Universums muß alle Beschreibungen der Theorie erfüllen, und das intuitiv vom Linguisten gemeinte Universum enthält alle Objekte, welche die Beschreibungen erfüllen. Das heißt, das intendierte Universum ist in gewisser Weise maximal bezüglich der Theorie.
Der klassische Ort für Unterspezifikation in der HPSG ist die Beschreibungssprache. Von den Objekten des Universums wird verlangt, daß sie vollständig sind, das heißt, sie sind totally well-typed und sort-resolved (Pollard und Sag 1994). Typische Mittel der Unterspezifikation in der Beschreibungssprache sind die Verwendung von Obersorten der Sortenhierarchie und von Relationen in den Beschreibungen der Theorie. Zum Beispiel trifft die Beschreibung *sign* sowohl auf Phrasen als auch auf Worte zu, und Relationen zwischen Teilobjekten eines linguistischen Zeichens können mehrere Lösungen haben. Das heißt, in der Denotation einer n-stelligen Relation sind alle n-Tupel von (vollständigen) Teilobjekten des Gesamtobjektes, welche die Relation erfüllen.
Für die Verwendung als semantische Repräsentation ist die Beschreibungssprache jedoch in mehrfacher Hinsicht problematisch. Zum einen haben Beschreibungen eine auf mengentheoretischer Basis definierte klassische denotationelle Semantik. Das bedeutet, daß eine unterspezifizierte Beschreibung nichts anderes als eine Disjunktion bedeutet, also eine Menge von ihrerseits vollspezifizierten Objekten. Auf der anderen Seite können einem Objekt, zum Beispiel einer Phrase, in der Regel unendlich viele Beschreibungen zugeordnet werden. Jede Phrase erfüllt zum Beispiel die Beschreibung, ein *sign* zu sein ebenso wie die Beschreibung, eine Phrase mit SYNSEM-Wert *synsem* zu sein.
Um der semantischen Repräsentation eines linguistischen Zeichens jedoch eine sinnvolle denotationelle Semantik geben zu können, ist es erforderlich, daß wir in der Lage sind, einer Phrase im linguistischen Universum eindeutig eine semantische Repräsentation zuzuordnen. Verfolgt man insbesondere die Idee der semantischen Unterspezifikation im Sinne von Bos (1995) und Pinkal (1995), so ergibt sich damit die Forderung, daß jedem Zeichen eindeutig eine Formel der semantischen Metasprache, also in anderen Worten eine unterspezifizierte semantische Repräsentation zuzuordnen ist.
Eine eindeutige Zuordnung einer unterspezifizierten semantischen Repräsentation ist somit nur auf der Objektebene der HPSG möglich. Das Teilobjekt eines sprachlichen Zeichens, das seine Semantik repräsentiert (üblicherweise der Wert des Merkmals CONTENT), interpretieren wir als Objekt der semantischen Metasprache. Im folgenden Abschnitt werden eine Signatur und Prinzipien definiert werden, so daß Sätzen mit PP-Adjunkten auf Objektebene eine eindeutige unterspezifizierte Semantik zugewiesen wird. Die hierzu verwendete Repräsentation lehnt sich an Bos (1995) an. Nach den eben angestellten Überlegungen ist bei der Definition der semantischen Prinzipien in der Beschreibungsprache darauf zu achten, daß es — abgesehen von lexikalischen Ambiguitäten — nicht syntaktisch gleiche Zeichen mit unterschiedlicher semantischer Repräsentation gibt, da dies der Idee der semantischen Unterspezifikation zuwiderliefe. Die semantische Repräsentation ist so zu definieren, daß sie durch die Prinzipien der Beschreibungssprache für Phrasen eindeutig bestimmt wird.

4 Eine unterspezifizierte Semantik für HPSG

4.1 Die Struktur des CONTENT-Wertes

Bos (1995) macht Gebrauch von verschieden Arten von Elementen: Formeln einer beliebigen Logik (Prädikatenlogik, DRT, etc.), Labels und Holes. Holes sind Metavariablen über Formeln, Labels sind lediglich Notationshilfen, die abkürzend für Formeln verwendet werden. Eine gelabelte Formel ist ein Paar aus einem Label und einer Formel.
Um dieses Grundinventar bereitzustellen, führen wir die Sorten *formula*, *label* und *hole* ein, wobei wir eine Sorte *meta* bereistellen, die die Obersorte von *hole* und *label* ist. Wie Objekte der Sorte *formula* intern strukturiert sind soll hier ignoriert werden, da das von der Entscheidung für eine bestimmte semantische Objektsprache (z.B. Prädikatenlogik) abhängt. Für gelabelte Formeln führen wir die Sorte *l(abelled)-formula* ein, mit den Attributen LABEL und FORMULA, die einen *label* bzw. eine *formula* als Wert nehmen. Abkürzend schreiben wir für ein *l-formula*-Objekt mit LABEL-Wert l und FORMULA-Wert ϕ auch '$l : \phi$'.

Neben diesen Grundbegriffen führt Bos Subordinationsbeschränkungen auf Labels und Holes ein. Wir verwenden dazu eine Sorte *constraint* mit den *meta*-wertigen Attributen ARG1 und ARG2. Als einzige Beschränkungsart benutzt Bos das Symbol '$\leq$'. In unserem Ansatz haben wir zwei Subsorten von *constraint*: *leq* ($\leq$) für die Bos'sche Subordination und *not-equ* ($\neq$) für Ungleichheiten. Eine *Unterspezifizierte Repräsentation* (UR) bei Bos ist ein Trippel $\langle H, L, C\rangle$, wobei H eine Menge von Holes, L eine Menge von gelabelten Formeln und C eine Menge von Beschränkungen ist. Da *content*-Objekte in unserem Ansatz den URs von Bos (1995) entsprechen, finden sich die drei Mengen in Form von *list*-wertigen Attributen an *content*-Objekten wieder. Daneben führen wir zwei Hilfsattribute MAIN und TOP ein, die globale Information lokal verfügbar machen (vergl. auch die Relationen **main** und **top** in Bos, Gambäck, Lieske, Mori, Pinkal und Worm (1996)). Damit ergeben sich die folgenden Attribute für *content*-Objekte:[1]

$$\begin{bmatrix} \textit{content} & \\ \textsc{holes} & \textit{list(hole)} \\ \textsc{labels} & \textit{list(l-formula)} \\ \textsc{constr(aints)} & \textit{list(constraint)} \\ \textsc{main} & \textit{l-formula} \\ \textsc{top} & \textit{hole-or-none} \end{bmatrix}$$

In (4) sind die CONTENT-Werte einiger Lexikoneinträge zusammengestellt.

(4) CONTENT-*Werte der Lexikoneinträge*

sieht

$$\begin{bmatrix} \textsc{holes} & \langle h_1 \rangle \\ \textsc{labels} & \langle l_1 : \texttt{sieht}'(x,y)\rangle \\ \textsc{constr} & \langle l_1 \leq h_1 \rangle \\ \textsc{main} & l_1 : \texttt{sieht}'(x,y) \\ \textsc{top} & h_1 \end{bmatrix}$$

Mann:

$$\begin{bmatrix} \textsc{holes} & \langle\ \rangle \\ \textsc{labels} & \langle l_2 : \texttt{mann}'(y)\rangle \\ \textsc{constr} & \langle\ \rangle \\ \textsc{main} & l_2 : \texttt{mann}'(y) \\ \textsc{top} & \textit{none} \end{bmatrix}$$

Fernglas:

$$\begin{bmatrix} \textsc{holes} & \langle\ \rangle \\ \textsc{labels} & \langle l_3 : \texttt{ferng}'(z)\rangle \\ \textsc{constr} & \langle\ \rangle \\ \textsc{main} & l_3 : \texttt{ferng}'(z) \\ \textsc{top} & \textit{none} \end{bmatrix}$$

einen:

$$\begin{bmatrix} \textsc{holes} & \langle h_2, h_3 \rangle \\ \textsc{labels} & \langle l_4 : \texttt{ein}'(y)[h_2 \wedge h_3]\rangle \\ \textsc{constr} & \langle\ \rangle \\ \textsc{main} & l_4 : \texttt{ein}'(y)[h_2 \wedge h_3] \\ \textsc{top} & \textit{none} \end{bmatrix}$$

jedem:

$$\begin{bmatrix} \textsc{holes} & \langle h_4, h_5 \rangle \\ \textsc{labels} & \langle l_5 : \texttt{jed}'(z)[h_4 \rightarrow h_5]\rangle \\ \textsc{constr} & \langle\ \rangle \\ \textsc{main} & l_5 : \texttt{jed}'(z)[h_4 \rightarrow h_5] \\ \textsc{top} & \textit{none} \end{bmatrix}$$

mit:

$$\begin{bmatrix} \textsc{holes} & \langle h_1, h_6 \rangle \\ \textsc{labels} & \langle l_6 : \texttt{mit}'(h_6,z)\rangle \\ \textsc{constr} & \langle l_6 \leq h_1 \rangle \\ \textsc{main} & l_6 : \texttt{mit}'(h_6,z) \\ \textsc{top} & h_1 \end{bmatrix}$$

Für *phrase*-Objekte ergibt sich der CONTENT-Wert aus den CONTENT-Werten der Töchter und dem Semantikprinzip. Die Teile 1 – 4 des Semantikprinzips in (5) regeln die Werte der Attribute MAIN, TOP, HOLES und LABELS, die entweder von einer Tochter 'vererbt' werden oder die Konkatenation der Werte der beiden Töchter darstellen. Teil 5 führt konstruktionsspezifische Beschränkungen auf die CONSTR-Liste des *phrase*-Objekts ein. Die Beschränkungen sind derart zu verstehen, daß die CONSTR-Liste genau die dort beschriebenen Elemente besitzt. Die sich an der Mutter ergebende CONSTR-Liste ist immer dann endlich, wenn die der Töchter es sind.

(5) *Das Semantikprinzip:*

1. Der MAIN-Wert der Mutter ist identisch mit dem MAIN-Wert des semantischen Kopfes (d.h. eines Quantors oder Adjektivs in nominalen Projektionen und des syntaktischen Kopfes sonst)

2. Der TOP-Wert der Mutter ist identisch mit dem TOP-Wert der Kopftochter. In verbalen Kopf-Adjunkt-Strukturen sind die TOP-Werte an Mutter und beiden Töchtern identisch.

3. Die HOLES-Liste der Mutter besteht genau aus den Elementen der HOLES-Listen der Töchter, wobei identische Element nur einmal auftreten.

4. Der LABELS-Wert der Mutter ist die Konkatenation der LABELS-Werte der Töchter.

5. Der CONSTR-Wert der Mutter enthält genau die in (a) – (c) beschriebenen Elemente:

[1] Der TOP-Wert ist von der Sorte *hole-or-none*, d.h. er ist entweder ein *hole* oder er hat den 'dummy'-Wert *none*.

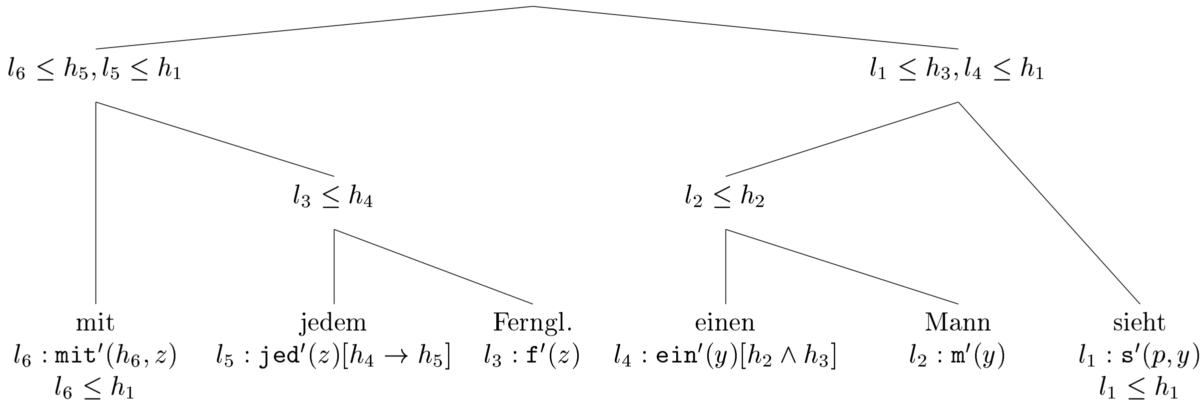
(a) Jedes Element der CONSTR-Listen der Töchter tritt einmal in der CONSTR-Liste der Mutter auf.

(b) Ist eine Tochter innerhalb einer nominalen Projektion ein Quantor Q, die andere Teil des Restriktors R, wobei der MAIN-Wert von Q die Gestalt $l_q : \mathtt{quant}'(x)[h_{restr} \bullet h_{nucl}]$ hat, l_m der MAIN-Wert von R ist, so enthält die CONSTR-Liste auch $l_m \leq h_{restr}$.

(c) Ist in einer Kopf-Komplement-Struktur das Komplement eine quantifizierte NP mit einem MAIN-Wert der Gestalt $l_q : \mathtt{quant}'(x)[h_{restr} \bullet h_{nucl}]$ und ist l_m der MAIN-Wert der Kopftochter und l_t der TOP-Wert der Kopftochter, so enthält die CONSTR-Liste auch $l_q \leq l_t, l_m \leq h_{nucl}$.

In (6) ist die syntaktische Struktur eines Satzes angedeutet, wobei wir das Subjekt und den Marker ignoriert haben (vergl. den Baum in (3)). An den terminalen Knoten sind die Elemente der LABLES- und der CONSTR-Listen angegeben. An den nonterminalen sind lediglich diejenigen *constraint*-Objekte angegeben, die durch Teil 5 des Semantikprinzips neu in die Struktur eingeführt werden. In (7) steht der sich so ergebende CONTENT-Wert des Gesamtsatzes.

(6) *Anwendung des Semantikprinzips auf den Satz:*
daß Peter einen Mann mit jedem Fernglas sieht



(7) *Der* CONTENT-*Wert des Beispielsatzes:*

$$\begin{bmatrix} \textit{content} & \\ \text{HOLES} & \langle h_1, h_2, h_3, h_4, h_5, h_6 \rangle \\ \text{LABELS} & \langle l_1 : \mathtt{s}'(p,y), l_2 : \mathtt{m}'(y), l_3 : \mathtt{f}'(z), l_4 : \mathtt{ein}'(y)[h_2 \wedge h_3], l_5 : \mathtt{jed}'(z)[h_4 \rightarrow h_5], l_6 : \mathtt{mit}'(h_6, z) \rangle \\ \text{CONSTR} & \langle l_1 \leq h_1, l_1 \leq h_3, l_2 \leq h_2, l_3 \leq h_4, l_4 \leq h_1, l_5 \leq h_1, l_6 \leq h_1, l_6 \leq h_5 \rangle \\ \text{MAIN} & l_1 : \mathtt{s}'(p, x) \\ \text{TOP} & h_1 \end{bmatrix}$$

4.2 Modelltheoretische Interpretation

Wir können die Definitionen zur Interpretation von semantischen Repräsentationen aus (Bos 1995) übernehmen, wobei zu beachten ist, daß Bos URs interpretiert, wir hingegen *content*-Objekte. Die semantische Interpretation eines *content*-Objekts ist dabei die Menge der semantischen Interpretationen der möglichen Lesarten.
Bos definiert eine Subordinationsrelation SUB auf URs. Seine Definition läßt sich leicht auf *content*-Objekte übertragen. Für jedes *content*-Objekt u definieren wir H_u als die Menge der *hole*-Objekte, die als Elemente im HOLES-Wert von u auftreten, und L_u als die Menge der *label*-Objekte, die als LABEL-Wert eines Elements der LABELS-Liste von u vorkommen.
Zwei *meta*-Objekte k, k' aus $H_u \cup L_u$ stehen in der SUB-Beziehung ($SUB_u(k, k')$) wenn k und k' identisch sind, oder $k \leq k'$ ein Element der CONSTR-Liste von u ist, oder k' ein *label* ist und k ein *hole*, das als Argument in einer gelabelten Formel der LABELS-Liste von u vorkommt, die die Form $k' : \phi$ hat. SUB_u ist abgeschlossen unter Transitivität.

Eine Lesart wird durch ein *Plugging* festgelegt. Ein Plugging ist eine bijektive Abbildung von H_u in L_u. Ein *mögliches* Plugging liegt vor, wenn es sämtliche Beschränkungen erfüllt, die sich aus SUB_u ergeben.
Unter einem gegebenen Plugging P gibt es genau ein *hole*-Objekt h, so daß für alle $k \in H_u \cup L_u$ gilt, wenn $SUB_u(h,k)$, dann $P(h) = k$ oder $h = k$. Dieses *hole*-Objekt nennen wir in Anlehnung an Bos (1995, Definition 6) $TOP_{P,u}$. Die Interpretation eines *content*-Objekts bezüglich eines semantischen Modells M ergibt sich dann wie bei Bos als:

$$[\![u]\!]_M = \{[\![TOP_{P,u}]\!]_{M,P} | P \text{ ist ein mögliches Plugging}\}$$

Bos (1995) zeigt, daß verschiedene Logiken (Prädikatenlogik, Dynamische Prädikatenlogik, Diskursrepräsentationstheorie) als semantische Objektsprache der von ihm eingeführten URs verwendet werden können. Wir gehen deshalb davon aus, daß die Interpretationsfunktion $[\![\ ]\!]_M$ der semantischen Objektsprache für *formula*-Objekte, in denen keine *holes* auftreten, definiert ist. Für ein *hole*-Objekt h mit $P(h) = l$, so daß es ein Element der LABELS-Liste der Form $l : \phi$ gibt, gilt $[\![h]\!]_{M,P} = [\![\phi]\!]_{M,P}$. Entsprechend gilt für *formula*-Objekte in denen *holes* $h_1, \ldots h_n$ als Argumente auftreten:

$$[\![\phi]\!]_{M,P} = [\![\ \phi\ [[\![h_1]\!]_{M,P}/h_1, \ldots, [\![h_n]\!]_{M,P}/h_n]\ ]\!]_{M,P}.$$

Mit den gegebenen Definitionen erhalten wir genau 5 mögliche Pluggings für den Beispielsatz, vier erwünschte und ein abwegiges.

(8) a. Verbale Modifikation, *einen* hat weiten Skopus:

$P_1 = \{h_1 = l_4, h_2 = l_2, h_3 = l_5, h_4 = l_3, h_5 = l_6, h_6 = l_1\}$

$\texttt{ein}'(y)[\texttt{mann}'(y) \wedge \texttt{jed}'(z)[\texttt{fern}'(z) \rightarrow \texttt{mit}'(\texttt{sieht}'(p,y), z)]]$

(Es gibt einen Mann, so daß Peter ihn mit jedem Fernglas sieht)

b. Verbale Modifikation, *jedem* hat weiten Skopus:

$P_2 = \{h_1 = l_5, h_2 = l_2, h_3 = l_6, h_4 = l_3, h_5 = l_4, h_6 = l_1\}$

$\texttt{jed}'(z)[\texttt{fern}'(z) \rightarrow \texttt{ein}'(y)[\texttt{mann}'(y) \wedge \texttt{mit}'(\texttt{sieht}'(p,y), z)]]$

(Für jedes Fernglas gilt, daß es einen Mann gibt, den Peter mit diesem Fernglas sieht.)

c. Nominale Modifikation, *einen* hat weiten Skopus:

$P_3 = \{h_1 = l_4, h_2 = l_5, h_3 = l_1, h_4 = l_3, h_5 = l_6, h_6 = l_2\}$

$\texttt{ein}'(y)[\texttt{jed}'(z)[(\texttt{fern}'(z) \rightarrow \texttt{mit}'(\texttt{mann}'(y), z)) \wedge \texttt{sieht}'(p,y)]]$

(Es gibt einen Mann, der jedes Fernglas bei sich hat, und den Peter sieht.)

d. Nominale Modifikation, *jedem* hat weiten Skopus:

$P_4 = \{h_1 = l_5, h_2 = l_6, h_3 = l_1, h_4 = l_3, h_5 = l_4, h_6 = l_2\}$

$\texttt{jed}'(z)[\texttt{fern}'(z) \rightarrow \texttt{ein}'(y)[\texttt{mit}'(\texttt{mann}'(y), z) \wedge \texttt{sieht}'(p,y)]]$

(Zu jedem Fernglas gibt es einen Mann, der dieses Fernglas bei sich hat, und den Peter sieht.)

e. Unerwünschtes Plugging

$P_5 = \{h_1 = l_5, h_2 = l_2, h_3 = l_1, h_4 = l_3, h_5 = l_6, h_6 = l_4\}$

$\texttt{jed}'(z)[\texttt{fern}'(z) \rightarrow \texttt{mit}'(\texttt{ein}'(y)[\texttt{mann}'(y) \wedge \texttt{sieht}'(p,y)], z)]$

(? Für jedes Fernglas gilt: mit ihm geschieht es, daß es einen Mann gibt, den Peter sieht.)

Die Pluggings P_1 – P_4 stellen genau die 4 Lesarten des Beispielsatzes dar, d.h. jede Anbindungsmöglichkeit der PP mit jeder Skopusmöglichkeit der Quantoren gepaart.
P_5 entsteht hier aufgrund der stark vereinfachten semantischen Strukturen. Wir könnten unsere *content*-Objekte mit Sorten für semantische Typen anreichern. Dann wäre es auch möglich, getypte *hole*-Objekte zuzulassen. So hätten wir für *sieht* und *mit* dann die folgenden Lexikoneinträge:

(9) $\begin{bmatrix} \textit{content} & \\ \text{HOLES} & \langle h_1, h_2 \rangle \\ \text{LABELS} & \langle l_1 : h_2(x,y), l_2 : \texttt{sieht}' \rangle \\ \text{CONSTR} & \langle l_1 \leq h_1, l_2 \leq h_2 \rangle \\ \text{MAIN} & l_1 : h_2(x,y) \\ \text{TOP} & h_1 \end{bmatrix}$ $\begin{bmatrix} \textit{content} & \\ \text{HOLES} & \langle h_1, h_3 \rangle \\ \text{LABELS} & \langle l_3 : \texttt{mit}'(h_3, z) \rangle \\ \text{CONSTR} & \langle l_3 \leq h_1 \rangle \\ \text{MAIN} & l_3 : \texttt{mit}'(h_3, z) \\ \text{TOP} & h_1 \end{bmatrix}$

In (9) wurde im Lexikoneintrag von *sieht* über das Prädikat abstrahiert, d.h. statt $l_1 : \texttt{sieht}'(x,y)$ haben wir jetzt eine gelabelte Formel mit den Argumenten ($l_1 : h_2(x,y)$) und eine mit dem Prädikat ($l_2 : \texttt{sieht}'$). Bei verbaler Modifikation wäre dann $h_2 = l_3$ und $h_3 = l_2$, d.h. *mit* verbindet sich semantisch mit dem Prädikat und nicht, wie bisher, mit einer Aussage.
Im Lexikoneintrag von *mit* hätten wir eine semantische Typenrestriktion derart, daß h_3 nicht vom Typ 'Wahrheitswert' sein darf. Eine solche Lösung wäre eine Generalisierung des (Bos 1995)-Ansatzes, die wir hier aus Platzgründen nicht weiter verfolgen.

4.3 Ergänzungen zum Semantikprinzip

Die lineare Position von PP-Adjunkten hat Einfluß auf die möglichen Lesarten. So ist in Satz (10a) ein Bezug der PP auf die NP *einen Mann* möglich, in (10b) nicht, obwohl beide Sätze die gleiche syntaktische Struktur haben.

(10) a. daß Peter einen Mann mit jedem Fernglas sieht

b. daß Peter mit jedem Fernglas einen Mann sieht

Innerhalb des Mittelfelds gilt, daß eine Adjunkt-PP eine nominale Projektion nur dann modifizieren kann, wenn sie ihr direkt phonologisch folgt. Dies können wir in einer Ergänzung zu Teil 5 des Semantikprinzips in (5) ausdrücken:

(d) In einem Satz gilt, für jedes dominierte Adjunkt A, dessen MAIN-Wert ein *hole*-Objekt h als Argument enthält, und für jede dominierte N-Projektion N mit MAIN LABEL-Wert l:

wenn die Phonologie von A in der Gesamtphonologie nicht direkt auf die von N folgt, so enthält der CONSTR-Wert des Satzes auch $h \neq l$.

Durch diese Erweiterung des Semantikprinzips, ergibt sich für den Satz in (10b) ein CONTENT-Wert, der dem in (7) gleicht, aber zudem noch die folgenden Ungleichungen in seiner CONSTR-Liste hat: $h_6 \neq l_2$ und $h_6 \neq l_4$. Wir müssen jetzt nur noch fordern, daß ein mögliches Plugging nicht nur die '$\leq$'-Bedingungen respektiert, sondern auch die '$\neq$'-Bedingungen der CONSTR-Liste. Damit sind die Pluggings P_3, P_4 (und P_5) für Satz (10b) korrekt ausgeschlossen.
Wenden wir uns zuletzt noch Sätzen mit mehreren Adjunkt-PPs zu. Die in (12) eingeführten Klammerungen sollen die intendierte Lesart andeuten.

(11) daß Peter den Mann mit dem Fernglas aus Köln sieht

(12) a. daß Peter den [Mann mit dem [Fernglas aus Köln]] sieht

b. daß Peter den [[Mann mit dem Fernglas] aus Köln] sieht

c. daß Peter den Mann [mit dem [Fernglas aus Köln]] sieht]

d. * [aus Köln [mit dem Fernglas sieht]]

In Satz (11) treten zwei Adjunkt-PPs auf. Nehmen wir an, die erste PP ist ein nominaler Modifikator. Die zweite PP kann *Fernglas* modifizieren, da *Fernglas* vom Gesamtsatz dominiert wird und phonologisch der zweiten PP direkt vorausgeht. Alternativ kann die zweite PP auch die gesamte erste PP zusammen mit deren Argument modifizieren. Wir erhalten also die Lesarten (12a) und (12b). Ist die erste PP ein verbaler Modifikator, dann kann die zweite PP wieder *Fernglas* modifizieren (12c).

Treten im Mittelfeld eines Satzes mehrere verbale Modifikatoren auf, so wird angenommen, daß ihre gegenseitige Einbettung sich aus ihrer relativen Reihenfolge ergibt (Jacobs 1988, S. 20). Deshalb ist die in (12d) angedeutete Lesart auch mit anderem lexikalischen Material nicht möglich.
Wir können diese Generalisierung in unseren Ansatz integrieren, indem wir die Sortendefinitionen für *constraint* erweitern. Die Sorten *leq* und *not-equ* werden zu Untersorten von *simple-constraint*, und wir führen die neue Sorte *complex-constraint* ein, die Listen von *constraint*-Objekten als Werte der Attribute ARG1 und ARG2 nimmt. Ein Objekt der Sorte *complex-constraint* wird vereinfacht als $[\langle c_1, \ldots c_n\rangle \rightarrow \langle c_{n+1}, \ldots c_m\rangle]$ notiert und ist intuitiv als Implikation über Beschränkungen aufzufassen.[2]
Die benötigte Ergänzung zu Teil 5 des Semantikprinzips lautet dann:

(e) In einem Satz, dessen MAIN LABEL-Wert l ist und der genau die Adjunkte $A_1, \ldots A_n$ mit den MAIN LABEL-Werten $l_1, \ldots, l_n$ und mit den *hole*-Objekten $h_1, \ldots, h_n$ in ihren MAIN-Werten dominiert, gilt für jedes i, j mit $1 \leq i, j \leq n$:

falls A_i phonologisch vor A_j steht, ist $[\langle h_i \leq l, l \leq h_i\rangle \rightarrow \langle h_j \neq l_i\rangle]$ ein Element der CONSTR-Liste des Satzes.

'$\rightarrow$'-Bedingungen sollen in intuitiver Weise die Menge der möglichen Pluggings einschränken: wenn ein Plugging P die Bedingungen der ARG1-Liste erfüllt, dann erfüllt es auch die der ARG2-Liste. Wir verzichten an dieser Stelle auf eine formale Definition.

5 Zusammenfassung

Zum Thema der unterspezifizierten Semantik liegt eine Reihe von Veröffentlichungen vor (wie beispielsweise Reyle (1993), Frank und Reyle (1995), Egg und Lebeth (1995), Bos (1995), Pinkal (1995), Schiehlen (1996), Bos et al. (1996)).
In Egg und Lebeth (1995) und Frank und Reyle (1995) werden Vorschläge innerhalb der HPSG gemacht. Während diese beiden Ansätze jedoch die semantische Unterspezifikation in der Beschreibungssprache ansiedeln, haben wir dafür argumentiert, den CONTENT-Wert von Zeichen als unterspezifizierte semantische Repräsentation zu betrachten. Zu diesem Zweck wurden HPSG-Prinzipien definiert, in deren Denotation Objekte liegen, die wie die von Bos (1995) vorgestellte UR als Metasprache einer beliebigen semantischen Objektsprache (z.B. Prädikatenlogik) interpretiert werden. Die damit zur Verfügung stehende unterspezifizierte semantische Repräsentation kann als eine Variante der von Pinkal (1995) diskutierten *Higher-Order Constraint Language* aufgefaßt werden.
Es wurde gezeigt, daß das Konzept der Linearisierungsgrammatiken in HPSG es erlaubt, Sätzen mit Adjunkten, die sowohl nominale als auch verbale Projektionen modifizieren können, eine eindeutige syntaktische Struktur zu geben. Adjunkte werden unabhängig von ihrem Skopus oberhalb der Komplemente des Verbs im Baum eingeführt. Zur semantischen Interpretation dieser Strukturen wurde der Vorschlag von Bos (1995) zu einer unterspezifizierten Semantik in HPSG integriert. Anhand von Beispielen wurde dargestellt, daß in diesem Ansatz nicht nur eine unterspezifizierte Anbindung von PP-Adjunkten, sondern auch die bekannten Skopusambiguitäten bei nominalen Quantoren behandelt werden können. Zusätzliche Prinzipien der HPSG-Beschreibungssprache können dazu verwendet werden, die semantische Constraintmenge so zu erweitern, daß bestimmte Lesarten bei manchen Wortreihenfolgen im Satz ausgeschlossen werden.
Der hier beschriebene Ansatz legt sich nicht auf eine serielle oder parallele Verarbeitung von syntaktischer und semantischer Information fest, sondern bleibt deklarativ. Allerdings werden lexikalische Ambiguitäten zum Beispiel von Präpositionen bereits in der Syntax aufgelöst. Desgleichen muß die Unterscheidung zwischen Adjunkten und Komplementen bereits beim syntaktischen Parsen getroffen werden.
Eine Variante der im vorliegenden Papier skizzierten Theorie der Behandlung von Adjunkt-PPs wurde im Fragmentprojekt B8 des SFB 340 im ConTroll-System (Götz und Meurers 1995) implementiert.

[2] Vergl. auch Schiehlen (1996) für einen Ansatz, der komplexe 'Kontextbedingungen' integriert.

Literatur

Bos, J.: 1995, Predicate logic unplugged, *Proceedings of the 10th Amsterdam Colloquium*.

Bos, J., Gambäck, B., Lieske, C., Mori, Y., Pinkal, M. und Worm, K.: 1996, Compositional semantics in Verbmobil, *Proceedings of COLING '96*.

Egg, M. und Lebeth, K.: 1995, Semantic underspecification and modifier attachment, *Proceedings der Jahrestagung der DGfS*.

Frank, A. und Reyle, U.: 1995, Principle based semantics for HPSG, *Proccedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, pp. 9 – 16.

Götz, T. und Meurers, D.: 1995, Compiling HPSG type constraints into definite clause programms, *ACL '95*.

Jacobs, J.: 1988, Probleme der freien Wortstellung im Deutschen, *Sprache und Pragmatik* **5**, 8–37.

Kallmeyer, L.: 1996, Unterbestimmtheit in Tree Description Grammars, Handout.

Kathol, A.: 1995, *Linearization-Based German Syntax*, PhD thesis, Ohio State University.

King, P. J.: 1989, *A logical Formalism for Head-Driven Phrase Structure Grammar*, PhD thesis, University of Manchester.

King, P. J.: 1994, An expanded logical formalism for Head-Driven Phrase Structure Grammar, *Arbeitspapiere des SFB 340 59*, Universität Tübingen.

Pinkal, M.: 1995, Radical underspecification, *Proceedings of the 10th Amsterdam Colloquium*, pp. 587 – 606.

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

Reape, M.: 1994, Domain Union and Word Order Variation in German, *in* J. Nerbonne, K. Netter und C. Pollard (eds), *German in Head-Driven Phrase Structure Grammar*, Vol. 46, CSLI, pp. 151–197.

Reyle, U.: 1993, Dealing with ambiguities by underspecification: Construction, representation and deduction, *Journal of Semantics* **10**(2), 123 – 179.

Richter, F. und Sailer, M.: 1995, *Remarks on Linearization. Reflections on the Treatment of LP-Rules in HPSG in a Typed Feature Logic*, Magisterarbeit, Universität Tübingen.

Schiehlen, M.: 1996, Semantic construction from parse forests, *Proceedings of COLING '96*.