

Dependenzbasiertes Parsing des Deutschen

Analysen und Abdeckungsbereich

Johannes Heinecke und Andreas Nolda
Humboldt-Universität zu Berlin

DAWAI-Memo B-1/98, 10. November 1999

Inhaltsverzeichnis

1	Ein dependenzbasierter Parser für das Deutsche	1
2	Benennung von Relationen	1
2.1	Benennung morphosyntaktischer Relationen	1
2.2	Benennung semantischer Relationen	4
3	Relationen zum Aufbau atomarer syntaktischer Einheiten	8
4	Komplementrelationen	9
4.1	Verbkomplemente	9
4.1.1	Komplemente von Vollverbformen	9
4.1.2	Komplemente von Auxiliar- und Modalverbformen	12
4.1.3	Komplemente von Kopulaverbformen	13
4.1.4	Komplemente von Kausativverbformen	14
4.1.5	Komplemente von Passiv-Auxiliarverbformen	15
4.1.6	Komplemente von Medial-Auxiliarverbformen	16
4.2	Adjektivkomplemente	16
4.3	Substantivkomplemente	17
4.4	Adpositionskomplemente	17
4.5	Subjunktionskomplemente	19
5	Angaberelationen	20
5.1	Adverbialbestimmungen	20
5.2	Attribute	21
5.2.1	Attribute von Substantivformen	21
5.2.2	Attribute von Adjektivformen	22
5.2.3	Attribute von Adverbformen	22
6	Determinationsrelationen	22

7 Skopusrelationen	23
8 Topologische Struktur	23

Abbildungsverzeichnis

1	Lokalrelationen	5
2	Transitive Konstruktionen (Syntax)	10
3	Transitive Konstruktionen (Semantik)	10
4	Kontrollverb-Konstruktion mit verbalem Objekt (Syntax/Semantik) .	11
5	Kontrollverb-Konstruktion mit subjunktionalem Objekt (Syntax/Semantik)	11
6	ACI-Konstruktion (Syntax und Semantik)	12
7	Auxiliarverb-Konstruktion (Syntax und Semantik)	13
8	Modalverb-Konstruktion (Syntax und Semantik)	13
9	Kopulaverb-Konstruktionen (Syntax)	14
10	Kopulaverb-Konstruktionen (Semantik)	14
11	Kausativ-Konstruktion (Syntax und Semantik)	15
12	Passiv-Konstruktion (Syntax und Semantik)	15
13	Medial-Konstruktion mit <i>lassen</i> (Syntax und Semantik)	16
14	Medial-Konstruktion ohne <i>lassen</i> (Syntax und Semantik)	16
15	Adjektivkomplemente (Syntax/Semantik)	17
16	Zirkumpositionskomplement (Syntax/Semantik)	18
17	Präpositionskomplement I (Syntax/Semantik)	18
18	Präpositionskomplement II (Syntax/Semantik)	19
19	Subjunktionkomplement (Syntax/Semantik)	19
20	Uhrzeitangabe-Konstruktion (Syntax/Semantik)	20
21	Datumsangabe-Konstruktion (Semantik)	21

1 Ein dependenzbasierter Parser für das Deutsche

Der im Rahmen des DFG-Projekts „Ressourcenbeschränkte Analyse gesprochener Sprache“ (DAWAI) entwickelte Parser ist in der Lage, eine wesentliche Teilmenge der Sätze (insbesondere umgangssprachlicher) standarddeutscher Idiolekte zu analysieren.¹ Dabei greift er auf linguistische Daten zurück, die im wesentlichen aus einem Vollformenlexikon und einer Dependenzgrammatik bestehen.² Die Lexikoneinträge — Merkmalstrukturen bestimmter Art — beschreiben neben der orthographischen Form die morphosyntaktischen und (in einem den Erfordernissen entsprechenden Umfang) die semantischen Eigenschaften eines relevanten Teils der deutschen Wortformen. Die mit Hilfe von statistisch bewerteten lokalen Constraints³ formulierte Grammatik beschränkt im Sinne des Eliminativen Parsings⁴ die auf der morphosyntaktischen und der rollensemantischen Beschreibungsebene angesetzten Dependenzstrukturen, wobei so weit wie möglich auf Informationen aus dem Lexikon zurückgegriffen wird. Die Strukturen werden — wie üblich — durch geordnete Baumgraphen modelliert, wobei jedem von der Wurzel verschiedenen Knoten ein orthographisches Wort zugeordnet ist und die Kanten mit den Namen der vorkommenden morphosyntaktischen bzw. rollensemantischen Relationen markiert sind; sie entsprechen, was die Morphosyntax betrifft, weitgehend den von Kunze (1975) motivierten Strukturen.

In dem vorliegenden Bericht werden die für die beiden Beschreibungsebenen angesetzten Analysen vorgestellt. Dadurch erfolgt zugleich eine Charakterisierung des Abdeckungsbereichs des DAWAI-Parsers.

2 Benennung von Relationen

2.1 Benennung morphosyntaktischer Relationen

Da die traditionelle Grammatik nicht für alle morphosyntaktischen Relationen, die für die Beschreibung des Deutschen relevant sind, Bezeichnungen zur Verfügung stellt, war es notwendig, im Rahmen von DAWAI eigene Bezeichnungen festzulegen. Dabei wurde aus Gründen der Einheitlichkeit ganz auf traditionelle Termini verzichtet. Die Benennung dieser Relationen ist in DAWAI durch die im folgenden dargestellte Kon-

¹Nicht erfaßt sind gegenwärtig Phänomene wie Relativsätze, idiomatische Wendungen, Koordination und Ellipse.

²Daneben werden eine Sorten- und eine Kategorien-Hierarchie verwendet.

³„Lokal“ sind diese Constraints dahingehend, daß ihre Aussagen auf maximal zwei Kanten relativiert sind.

⁴Vgl. hierzu u. a. Maruyama (1990) und Menzel (1995).

vention geregelt.⁵

Die Relationsnamen werden nach dem folgenden Schema gebildet:

$$\textit{Relationsname} ::= \\ \textit{Dependenskategorie?} _? \textit{Regenskategorie} _ \textit{Relationstyp} \textit{Relationsnummer?}$$

Der *Relationstyp*-Teil des Relationsnamen kennzeichnet den allgemeinen Typ einer morphosyntaktischen Relation. Im einzelnen kann er die folgende Werte annehmen:

$$\textit{Relationstyp} ::= \\ \text{ANT} \mid \text{C} \mid \text{D} \mid \text{EXPL} \mid \text{M} \mid \text{P} \mid \text{SC}$$

Die Werte von *Relationstyp* stehen dabei für die (englischen) Ausdrücke:

ANT	dislocated antecedent
C	complement
D	determiner
EXPL	expletive
M	modifier
P	word part
SC	scope particle

P ist also Bestandteil eines Namens einer Relation, die zum Aufbau atomarer syntaktischer Einheiten dient, während es sich bei ANT, C, D, EXPL, M und SC um Namensteile von Relationen handelt, mit deren Hilfe komplexe syntaktische Einheiten beschrieben werden — also um *syntaktische Relationen*.

Die *Relationsnummer* wird angegeben, wenn eine bestimmte Anzahl > 1 von verwandten Relationen zu unterscheiden ist.

$$\textit{Relationsnummer} ::= \\ 1 \mid 2 \mid 3 \mid 4$$

Eine *Relationsnummer* wird nur im Zusammenhang mit dem *Relationstyp* C verwendet.

Die *Regenskategorie* kennzeichnet die Elemente des Nachbereichs der betreffenden Relation in bezug auf bestimmte Wortkategorien.⁶

⁵Diese wurde von traditionellen französischen Bezeichnungen wie „complément de verbe“ oder „complément de nom“ inspiriert.

⁶Zum Begriff der Wortkategorie vgl. Lieb (1993).

Regenskatégorie ::=
 ADJ | ADP | AUX | CARD | CAUS | CIRCP | COP | MOD | N | PASS | PRO | SUBJ | V

Die Werte von *Regenskatégorie* stehen dabei für die Ausdrücke:

ADJ	adjective
ADP	adposition
AUX	auxiliary
CARD	cardinal number word
CAUS	causative verb
CIRCP	circumposition
COP	copula verb
MOD	modal verb
N	noun
PASS	passive auxiliary
PRO	pronoun
SUBJ	subjunction
V	full verb

Die *Dependenskatégorie* dient dazu, bei Bedarf Subtypen von syntaktischen Relationen zu spezifizieren, indem die Elemente des Vorbereichs der betreffenden Relation in bezug auf bestimmte Wortkategorien, Einheitenkategorien⁷ oder orthographische Wörter gekennzeichnet werden.

Dependenskatégorie ::=
 ADJ | ADP | ADV | AM | GEN | N | PART | RIGHT | ZU

Die Werte von *Dependenskatégorie* stehen für die Ausdrücke:

ADJ	adjective
ADP	adposition
ADV	adverb
AM	<i>am</i>
GEN	genitive noun
N	noun
PART	verb particle
RIGHT	'right' part of a complex wordform
ZU	<i>zu</i>

Entsprechend dieser Grundsätze sind beispielsweise „V_C1“ als „verb complement 1“ („Subjekt“ in traditioneller Terminologie) und „ADJ_N_M“ als „adjectival noun modifier“ („Adjektivattribut“) zu lesen.

⁷Zum Begriff der Einheitenkategorie vgl. Lieb (1993).

„S“ (für „Satz“) kennzeichnet ‚Wurzelkanten‘.⁸

2.2 Benennung semantischer Relationen

Die Semantikkomponente in DAWAI beschränkt sich auf die Selektion auf Rollen bzw. überprüft die sortale Restriktion bei Prädikat-Aktant Kanten (in der Regel Verb/prädikatives Adjektiv/Prädikatsnomen-Substantiv).

Im Falle von Kanten, die Aktantifizierungen von Argumentstellen bezeichnen, werden die Kanten nach der vom Regens selektierten Rolle benannt:

EXP	Experiencer
THEME	Theme
AG	Agent
PAT	Patient
REC	Recipient
GOAL	Goal
COMIT	Komitativ (präpositional)

In einigen Partizipial- oder Infinitiv-Konstruktionen (*die das Buch lesende Frau*) wird die Rolle vom Dependens selektiert:

EXP_REV	Experiencer
THEME_REV	Theme
AG_REV	Agent
PAT_REV	Patient
REC_REV	Recipient
GOAL_REV	Goal

Lokale Relationen werden unabhängig davon, ob sie obligatorisch (*wohnen in*) oder fakultativ (*essen in*) sind mit folgenden Kantennamen bezeichnet, die die Lokalisierung des Dependens ausdrücken. Lokale Relationen können statisch oder direktional (in Richtung auf die lokale Position, bzw. von der Position kommen) sein:

⁸Diese Bezeichnung unterscheidet sich qualitativ von den anderen, da es sich hierbei nicht um einen Relationsnamen, sondern um einen Konstituentenkategorienamen handelt.

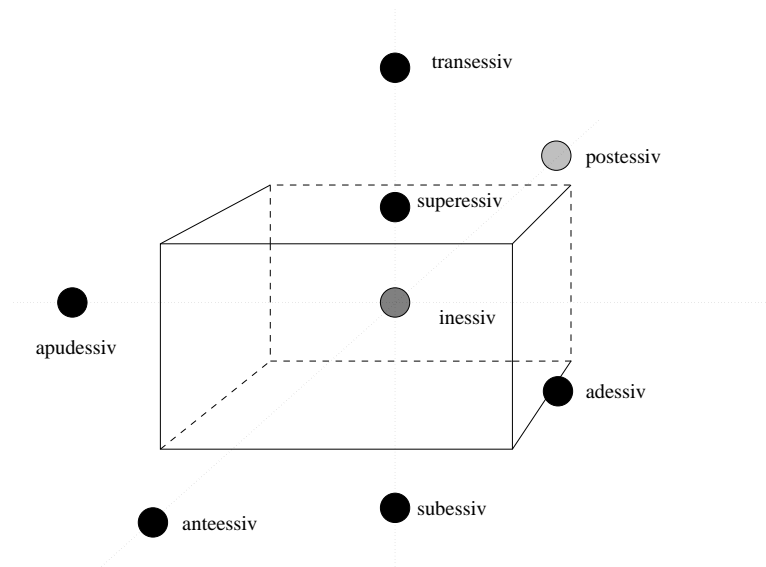


Abbildung 1: Lokalrelationen

Die dadurch definierten Lokalisierungsrelationen werden im Deutschen in der Regel durch Präpositionalphrasen ausgedrückt. Der Unterschied von statischen (-essiv) und direktionalen (-lativ) wird dabei vom Kasus des von der Präposition regierten Dependens markiert:

Rollenpräfix	Präposition	Kasus des Dependens			
		-ESSIV	-LATIV		
			Bewegungstyp		
			<i>hin</i>	<i>weg</i>	<i>via</i>
IN-, IL- AD-	in	+DAT	+ACC		
	nach ⁹ zu ¹⁰		+DAT +DAT		
TRANS- SUPER- EX- SUB- ANTE- POST-	an	+DAT	+ACC		
	über	+DAT			+ACC
APUD-	auf	+DAT	+ACC		
	aus			+DAT	
AB- VIA-	unter	+DAT	+ACC		
	von	+DAT	+ACC	+DAT	
	vor	+DAT	+ACC		
	hinter	+DAT	+ACC		
	neben	+DAT	+ACC		
	bei	+DAT			
	zu		+DAT		
	durch			+DAT	+ACC

Tabelle 1: Lokalrelationen und ihre Präpositionen

Temporale Kanten: Ein Problem besteht für die Semantik darin, daß temporale Ausdrücke aus mehreren Konstituenten bestehen. Der gesamte temporale Ausdruck kann mit einer Kante DATE oder TIME an sein Regens angeschlossen werden. Jedoch müssen die einzelnen Konstituenten auch innerhalb des temporalen Ausdrucks verknüpft werden, was zusätzliche Kantennamen erfordert (vgl. Beispiel 20 auf Seite 20).

TIME	Temporale Rollen (Uhrzeiten, Datum)
DATE	Verb — <i>um/am/nachmittags, wann</i>
TIME_SPEC	Anbindung der Kardinalzahl an das temporale <i>um</i>
TIME_MOD	Anbindung von <i>halb</i> in Uhrzeiten
DATE_MOD	Anbindung von Datumsmodifizierungen (<i>nach Ostern</i>)
CLOCK	Anbindung von <i>Uhr</i> in Uhrzeiten
WEEK_DAY	Anbindung von Wochentagen an <i>am</i>
MONTH	Anbindungen von Monatsnamen (<i>den zweiten April</i>)

Andere im Deutschen durch präpositionale Konstruktionen ausgedrückte Kanten:

⁹bei Ortsnamen

¹⁰bei Personennamen

CAUSE	Kausalrelationen (<i>wegen, zwecks</i>)
ARG_LIMIT	<i>das ist günstig für mich</i>
BELIEVE	<i>das ist leicht für mich</i>

Kanten für periphrastische Prädikate:

ANTE	Perfektauxiliar — Partizip
POST	Futurauxiliar — Infinitiv
MODAL	Anbindung von Modalverben
PRED	Anbindung prädikativer Adjektive und Prädikatsnomen: <i>ist — PRED — schön — THEME — Karin</i> <i>ist — PRED — Lehrerin — THEME — Karin</i>
W_PASSIV	<i>werden</i> -Passiv

Modifikanden werden durch folgende Kanten benannt:

DET	Definitheitsrelation
DEM	Demonstrativrelation
MOD	attributive Adjektive, Partizipien
ADVMOD	Adverbien
GENMOD	Genitivmodifikanden (muß feiner aufgliedert werden)
AMOD	Appositionen mit Kasus Kongruenz (<i>die Tochter Karins der Lehrerin kommt</i>)
POSS	Possesivrelationen

Subordinierende Kanten:

SUBORD	daß
CAUS	Kausativ mit <i>lassen</i>
IF	— IF — <i>wenn</i> (konditional) — SUBORD — finites Verb

Diathesenbezogene Kanten:

W_PASSIV	Anbindung des Partizips an das <i>werden</i> -Passiv Auxiliar
MEDIAL	Anbindung des Infinitivs an das Medial-Auxiliar <i>lassen</i>
M_SICH	Anbindung des medialen <i>sich</i> an <i>lassen</i> bzw. das Vollverb bei Medial-Konstruktionen ohne <i>lassen</i>

Andere Kanten binden Knoten parallel zur Syntax an, ohne daß eine semantische Relevanz vorliegt:

IDENT	Identitätsrelationen (<i>Karin ist die Lehrerin</i>)
SPEC	Präposition — Dependens (Nominale, Kardinalzahlen)
ZPOS2	rechter Teil von Zirkumpositionen
EXPL	explikatives <i>es</i>
VERB_ZUSATZ	Anbindung der Verbzusätze
INFZU	Anbindung des <i>zu</i> vor Infinitiven
ADJ_AM	Anbindung des Superlativ <i>am</i>
SCOPE	Skopusrelationen
PRO_ANTEZ	Antezedens, an ein Pronomen angebunden

3 Relationen zum Aufbau atomarer syntaktischer Einheiten

Ein wesentliches Kennzeichen des gewählten Formalismus besteht darin, daß einem Knoten auch dann nicht mehr als ein orthographisches Wort zugeordnet werden kann, wenn dieses gemeinsam mit einem oder mehreren anderen eine zusammengesetzte Form eines Lexems bildet. Eine solche Belegung ist insbesondere ausgeschlossen, wenn ein diskontinuierliches Vorkommen einer Form vorliegt. Aus diesem Grunde werden — zusätzlich zu den syntaktischen und rollensemantischen Relationen — Beziehungen für die Beschreibung des internen Aufbaus bestimmter atomarer syntaktischer Einheiten angesetzt:

- die Relation zwischen einem ‚trennbaren‘ Verbzusatz und der entsprechenden Verbform (PART_V_P bzw. VERB_ZUSATZ),
- die Relation zwischen einer *zu*-Partikel und einer ‚einfachen‘ Infinitivform (ZU_V_P bzw. INFZU) sowie
- die Relation zwischen einer *am*-Partikel und einer Adjektivform zur Bildung des Superlativs (AM_ADJ_P bzw. ADJ_AM).
- die Relation zwischen dem linken und dem rechten Teil einer Zirkumposition (RIGHT_CIRC_P bzw. ZPOS2).

4 Komplementrelationen

Im Deutschen treten Komplemente¹¹ insbesondere zu Verbformen, Adjektivformen, Substantivformen, Adpositionsformen¹² und Subjunktionsformen¹³ auf. Der Name einer syntaktischen Relation, die zwischen einem Komplement und dessen Regens vorkommt, entspricht der Ordinalzahl der Argumentstelle, die durch das Komplement ausgefüllt wird. Als semantische Relation, die zwischen einem Komplement und dessen Regens vorkommt, wird die semantische Rolle angesetzt, die das Regens an das Komplement vergibt.

4.1 Verbkomplemente

4.1.1 Komplemente von Vollverbformen

Vollverbformen haben je nach den in ihrem Lexikoneintrag aufgeführten Werten der $\langle \text{syn catn} \rangle$ -Merkmale¹⁴ ein bis etwa vier nominale, adpositionale, subjunktionale, verbale oder adverbiale Komplemente: V_C1, V_C2, V_C3 und V_C4. Diese sind entweder obligatorisch oder fakultativ. Neben der syntaktischen Kategorie legt der Lexikoneintrag der Vollverbform die semantische Rolle (die Werte der $\langle \text{sem ro-len} \rangle$ -Merkmale),¹⁵ in der das Komplement zu ihm steht, sowie gegebenenfalls weitere Eigenschaften seiner Komplemente wie den Kasus (die Werte der $\langle \text{syn casen} \rangle$ -Merkmale) und die Sorte (die Werte der $\langle \text{sem sortn} \rangle$ -Merkmale)¹⁶ fest. Das Regens eines V_C1 kongruiert mit seinem Dependens in Person und Numerus. In unmarkierter Wortstellung geht das V_C1 anderen Verbkomplementen voraus. Im Gegensatz zur semantischen Beschreibungsebene haben infinite Verbformen — also auch infinite Auxiliar-, Modal-, Kopula- und Kausativverbformen — auf der syntaktischen Beschreibungsebene kein V_C1.¹⁷

¹¹Soweit dies nicht zu Ambiguitäten führt, werden im folgenden Relationsausdrücke wie „Komplement“ oder „Angabe“ stillschweigend zur Bezeichnung einer syntaktischen Relation als auch zur Bezeichnung einer Konstituente (also eines Teilbaumes) verwendet.

¹²Adpositionen (z. B. *in*) können mehrere, unterschiedlich kategorisierte Formen haben (hier: *in*, *im* und *ins*).

¹³Eine *Subjunktion* ist eine subordinierende Konjunktion.

¹⁴Der Wert des *syn*-Merkmals im Lexikoneintrag einer Form entspricht in etwa der Rektionskategorie dieser Form.

¹⁵Der Wert des *sem*-Merkmals im Lexikoneintrag einer Form enthält Angaben zu den thematischen Eigenschaften und zur sortalen Selektion dieser Form.

¹⁶Ein Komplement erfüllt die Sortenselektion seines Regens genau dann, wenn die ihm zugeordnete Sorte mit der selegierten Sorte ‚kompatibel‘ in einem bestimmten Sinn ist.

¹⁷Vgl. Abschnitte 4.1.2, 4.1.3 und 4.1.4.

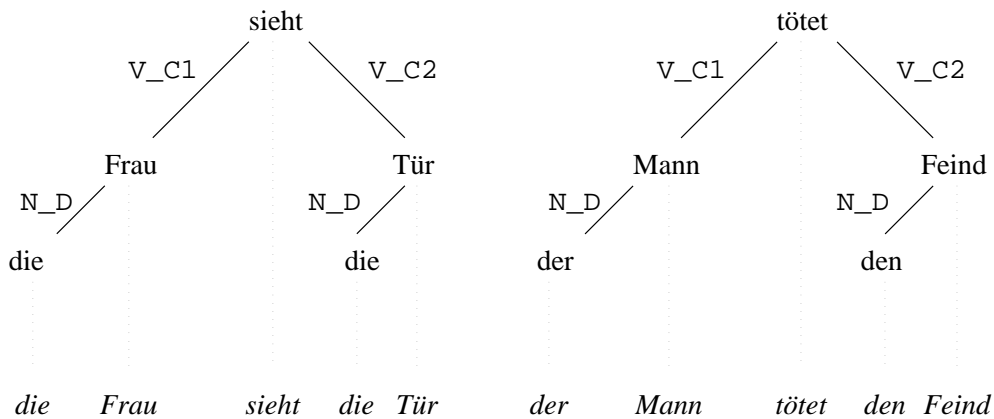


Abbildung 2: Transitive Konstruktionen (Syntax)

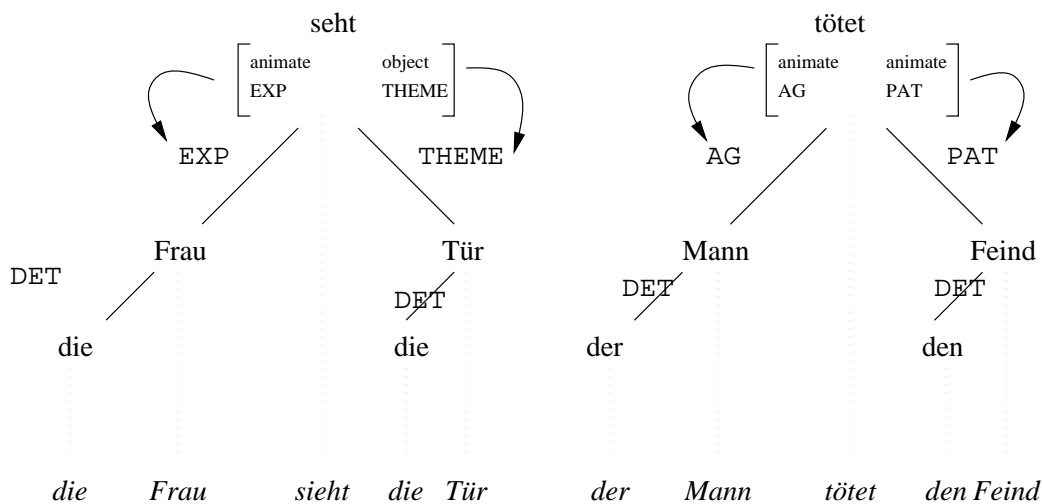


Abbildung 3: Transitive Konstruktionen (Semantik)

Die Antezedensrelation, die zwischen dem Subjekt oder dem indirekten Objekt und dem (bei Infinitiven nicht ausgedrückten) Subjekt eines verbalen oder subjunktionalen direkten Objekt sogenannter ‚Kontrollverbformen‘ besteht, wird weder auf der syntaktischen noch auf der semantischen Beschreibungsebene unmittelbar dargestellt. Stattdessen könnte sie auf einer zusätzlichen Beschreibungsebene repräsentiert werden, wo außerdem die von den jeweiligen Verbformen geforderten Eigenschaften der in einer Antezedensrelation stehenden Komplemente abzugleichen wären.

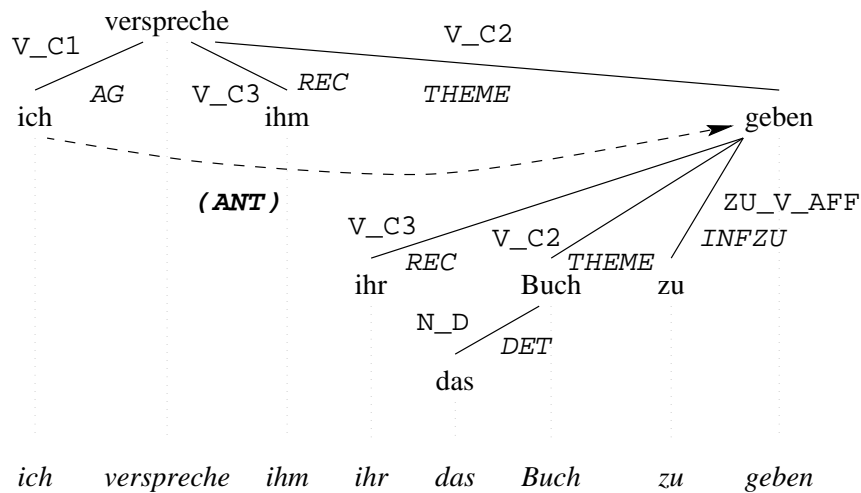


Abbildung 4: Kontrollverb-Konstruktion mit verbalem Objekt (Syntax/Semantik)

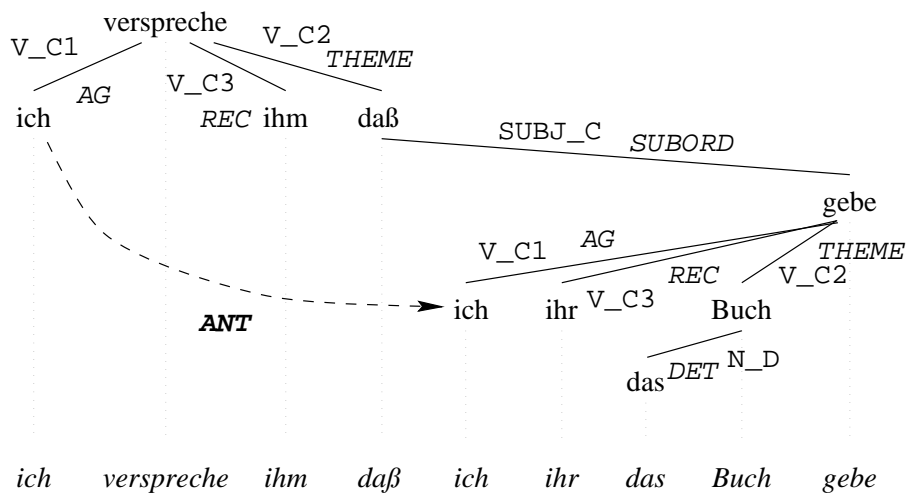


Abbildung 5: Kontrollverb-Konstruktion mit subjunktionalem Objekt (Syntax/Semantik)

Bei den sogenannten *accusativus cum infinitivo* (ACI)-Konstruktionen wird die Antezedensrelation zwischen dem direkten Objekt und dem nicht ausgedrückten Subjekt des Infinitiv-Objekts auf der syntaktischen Beschreibungsebene ebenfalls nicht repräsentiert (dafür könnte die für Kontrollverb-Konstruktionen angesetzte Beschreibungsebene verwendet werden). Auf der semantischen Beschreibungsebene hingegen führt die Rollenselektion des Infinitivs bezüglich seines ersten Komplements zum Ansetzen einer „_REV“-Rolle.

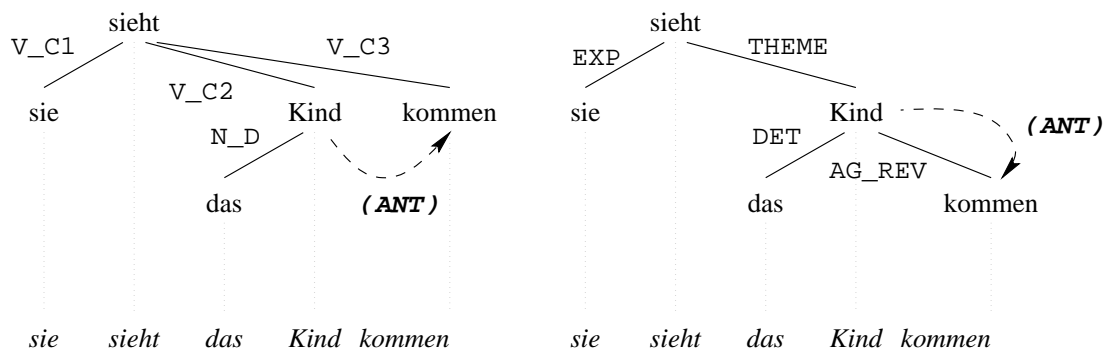


Abbildung 6: ACI-Konstruktion (Syntax und Semantik)

4.1.2 Komplemente von Auxiliar- und Modalverbformen

Auf der syntaktischen Beschreibungsebene hat eine Auxiliar- und Modalverbform neben einem ersten Komplement (dem V_C1), dessen Kasus durch den Wert des $\langle \text{syn case1} \rangle$ -Merkmals ihres Lexikoneintrags und dessen Kategorie per ‚Merkmalvererbung‘ durch den Lexikoneintrag der untergeordneten infiniten Vollverbform bestimmt werden, ein weiteres verbales Komplement (das AUX_C bzw. MOD_C), dessen syntaktische Kategorie der Wert des $\langle \text{syn catv} \rangle$ -Merkmals festlegt. Auf der semantischen Beschreibungsebene haben Auxiliar- und Modalverbformen kein erstes Komplement: diejenige Wortform, die auf der syntaktischen Beschreibungsebene erstes Komplement einer Auxiliar- oder Modalverbform ist, ist auf der semantischen Beschreibungsebene Komplement der indirekt davon abhängenden infiniten Vollverbform.¹⁸ Durch diese Analyse wird der zugrundeliegenden ‚logischen Struktur‘ Rechnung getragen. Jedoch ergeben sich bei solchen Konstruktionen auf der semantischen Beschreibungsebene häufiger nicht-projektive Strukturen als auf der syntaktischen Beschreibungsebene.¹⁹ Die semantische Rolle des zweiten Komplements (ANTE, POST oder MODAL) wird ebenfalls lexikalisch als Wert des $\langle \text{sem rolev} \rangle$ -Merkmals bestimmt.

¹⁸Vgl. Abschnitt 4.1.1.

¹⁹Projektivität wird nur auf der syntaktischen Beschreibungsebene eingefordert, da es sich dabei um ein Wortstellungsphänomen handelt. Konstruktionen bestimmten Typs bleiben davon jedoch ausgenommen (vgl. Kunze 1972, Kap. 2).

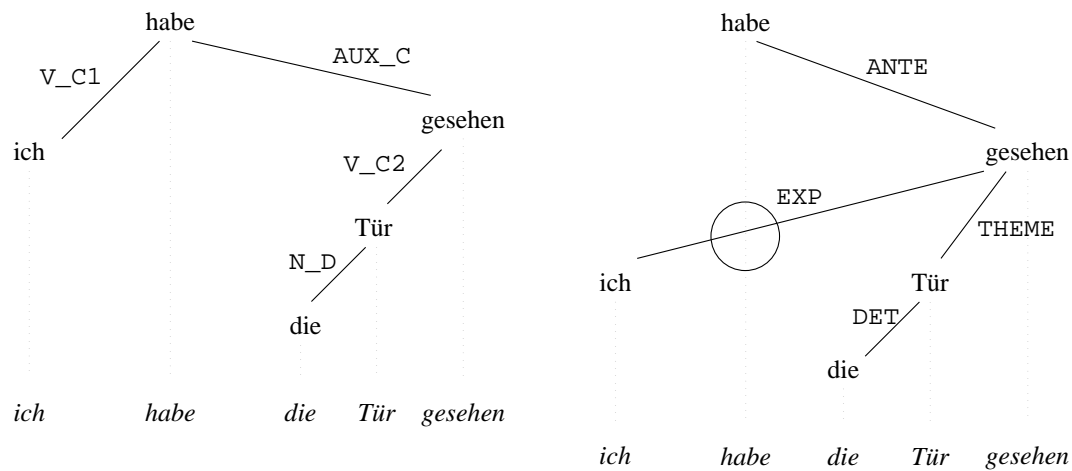


Abbildung 7: Auxiliärverb-Konstruktion (Syntax und Semantik)

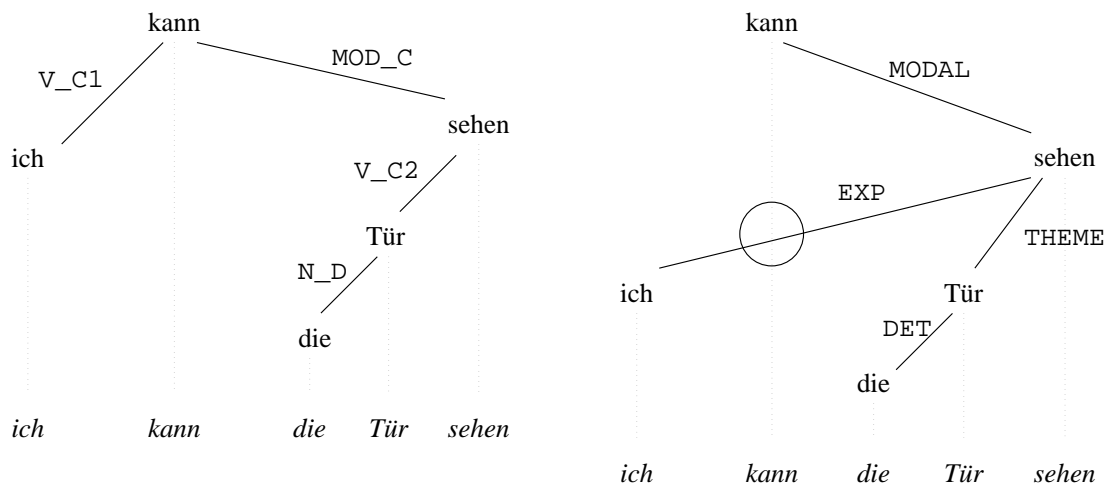


Abbildung 8: Modalverb-Konstruktion (Syntax und Semantik)

4.1.3 Komplemente von Kopulaverbformen

Strukturen mit Kopulaverbformen²⁰ sind solchen mit Auxiliär- oder Modalverbformen²¹ ähnlich. Dem AUX_C/ANTE/POST und MOD_C/MODAL dort entspricht das

²⁰Von dem Kopulaverb *sein_{cop}* werden die Vollverben *sein_{sit}* mit lokal oder temporal situierender Wortbedeutung (vorkommend beispielsweise in *die Konferenz ist am Freitag*) und *sein_{ident}* mit identifizierender Wortbedeutung (z. B. in *dies ist mein Büro*) unterschieden.

²¹Vgl. Abschnitt 4.1.2.

COP_C bzw. PRED hier, dessen syntaktische Kategorie durch den Wert des $\langle \text{syn catn} \rangle$ -Merkmals, dessen Kasus durch den Wert des $\langle \text{syn casen} \rangle$ -Merkmals und dessen semantische Rolle (trivialerweise) durch den Wert des $\langle \text{sem rolen} \rangle$ -Merkmals im Lexikoneintrag der Kopulaverbform spezifiziert wird. Für das V_C1 von Kopulaverbformen gilt dasselbe wie für das V_C1 von Auxiliar- oder Modalverbformen.

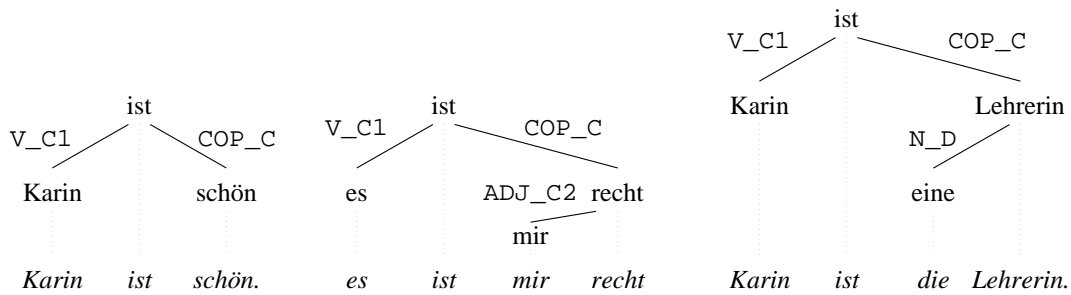


Abbildung 9: Kopulaverb-Konstruktionen (Syntax)

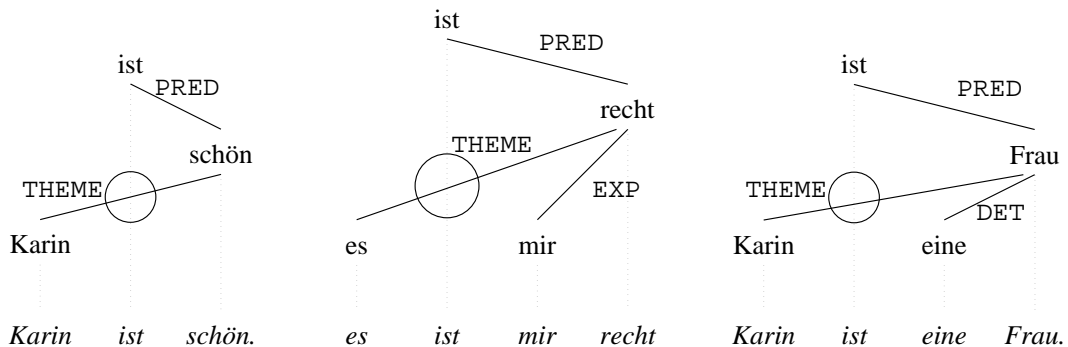


Abbildung 10: Kopulaverb-Konstruktionen (Semantik)

4.1.4 Komplemente von Kausativverbformen

Für Kausativverbformen werden bis zu drei Komplemente angesetzt: ein V_C1 bzw. AG, ein fakultatives V_C2 (auf der semantischen Beschreibungsebene wird der *causee* als Komplement des untergeordneten Infinitivs analysiert) und ein verbales CAUS_C bzw. CAUS. Die Eigenschaften des letzteren werden durch die Werte des $\langle \text{syn catv} \rangle$ - und des $\langle \text{sem rolev} \rangle$ -Merkmals im Lexikoneintrag der Kausativverbform festgelegt; sein eigener Lexikoneintrag bestimmt per ‚Merkmalvererbung‘ die Kategorie des V_C2.

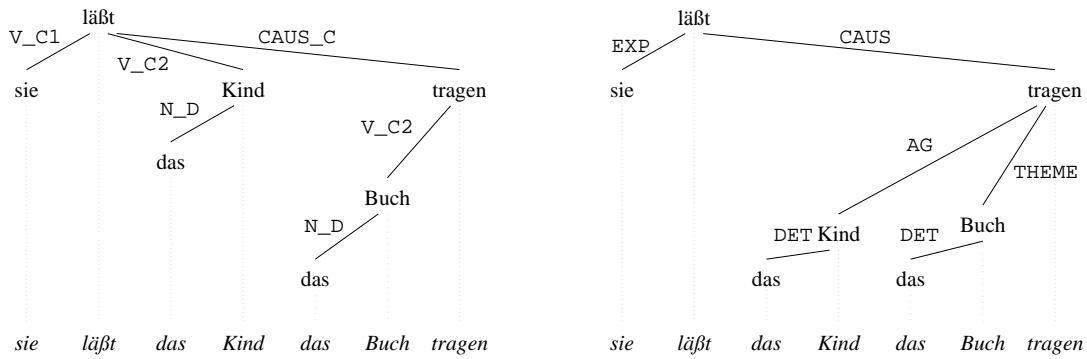


Abbildung 11: Kausativ-Konstruktion (Syntax und Semantik)

4.1.5 Komplemente von Passiv-Auxiliarverbformen

Auf der syntaktischen Beschreibungsebene haben Passiv-Auxiliarverbformen ein V_C1, ein fakultatives V_C2 (die *von*-Gruppe)²² und ein verbales PASS_C. Die Kategorie des V_C1 wird per ‚Merkmalvererbung‘ vom Lexikoneintrag des untergeordneten Partizips spezifiziert. Dieses bildet auf der semantischen Beschreibungsebene als W_PASSIV das einzige Komplement des Passiv-Auxiliars und nimmt die dem V_C1 und V_C2 entsprechenden Gruppen selbst als Komplement.

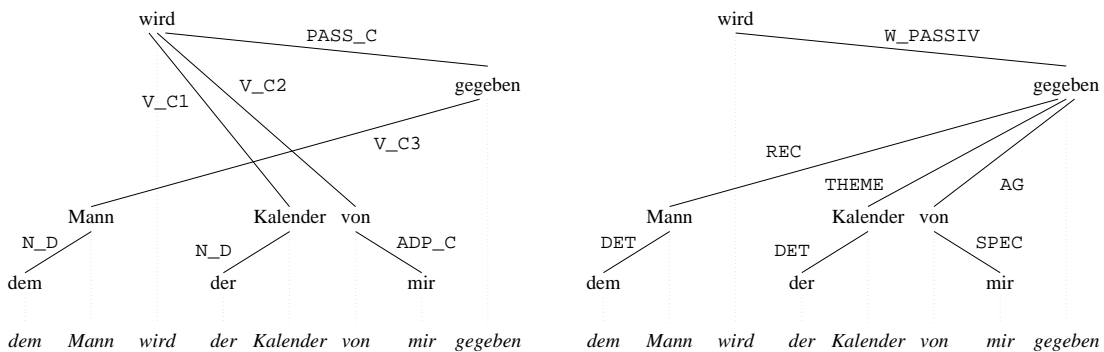


Abbildung 12: Passiv-Konstruktion (Syntax und Semantik)

²²Das potentielle Auftreten einer agentivischen *von*-Gruppe kann als Eigenschaft des Passiv-Auxiliars betrachtet werden. Dieser Ansatz hat den Vorteil, daß die Rektionskategorie einer Vollverbform in Passiv-Konstruktionen unverändert bleiben kann.

4.1.6 Komplemente von Medial-Auxiliarverbformen

Die für Medial-Konstruktionen charakteristische Reflexivpronomenform tritt als V_C2 bzw. M_SICH an eine Form des Medial-Auxiliarverbs *lassen* oder, sofern eine solche nicht vorkommt, an die entsprechende Vollverbform. Wenn eine *lassen*-Form auftritt, so hat diese außerdem ein infinitives MED_C bzw. MEDIAL und — ausschließlich auf der syntaktischen Beschreibungsebene — ein V_C1, dessen Kategorie per ‚Merkmalererbung‘ vom Lexikoneintrag des untergeordneten Infinitivs spezifiziert wird, sowie ein fakultatives V_C3 (eine *von*-Gruppe).²³

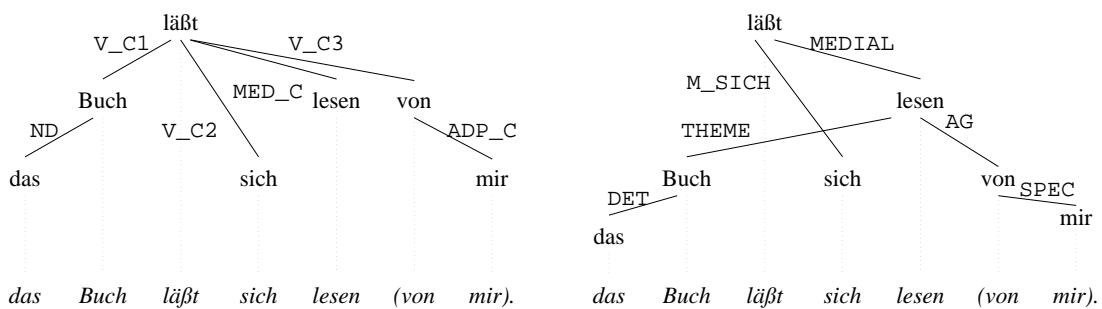


Abbildung 13: Medial-Konstruktion mit *lassen* (Syntax und Semantik)

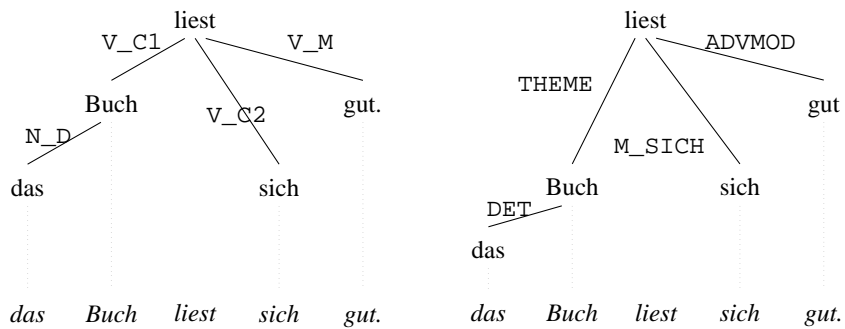


Abbildung 14: Medial-Konstruktion ohne *lassen* (Syntax und Semantik)

4.2 Adjektivkomplemente

Neben Verbformen können unter anderem auch prädikative und attributive Adjektivformen oder Partizipien Komplemente haben. Die syntaktischen Eigenschaften von

²³Das potentielle Auftreten einer agentivischen *von*-Gruppe kann als Eigenschaft des Medial-Auxiliars betrachtet werden.

ADJ_C2, ADJ_C3 und ADJ_C4²⁴ werden wie bei den Verbformen durch die Werte von Merkmalen wie $\langle \text{syn catn} \rangle$ oder $\langle \text{syn casen} \rangle$ im Lexikoneintrag der Adjektivform bzw. des Partizips festgelegt; die Werte der $\langle \text{sem rolen} \rangle$ - und $\langle \text{sem sortn} \rangle$ -Merkmale spezifizieren die semantischen Eigenschaften dieser Komplemente.

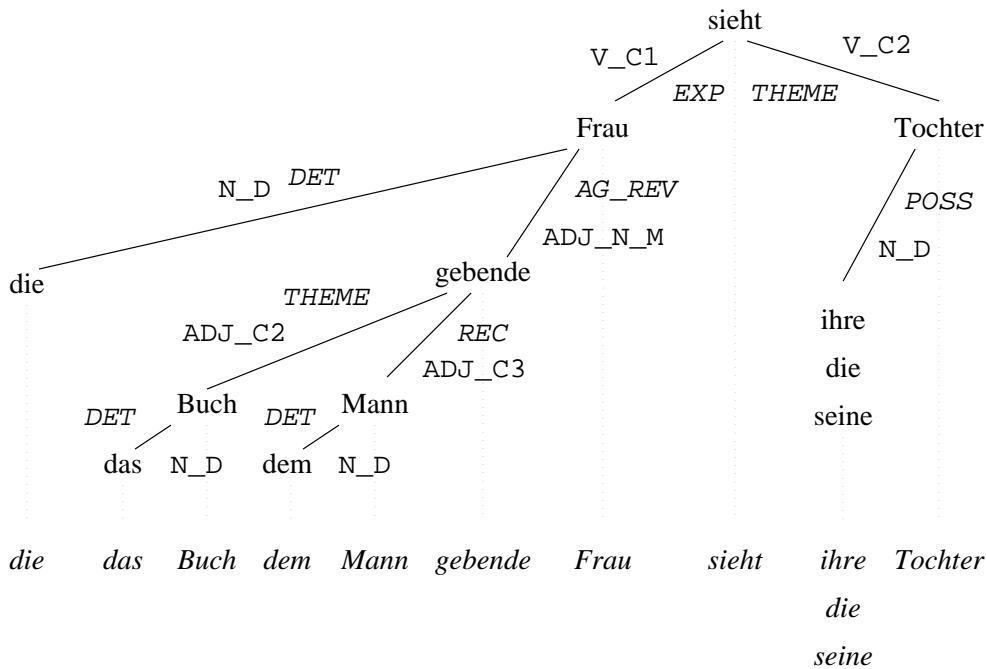


Abbildung 15: Adjektivkomplemente (Syntax/Semantik)

4.3 Substantivkomplemente

Komplemente von Substantivformen sind gegenwärtig noch nicht implementiert; ihre Erfassung scheint jedoch keine prinzipiellen Probleme aufzuwerfen.

4.4 Adpositionskomplemente

Adpositionsformen — Präpositionsformen wie *in*, Postpositionsformen wie *wegen*²⁵ und Zirkumpositionsformen wie *um willen* — haben genau ein Komplement: ein

²⁴Eine ADJ_C1-Relation — bezogen auf die erste Argumentstelle der Adjektivform bzw. des Partizips — wird nicht angesetzt, da diese mit der Attributsrelation (ADJ_N_M) bzw. mit der Prädikatsnomenrelation (COP_C) in Verbindung mit der Subjektsrelation (V_C1) zusammenfallen würde.

²⁵Es gibt auch eine Präposition *wegen*.

ADP_C bzw. SPEC.²⁶ (In bestimmten Fällen können auch die semantischen Relationen TIME_SPEC, WEEK_DAY oder MONTH auftreten.) Die syntaktische Kategorie (in der Regel Nominal²⁷) und der Kasus des Komplements wird durch den Wert des <syn cat1>- bzw. des <syn case1>-Merkmals im Lexikoneintrag der Adpositionsform festgelegt; seine Sorte spezifiziert der Wert des <sem sort1>-Merkmals.

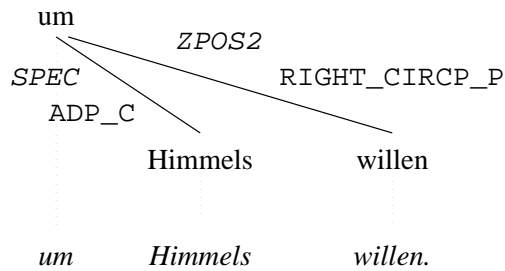


Abbildung 16: Zirkumpositionskomplement (Syntax/Semantik)

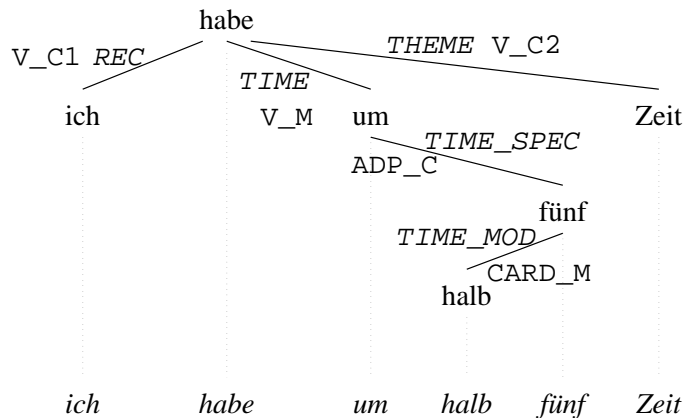


Abbildung 17: Präpositionskomplement I (Syntax/Semantik)

²⁶Zur Beziehung zwischen dem linken und rechten Teil einer Zirkumposition vgl. Abschnitt 3.

²⁷Auch Präpositionalgruppen können Komplemente von bestimmten Präpositionen sein (vgl. *bis nach Mitternacht*).

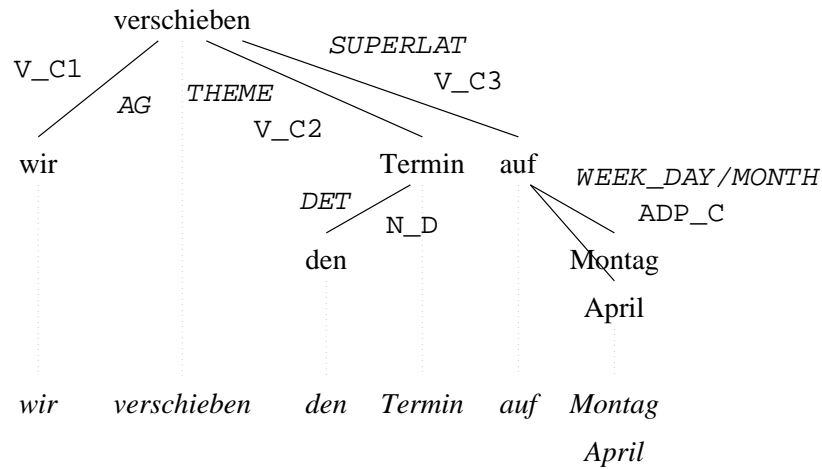


Abbildung 18: Präpositionskomplement II (Syntax/Semantik)

4.5 Subjunktionkomplemente

Wie Adpositionen haben Subjunktionen mit dem SUBJ_C bzw. SUBORD ein einziges Komplement. Dabei handelt es sich immer um finite Verbformen.

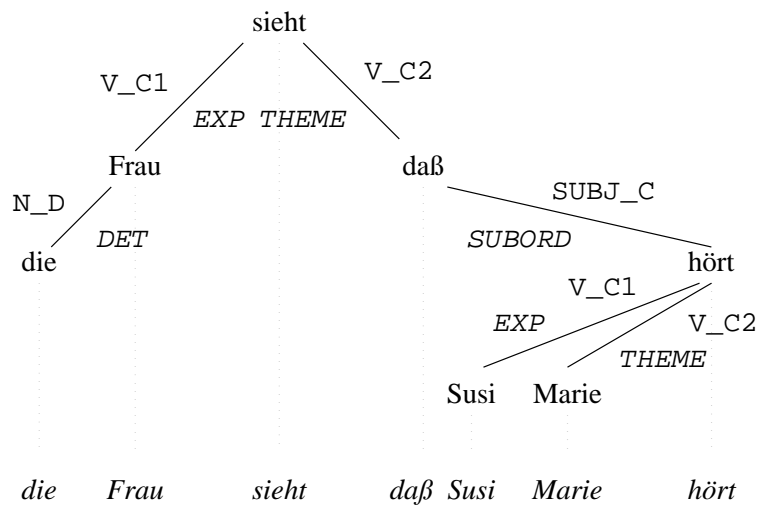


Abbildung 19: Subjunktionkomplement (Syntax/Semantik)

5 Angaberelationen

Zwei Typen von freien Angaben lassen sich im Deutschen unterscheiden: als Adverbialbestimmungen modifizieren freie Angaben Verbformen, während sie als Attribute insbesondere Substantiv-, Adjektiv- und Adverbformen²⁸ näher bestimmen.

5.1 Adverbialbestimmungen

Zwei Hauptarten von Adverbialbestimmungen sind zu unterscheiden: Adverbiale, die bestimmte Parameter eines Sachverhalts näher bestimmen (z. B. Ort und Zeit), und Adverbiale, die eine Sprechereinstellung zu einem Sachverhalt kennzeichnen (sogenannte ‚modale Satzadverbiale‘). Während Erstere Dependents der relevanten Vollverbform sind, hängen Letztere an der entsprechenden finiten Verbform; bei nicht periphrastischen Konstruktionen fallen ihre Positionen zusammen. Adverbiale des zweiten Typs sind gegenwärtig noch nicht implementiert. Der wesentliche Teil der Erfassung von Adverbialbestimmungen des ersten Typs (V_M auf der syntaktischen Beschreibungsebene) wird von der semantischen Beschreibungsebene geleistet.

Dort werden Adverbialbestimmungen lokale, temporale u. ä. Relationen zugeordnet. Die Disambiguierung kann jedoch häufig nur vorläufigen Charakter haben, da bislang ein Zugriff auf domänenspezifisches Weltwissen nicht möglich ist. Außerdem bereiten Zeitausdrücke (z. B. in Form von Uhrzeit- oder Datumsangaben) Schwierigkeiten, als ihre interne Struktur nicht unbedingt ihre semantische Struktur reflektiert.

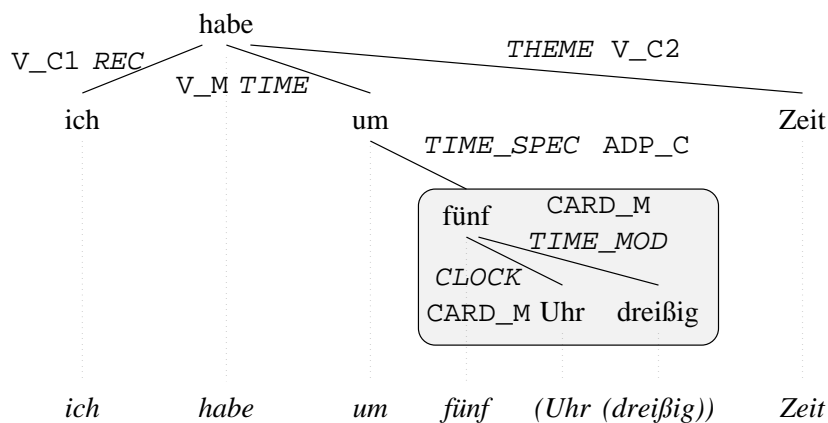


Abbildung 20: Uhrzeitangabe-Konstruktion (Syntax/Semantik)

²⁸Im Deutschen enthält das ‚uneigentliche‘ Paradigma (im Sinne von Lieb 1993) eines Adverb genau eine Form.

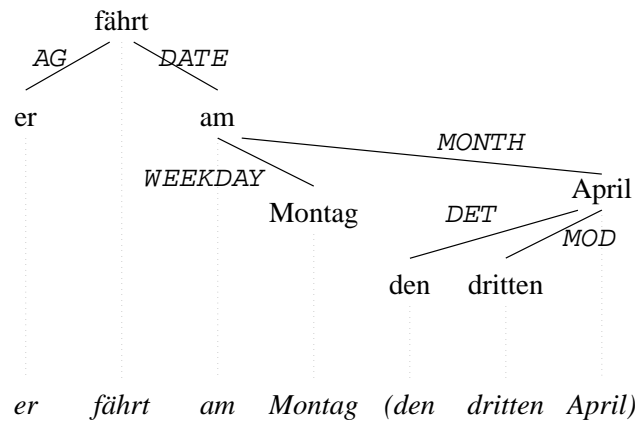


Abbildung 21: Datumsangabe-Konstruktion (Semantik)

5.2 Attribute

5.2.1 Attribute von Substantivformen

Substantivformen können durch Attribute unterschiedlicher Art modifiziert werden: durch Adjektivattribute,²⁹ Genitivattribute, Adpositionalattribute, Adverbattribute,³⁰ Appositionen, Subjunktionsattribute³¹ und Relativsätze.

Ein Adjektivattribut (ADJ_N_M auf der syntaktischen Beschreibungsebene und MOD oder bei attributiven Partizipien eine _REV-Rolle³² auf der semantischen Beschreibungsebene) kongruiert mit seinem Regens in Kasus, Numerus und Genus sowie mit dessen Determinanten in ‚Stärke‘. Es geht ihm — von archaischen Idiolekten abgesehen — ausnahmslos voraus. Attributive Partizipien selegieren außerdem die Sorte der Substantivform.

Genitivattribute (GEN_N_M bzw. GENMOD)³³ folgen normalerweise ihrem Regens. Wenn sie ihm vorausgehen, kann dieses nicht durch Artikelformen determiniert werden.

Sowohl Adpositionalattribute (ADP_N_M bzw. eine Lokalrelation, Temporalrelation etc.) als auch Adverbattribute (ADV_N_M bzw. eine Lokalrelation, Temporalrelation etc.) folgen ihrem Regens obligatorischerweise.

²⁹Darunter fallen auch partizipiale Attribute.

³⁰Z. B. *der Mann hier*.

³¹Z. B. *der Tag, als die Mauer fiel*.

³²Vgl. Abb. 15.

³³Genitivattribute sind von genitivischen Substantivkomplementen zu unterscheiden.

Appositionen (N_N_M bzw. AMOD) kongruieren mit ihrem Regens zumindest im Kasus;³⁴ ihre Sorten sind in einem bestimmten Sinn miteinander ‚kompatibel‘. Ob Appositionen immer auf ihr Regens folgen, ist unklar.

Subjunktionsattribute und Relativsätze sind noch nicht implementiert. Da das Finitum — und nicht das Relativpronomen — als Regens des Relativsatzes aufgefaßt wird, muß die zwischen dem Regens und dem Relativpronomen bestehende Antezedensrelation auf der für Kontrollverb- und ACI-Konstruktionen angesetzten zusätzlichen Beschreibungsebene repräsentiert werden.³⁵

5.2.2 Attribute von Adjektivformen

Während es sich bei Attributen von prädikativen Adjektivformen und Partizipien (ADJ_M bzw. ADVMOD) nur um Adverbformen in einem weiten Sinn handeln kann, lassen sich attributive Adjektivformen und Partizipien durch Attribute vielfältiger Art modifizieren (ADJ_M; auf der semantischen Beschreibungsebene kommen im wesentlichen dieselben Relationen wie bei Adverbialbestimmungen in Frage³⁶).

5.2.3 Attribute von Adverbformen

Attribute von Adverbformen sind gegenwärtig noch nicht implementiert.

6 Determinationsrelationen

Artikelformen und verschiedene adjektivische Pronomenformen stehen in einer Determinationsrelation (N_D bzw. DET/DEM/POSS etc.) zu einer Substantivform. Sie kongruieren mit dieser in Kasus, Numerus und Genus und gehen seinen anderen Dependents voraus. Sofern die Relation von Possessivpronomina zu ihrem Antezedenten satzintern ist, könnte diese auf der für Kontrollverb- und ACI-Konstruktionen angesetzten zusätzlichen Beschreibungsebene repräsentiert werden³⁷ — obgleich diese Identifikationen bei Pronomina der dritten Person nur mit einer Wahrscheinlichkeit < 1 zutreffen können.

³⁴Von der (‚weiten‘) Apposition ist die ‚enge‘ Apposition zu unterscheiden, die stets im Nominativ steht.

³⁵Vgl. Abschnitt 4.1.1.

³⁶Vgl. Abschnitt 5.1.

³⁷Vgl. Abschnitt 4.1.1.

7 Skopusrelationen

Wortformen wie *auch* oder *nicht* treten in eine doppelte Relation: in *das ist nicht schlecht* hat *nicht* neben einem Bereich (*das ist schlecht*) auch einen Bezug (hier *schlecht*).³⁸ Von diesen beiden Teilrelationen wird in DAWAI gegenwärtig nur die erste repräsentiert (V_SC/ADJ_SC bzw. SCOPE); die zweite ist ohne Rückgriff auf die Intonationsstruktur nicht hinreichend identifizierbar.

8 Topologische Struktur

Beschränkungen der topologischen Struktur des Deutschen sind in DAWAI erst teilweise formuliert. Dies hat seinen Grund vor allem in der Schwierigkeit, mit lokalen Constraints ‚globale‘ Eigenschaften von Sätzen zu erfassen. Dennoch lassen sich bereits Verb-Erststellung, Verb-Zweitstellung und Verb-Endstellung weitgehend beschreiben — Voraussetzung für die Bestimmbarkeit des Satzmodus.

Konstruktionen mit Vorfeld-*es* werden durch das Ansetzen einer V_EXPL- bzw. EXPL-Relation erfaßt. Für die Abdeckung von Herausstellungs-Konstruktionen ist die PRO_ANT- bzw. die PRO_ANTEZ-Relation bestimmt.

Literatur

- Kunze, J. 1972. *Die Auslaßbarkeit von Satzteilen bei koordinativen Verbindungen im Deutschen*. Nr. 14 in Schriften zur Phonetik, Sprachwissenschaft und Kommunikationsforschung. Berlin: Akademie-Verlag.
- Kunze, J. 1975. *Abhängigkeitsgrammatik*. Nr. 12 in *Studia Grammatica*. Berlin: Akademie-Verlag.
- Lieb, H.-H. 1993. Integrational Linguistics. In *Syntax: Ein internationales Handbuch zeitgenössischer Forschung/An International Handbook of Contemporary Research*, Hg. J. Jacobs, A. von Stechow, W. Sternefeld und T. Vennemann, Nr. 9 in Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 1. Berlin: de Gruyter, S. 430–468.
- Maruyama, H. 1990. Structural disambiguation with constraint propagation. In *Proceedings 28th Annual Meeting of the ACL*.
- Menzel, W. 1995. Robust processing of natural language. In *Proceedings 19th Annual German Conference on Artificial Intelligence*, Nr. 981 in Lecture Notes in Computer Science. Berlin.

³⁸Vgl. Lieb (1993).