

# **Auxiliarkonstruktionen und maschinelle Sprachverarbeitung**

Magisterarbeit im Fach Informationsverarbeitung

Bei: Prof. Dr. Jürgen Rolshoven  
Institut für Sprachliche Informationsverarbeitung  
Universität zu Köln

Vorgelegt am 1.10.2001 von:  
Jürgen Hühn  
Weißenburgstr. 14  
50670 Köln  
Tel.: 0221 7325689  
Matrikelnummer: a2669480

# Inhaltsverzeichnis

<b>1. Einleitung: Maschinelle Sprachverarbeitung und natürliche Sprache</b>	<b>1</b>
1.1 Die (klassische) Kognitionswissenschaft	1
1.2 Der Konnektionismus	1
1.3 Generative Grammatik	3
1.4 Linguistische Objektorientierung	5
1.5 Das System LPS	6
1.6 Gliederung der Arbeit	6
<b>2. Deskriptive Skizzierung des sprachlichen Phänomens Auxiliar</b>	<b>8</b>
2.1 Definition und Abgrenzung von Auxiliaren	10
2.2 Auxiliarfunktion: Träger der Finitheitsmerkmale	12
2.2.1 <i>Das topologische Feldermodell</i>	13
2.2.2 <i>Das X-bar-Schema und die syntaktische Kategorie AUX</i>	14
2.2.3 <i>Funktionale Kategorien und Auxiliare</i>	15
2.3 Auxiliare als Ausdruck verbaler Kategorien: Analytische Bildungen	18
2.3.1 <i>Tempus</i>	20
2.3.2 <i>Genus Verbi</i>	22
2.4 Auxiliare und Theta-Rollen: Auxiliarselektion im Deutschen	24
2.5 Zusammenfassung	28
<b>3. Auxiliare und Generative Grammatik</b>	<b>29</b>
3.1 Verbpositionen	30
3.1.1 <i>Die traditionellen Positionen</i>	31
3.1.2 <i>Pollocks Split-Infl-Hypothese und ihre Erweiterung durch Chomsky</i>	33
3.1.3 <i>Bowers Predication Phrase</i>	36
3.2 Verbbewegung	38
3.2.1 <i>Verbbewegung - prinzipiell</i>	38
3.2.2 <i>Verbbewegung - parametrisiert</i>	42
3.3 Feature-Checking	44
3.4 Zur Basisposition des Auxiliars	47
3.5 Zusammenfassung	48

<b>4. Auxiliare und LPS-Prolog</b>	<b>49</b>
4.1 LPS-Prolog und linguistische Objektorientierung	50
4.2 LPS-Module	56
4.2.1 <i>Das Agreement-Modul</i>	56
4.2.2 <i>Das Tempus-Modul</i>	57
4.2.3 <i>Das Kasus-Modul</i>	60
4.2.4 <i>Das Theta-Modul</i>	62
4.3 Auxiliare und das LPS-Lexikon	64
4.4 Zusammenfassung	65
<b>5. Zusammenfassung und Ausblick</b>	<b>68</b>
<b>Bibliographie</b>	<b>70</b>
<b>Anhang A: Module in LPS-Prolog</b>	<b>74</b>
<b>Anhang B: Vererbungshierarchie der lexikalischen und funktionalen Verbklassen</b>	<b>79</b>

# 1. Einleitung: Maschinelle Sprachverarbeitung und natürliche Sprache

Das Ziel der maschinellen Sprachverarbeitung ist es, natürliche Sprachen verarbeiten und generieren zu können. Um dieses Ziel zu erreichen, erscheint es zweckmäßig, eine kognitiv plausible Theorie aufzustellen, die das menschliche Sprachvermögen beschreiben und erklären kann. Wenn die menschliche Sprachfähigkeit auf ihr zugrundeliegende Prinzipien zurückgeführt werden kann und es möglich ist, diese Prinzipien zu ermitteln, so müßten diese nur mehr für die maschinelle Interpretation aufgearbeitet werden, um den Umgang der Maschine mit menschlicher Sprache zu gewährleisten.

## 1.1 Die (klassische) Kognitionswissenschaft

Zunächst ist zu klären, was unter 'kognitiv plausibel' zu verstehen ist. Die Kognitionswissenschaft ist ein relativ junger Wissenschaftszweig, der seine Entstehung Mitte der siebziger Jahre zwei grundlegenden Thesen verdankt: der Hypothese über physikalische Symbolsysteme (*physical symbol systems hypothesis* - PSSH) von Newell und Simon (1976), sowie der Computertheorie des Geistes (*computational theory of mind* - CTM) von Fodor (1975). Als Kognition wird der Bereich bezeichnet, der die Schnittstelle zwischen Reizaufnahme (Wahrnehmung) und Verhalten (Motorik) bildet. Die Art der Intervention wird als nachvollziehbare Berechnung, als symbolmanipulierender<sup>1</sup> Prozeß angenommen. Gemäß der CTM ist der menschliche Geist ein berechnendes (kognitives) Programm, welches auf der menschlichen Hardware, dem Gehirn, implementiert ist. Entscheidend ist das Programm, die Software, für die es unerheblich ist, auf welcher Art von Hardware sie implementiert wird. Die PSSH besagt, daß ein physikalisches Symbolsystem, ob nun natürlich gezeugt oder künstlich geschaffen, die notwendigen und hinreichenden Bedingungen für intelligente Handlungen, wie beispielsweise auch den Umgang mit Sprache, bereitstellt. CTM und PSSH bestimmen noch immer den *status quo* der Kognitionswissenschaft, obwohl sie sich inzwischen einiger Kritik gegenübergestellt sahen und leicht modifiziert werden mußten.

## 1.2 Der Konnektionismus

Sehr plausible Kritik kam aus der Richtung des Konnektionismus. Dessen Vertreter wiesen vor allem darauf hin, daß die PSSH vor allem anhand sequenziell arbeitender Systeme, der

---

<sup>1</sup> Symbol ist hier nicht ikonisch zu interpretieren, sondern als Zeichenträger im allgemeinen Sinn. Symbolmanipulation bzw. -verarbeitung setzt eine definierte Menge an Elementarsymbolen sowie Operatoren zu deren Transformation voraus.

sogenannten Rechner im *von Neumann*-Design, verifiziert wurde, sich aber diese Systeme in Architektur und Arbeitsweise grundlegend von einem neuronalen Netzwerk unterscheiden. Ein solches neuronales Netzwerk liegt dem menschlichen Gehirn zugrunde. Während ein *von Neumann*-Rechner neben Speicher und Ein-/Ausgabeeinheiten einen zentralen Prozessor besitzt, auf dem Anweisungen sequenziell abgearbeitet werden, besteht ein neuronales Netz aus einer großen Anzahl - das zentrale Nervensystem des Menschen besitzt  $10^{10}$  Neuronen - von Subprozessoren, die untereinander verbunden sind und Informationen massiv parallel austauschen. Wie nachgewiesen werden kann, haben konnektionistisch arbeitende Systeme bei Prozessen der Musterverarbeitung<sup>2</sup> Vorteile gegenüber sequenziell arbeitenden Systemen. In einem solchen konnektionistischen System sind Symbole nicht länger Bestandteile der kognitiven Ebene, sondern Abstraktionen von dieser; die eigentliche kognitive Arbeit vollzieht sich auf subsymbolischer Ebene:

Inferenzen und die Ableitung von Wissen sind die Interaktion von einer großen Anzahl von Prozessoren (Units). Diese Interaktion erlaubt keine exakte Beschreibung auf konzeptueller Ebene, sondern muß direkt durch die Modellprozessoren verwirklicht werden.

(Smollensky, zitiert nach Dorffner, 1991:146)

Die Schwierigkeit des konnektionistischen Ansatzes liegt auf der Hand: Durch das Fehlen der konzeptionellen Ebene können kognitive Vorgänge nicht beschrieben, sondern nur nachgeahmt werden, was bei Untersuchungsobjekten von der Größe des menschlichen Gehirns nicht durchführbar ist und möglicherweise auch nie sein wird. Stattdessen beschränkt sich die konnektionistische Forschung darauf, kleinere neuronale Netzwerke auf *von Neumann*-Rechnern zu simulieren, zu beobachten und zu modifizieren. Rein subsymbolisch arbeitende Systeme sind dabei eher selten, hierzu zählt wohl als bekanntestes das oft zitierte *past tense*-Modell (Rummelhard/McClelland 1986b). Doch finden sich in modernen KI-Systemen häufig konnektionistische Komponenten (z.B. in der Bildbearbeitung, aber auch beim Parsing menschlicher Sprache, s.a. Jain & Waibel 1991); diese arbeiten jedoch mit symbolischen Repräsentationen, auch um mit den nicht-konnektionistischen Modulen zu interagieren.

Der konnektionistische Ansatz wurde hier vor allem wegen seiner Plausibilität dargestellt; in der vorliegenden Arbeit - wie auch in einem großen Teil der modernen Kognitionswissenschaft - wird aus Gründen der Praktikabilität auf der Grundlage von PSSH und CTM gearbeitet.

---

<sup>2</sup> Diese umfaßt Mustererkennung, -klassifikation und -vervollständigung.

### 1.3 Generative Grammatik

Ein zentrales Teilgebiet der Kognitionswissenschaft stellt die von Chomsky in den fünfziger Jahren entwickelte allgemeine Sprachtheorie, die sogenannte *Generative Grammatik*, dar. Chomsky wandte sich mit ihr gegen den Behaviorismus, das bis dahin herrschende Paradigma der Sprachphilosophie. Dessen methodologisches Ziel war die exakte Beobachtung und Beschreibung sprachlicher Äußerungen, ohne auf das Problem der mentalen Grundlage des menschlichen Sprachvermögens einzugehen. Chomsky wies nach, daß das vom Behaviorismus postulierte *Lernen durch Konditionierung* die Kreativität der menschlichen Sprache nicht erklären kann. Prinzipiell ist ein Sprecher in der Lage, eine unendlich große Anzahl von Sätzen zu produzieren, von denen er naturgemäß einen Großteil noch nie gehört haben kann. Zudem ist es nicht möglich, aus den durch die Erfahrung gewonnenen Sprachdaten durch Induktion die nötigen Generalisierungen über grammatische Regularitäten zu treffen (vgl. Fodor 1975), das bekannte *logische Problem des Spracherwerbs*. Eine Sprachtheorie muß in der Lage sein, dieses Problem zu lösen; sie muß das sogenannte *Lernbarkeitskriterium* erfüllen.

Der Sprecher muß also über sprachliches Wissen verfügen, welches er nicht allein aus der Erfahrung gewonnen haben kann. Dieses Wissen bezeichnet Chomsky (1957) als *Kompetenz*; diese ermöglicht dem (idealisierten) Sprecher einer Sprache ebendiese zu benutzen, mithin zwischen wohlgeformten und nicht wohlgeformten Sätzen zu unterscheiden. Grundlage der Kompetenz bildet die Universalgrammatik (UG), ein genetisch determiniertes, autonomes Modul der menschlichen Kognition, das aus einem Inventar von Wohlgeformtheitsregeln besteht. Durch Konfrontation mit Sprachdaten ist der Mensch im Zuge des Spracherwerbs in der Lage, die Kompetenz, seine Muttersprache verstehen und sprechen zu können, auszubilden. Die genetische Determinierung schließt ein, daß jeder Mensch auf die gleiche UG zugreift, was bedeutet, daß alle menschlichen Sprachen letztlich auf die gleichen Wohlgeformtheitsregeln zurückzuführen sein müssen. Eine solche Annahme war hinsichtlich der Tatsache, daß Sprachen sehr unterschiedliche Strukturen haben, nicht haltbar. Aufgrund dessen postuliert die jüngere Generative Grammatik das *Prinzipien- und Parametermodell* (PP-Modell; vgl. Chomsky 1986). Demnach sind *Prinzipien* universell gültig, was beispielsweise für das von Jackendoff (1977) ausgearbeitete X-bar-Schema, eine Generalisierung der Phrasenstruktur, gilt. U.a. besagt dieses Schema, daß alle Phrasen einen Kopf haben, welcher der gleichen Kategorie angehört wie die Phrase selbst (*Prinzip der Endozentrität*). *Parameter* hingegen sind zwar auch universelle Attribute, ihr Wert wird jedoch einzelsprachlich gesetzt. Ein Beispiel für einen Parameter ist die phraseninterne

Kopfstellung: Köpfe können Komplemente nach rechts oder links regieren. Deutsche Verben regieren ihr Komplement (das Objekt) nach links<sup>3</sup>, englische Verben nach rechts:

(1.1) ..., weil er [[den Mann]<sub>NP</sub> sah<sub>V<sup>o</sup></sub>]<sub>VP</sub>

(1.2) ..., because he [saw<sub>V<sup>o</sup></sub> [the man]<sub>NP</sub>]<sub>VP</sub>

Die Chomskysche Kompetenz liefert ein Berechnungsmodell, mit dem Zeichenketten hinsichtlich ihrer Zugehörigkeit zu einer Sprache sowie bezüglich ihrer Wohlgeformtheit beurteilt werden können. Die Komponenten der Sprache, welche die mögliche Kombinatorik und die möglichen Abhängigkeitsbeziehungen von Inventarelementen dieser Sprache erfaßt, wird als Syntax bezeichnet. Die Syntax wird von der Generativen Grammatik als eigenständiges, spezifisch menschliches Modul der Kognition aufgefaßt.

Im Laufe ihrer seit Chomsky (1957) währenden Entwicklung wurde die generative Grammatik mehrfach modifiziert; in ihr wurden mehrere Modelle entwickelt, die sich vor allem in der Art der Formulierung ihrer grammatischen Prinzipien unterscheiden. Das sogenannte *Standardmodell* (STM) ging strikt derivationell vor: Die Idee war, daß jeder grammatische Satz aus Regeln abgeleitet werden kann. Auf der Ebene der Syntax sind dafür zwei Regeltypen notwendig: einerseits *Phrasenstrukturregeln*, aus denen die sogenannte *Tiefenstruktur* gebildet wird, andererseits *Transformationsregeln*, die aus einer zugrundeliegenden Tiefenstruktur verschiedene *Oberflächenstrukturen* ableiten. Das Problem dieses Modells war, daß die Transformationsregeln nicht restriktiv genug formuliert werden konnten, als daß durch ihre Anwendung ungrammatische Sätze hätten ausgeschlossen werden können. Zur Lösung des Problems wurde angenommen, daß Transformationsregeln in bestimmten Kontexten nur in bestimmten Abfolgen angewendet werden können, was aber eine starke Zunahme der zu lernenden Information zur Folge hatte, womit die Theorie ihren Anspruch, das Lernbarkeitskriterium zu erfüllen, verlor. Auch aus deskriptiver Perspektive war das STM unzulänglich, da bestimmte sprachliche Phänomene (beispielsweise die Bindung, vgl. 3.2) vom derivationellen Regelsystem nicht erfaßt werden konnten.

Ab Ende der siebziger Jahre wurde deshalb von der Generativen Grammatik ein repräsentationelles Modell entwickelt, das statt auf der derivationellen Regableitung auf der Überprüfung der Zulässigkeit von bestehenden syntaktischen Strukturen basiert (vgl. Chomsky 1981). Um diese Überprüfung vornehmen zu können, bedurfte es der Einführung restriktiver Prinzipien, die grammatischen Erzeugungsregeln Beschränkungen auferlegen, damit keine ungrammatischen Strukturen erzeugt werden können. Dieses Modell erlaubt

---

<sup>3</sup> Die für das Deutsche angenommene Ausgangsstruktur mit Verbendstellung findet sich auf der Oberflächenstruktur (s.u.) nur im Nebensatz. Zur Begründung sei hier auf (3.1) verwiesen.

stärkere Generalisierungen, zudem können umfangreichere empirische Daten erfaßt werden. Die Prinzipien sind in einzelnen Modulen zusammengefaßt, zwei dieser Module gaben der Theorie ihren Namen Rektions- und Bindungs-Theorie (*Theory of Government & Binding* - GB; Chomsky 1981). Die GB-Theorie basiert auf dem PP-Modell.<sup>4</sup> Auf diesem baut auch die neueste Weiterentwicklung der Generativen Grammatik, das Minimalistische Programm (*Minimalist Program* - MP; Chomsky 1995), auf. Seine Entwicklung begann Anfang der neunziger Jahre. Es ist noch nicht voll ausformuliert und nicht in dem Maße empirisch überprüft wie die GB-Theorie, welche als grundlegende Theorie für diese Arbeit ausgewählt wurde. Dort, wo sie brauchbar erscheinen, werden jedoch auch Entwicklungen, die im Rahmen des MP gemacht wurden, mit in die Analyse einfließen.

#### **1.4 Linguistische Objektorientierung**

Das Programm der Generativen Grammatik und damit auch die GB-Theorie haben Anspruch auf kognitive Plausibilität; gemäß CTM und PSSH muß eine Umsetzung auf einer anderen Hardware als dem menschlichen Gehirn möglich sein. Dazu bedarf es zunächst einer Einigung über die Form, in der die linguistischen Hypothesen der Generativen Grammatik codiert werden können. Eine solche Form liefert u.a. auch die von Rolshoven (1987) entwickelte, modular konzipierte, deklarative, objektorientierte linguistische Programmiersprache LPS-Prolog.<sup>5</sup> Im Vergleich zu anderen Formalisierungen der GB-Theorie, wie z.B. der auch deklarativen Umsetzung von Stabler (1992), wurde LPS-Prolog eigens für Linguisten konzipiert und ist nah an der linguistischen Notation angelehnt. Daraus ergeben sich zwei Vorteile: Zum einen ist LPS-Prolog für in Programmierung ungeübte Linguisten weniger gewöhnungsbedürftig und daher leichter zu erlernen. Zum anderen wird das linguistische Wissen nicht durch für die sprachwissenschaftliche Analyse unnötigen Programmcode verschleiert. Vor allem aber führt LPS-Prolog das Konzept der Objektorientierung in die linguistische Beschreibung ein. Dies ermöglicht, das sprachliche Wissen kompakt und transparent zu formulieren (vgl. 4.1).<sup>6</sup>

---

<sup>4</sup> Wissenschaftstheoretisch betrachtet ist die Nomenklatur hier etwas kontraintuitiv: Gemeinhin stellt die Theorie den zugrundeliegenden Diskursraum, in dem verschiedene Modelle entwickelt werden können. So müßte man korrekterweise von der Prinzipien- und Parameter-*Theorie* und dem Rektions- und Bindungs-*Modell* sprechen.

<sup>5</sup> LPS-Prolog ist in seiner ursprünglichen Form ein Prolog-Dialekt (vgl. Rolshoven 1987; für eine Einführung in Prolog vgl. Clocksin & Mellish 1987). Inzwischen wurde es dem objektorientierten Paradigma gemäß weiterentwickelt (s.u.).

<sup>6</sup> Objektorientierung ist ursprünglich ein überaus erfolgreiches Konzept moderner Programmiersprachen (vgl. Mössenböck 1998). Es ist jedoch auf verschiedenste andere Bereiche übertragbar, u.a. auch auf die Organisation der linguistischen Beschreibung (vgl. Rolshoven 2001).

Die Module der GB-Theorie können als (aus Klassen bestehende) Module in LPS-Prolog umgesetzt werden. Rolshoven weist darauf hin, daß diese theoretische Fundierung zwar eine notwendige, keinesfalls jedoch hinreichende Bedingung für erfolgreiches computerlinguistisches Arbeiten darstellt:

Sie ist deshalb nicht hinreichend, weil

1. sprachwissenschaftliche Theorien sprachliches Wissen ausdrücken, nicht aber Sprachverwendungswissen. Sprachverwendungswissen gibt an, wie sprachliches Wissen - Wissen über sprachliche Systeme - in der Rede [parole] anzuwenden ist.
2. gut ausgearbeitete sprachwissenschaftliche Theorien sind im Hinblick auf ein Ziel, z.B. Beschreibung von Sprachkompetenz, nicht aber im Hinblick auf Anwendung formuliert, wie unter 1. beschrieben. Anwendung aber erzwingt die Formulierung beispielsweise morphologischer oder orthographischer Prozesse, die Definition eines Lexikonformats und die Erstellung eines Lexikons, im Hinblick auf automatische Übersetzung die Formulierung von Transferregeln für die Übersetzung aus einer natürlichen oder formalen Quellsprache in eine natürliche oder formale Zielsprache (bei formalen Sprachen ist z.B. an Datenbankabfragesprachen zu denken).“

Rolshoven (1991:138f)

Um das in LPS-Prolog codierte linguistische Wissen anwenden zu können, sind eine Einbettung des Moduls GB in ein Programm notwendig, das weitere Module [Morphophonologie, Orthographie, Lexikon, Semantik] bereitstellt, sowie ein Interpreter, der dieses Programm auf einem Rechner ausführen kann. Beides stellt das von Rolshoven entwickelte LPS-System.<sup>7</sup>

## 1.5 Das System LPS

Das LPS-System eignet sich für die angewandte computerlinguistische Arbeit in verschiedenen Bereichen, die momentane Entwicklung zielt v.a. auf maschinelle Übersetzung ab. Module des LPS-Programms werden, sofern sie relevant für das bearbeitete Thema sind, in ihrer Struktur und Funktion unter Kapitel (4.1) angeführt.

Die Beschreibung der Implementation des LPS-Programms auf dem Rechner würde den Rahmen dieser Arbeit bei weitem sprengen und ist aus linguistischer Sicht auch unnötig. Zur Dokumentation und Entwicklung des LPS-Systems sei vor allem auf Rolshoven (1986, 1991, 1996, 2001) verwiesen.

---

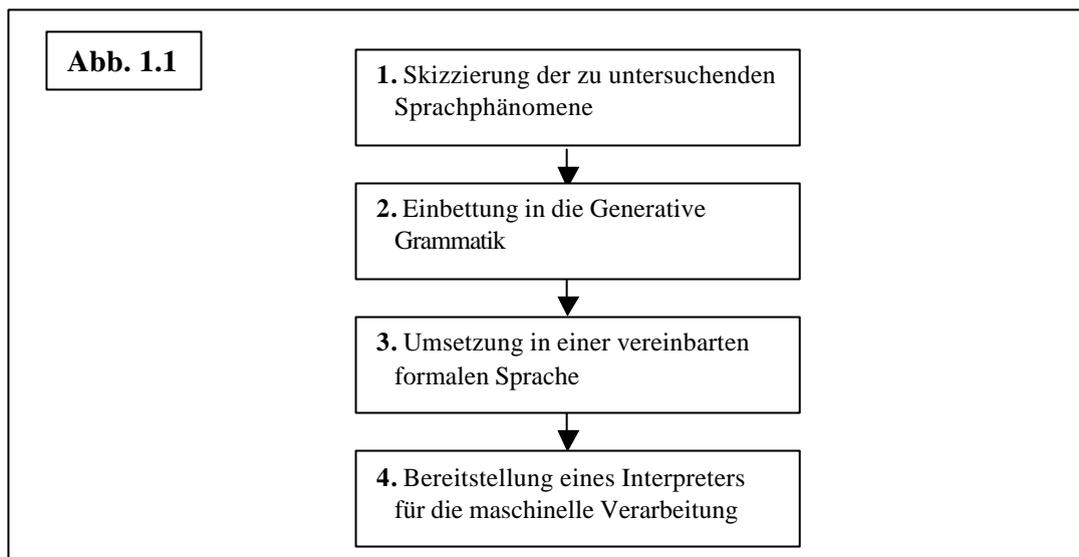
<sup>7</sup> Als eine weitere Anwendungsmöglichkeit für LPS-Prolog-codiertes Wissen ist das von Lalande (1997) entwickelte GBX-System zur Überprüfung linguistischer Hypothesen zu nennen, welches auch für die Erstellung aller in dieser Arbeit aufgeführten Strukturbäume genutzt wurde.

Dennoch soll hier kurz erwähnt werden, welche der in der Programmiersprache *Blackbox Component Pascal*<sup>8</sup> realisierten Module der Interpreter bereitstellen muß, um das theoretisch gewonnene und formal umgesetzte linguistische Wissen anwenden zu können, etwa zum Zwecke der maschinellen Übersetzung.

Maschinelle Übersetzung besteht in der Regel mindestens aus zwei Teilen: Einem Parsingprozeß über die Eingabekette der Ausgangssprache und einem Generierungsprozeß der zielsprachlichen Ausgabekette. Strukturell kann die Generierung als eine Umkehrung des Parsing-Prozesses gesehen werden, weshalb beide Aufgaben von einem einzigen Modul übernommen werden können: dem Parser/Generator (P/G) des Systems. Diesen stellt ein modifizierter *active-chart-parser* (vgl. Winograd 1983), der im Falle einer mehrdeutigen Eingabekette unterschiedliche Strukturen erzeugt und in der Lage ist, diese zu verwalten. Gesteuert wird der P/G durch das objektorientiert in Klassen (vgl. 4.1) in LPS-Prolog gespeicherte linguistische Wissen, sowie durch lexikalische Information aus dem LPS-Lexikon. Die Regeln des linguistischen Wissens müssen wie die lexikalische Information zur Verarbeitung in geeignete Datenstrukturen überführt werden. Zuständig hierfür sind zwei weitere Module: der Lexikoncompiler und der Regelcompiler. Schließlich bedarf es noch einer Inferenzmaschine, die den compilierten Code interpretiert und abarbeitet.

## 1.6 Gliederung der Arbeit

Die Stationen, die zwischen der menschlichen Sprachfähigkeit und ihrer Modellierung auf einem künstlichen Berechnungssystem liegen, werden in *Abb. 1.1* dargestellt:



<sup>8</sup> Die Programmierung von nichtlinguistischen Modulen in *Blackbox Component Pascal* (ehemals in MODULA-2) wurde vor allem aus Effizienz- und Portabilitätsgründen einem reinen PROLOG-System vorgezogen (vgl. Rolshoven 1991:157).

In dieser Arbeit werden die ersten drei Schritte aus *Abb. 1.1* jeweils in einem eigenen Kapitel behandelt: Kapitel (2) setzt sich weitestgehend deskriptiv mit Auxiliarkonstruktionen auseinander. Die gewonnenen Erkenntnisse werden in Kapitel (3) im Rahmen der GB-Theorie interpretiert. Kapitel (4) führt das Konzept der Objektorientierung in die linguistische Beschreibung ein und formalisiert, basierend auf diesem Konzept, die aus den vorangehenden Kapiteln ermittelten Erkenntnisse in LPS-Prolog. Das formalisierte Wissen kann vom LPS-System interpretiert und für computerlinguistische Anwendung genutzt werden.

## 2. Deskriptive Skizzierung des sprachlichen Phänomens Auxiliar

Gemeinhin wird der englische Terminus *auxiliary* mit *Hilfsverb* ins Deutsche übersetzt. Doch auch was im Deutschen als *Kopulativverb* bezeichnet wird, trägt im Englischen die Bezeichnung *auxiliary*. Um der zum Großteil in englischer Sprache verfaßten linguistischen Literatur Rechnung zu tragen, wird in dieser Arbeit der eingedeutschte Terminus *Auxiliar* verwendet, den es im Folgenden noch zu definieren gilt. Diese Arbeit wird sich vorwiegend mit dem deutschen Auxiliar befassen; da ihr Resultat jedoch in ein maschinelles Übersetzungssystem eingebunden werden soll, wird vereinzelt auch auf andere (europäische) Sprachen eingegangen, vor allem, wenn deren Auxiliarverhalten vom deutschen abweicht. Unter (4.1) wird eine objektorientierte linguistische Sichtweise eingeführt, welche eine Modellierung ermöglicht, die sowohl einzelsprachliche als auch sprachübergreifende Aspekte erfaßt.

In der Literatur findet sich eine Fülle von widersprüchlichen Aussagen hinsichtlich des Status und der Eigenschaften von Auxiliaren, sowie der Zugehörigkeit sprachlicher Ausdrücke zu dieser Klasse. Gängiger Grundtenor ist, daß Auxiliare nur über eine abgeschwächte lexikalische Bedeutung verfügen und vor allem eine grammatische Funktion ausüben. Sie treten typischerweise als Exponenten morphologischer Kategorien auf, wie im Deutschen *sein* und *haben* bei der Bildung analytischer Vergangenheitstempora (im Englischen lediglich *have*) oder *werden* bei der Bildung des deutschen, *be* bei der Bildung des englischen Passivs:

(2-1a) Barry *ist* in die Kneipe gegangen.

(2-1b) Barry *has* gone into the pub.

(2-2a) Dick *hat* eine Oper gehört.

(2-2b) Dick *has* been listening to an opera.

(2-3a) Die Ausstellung *wurde* um 5 Uhr geschlossen.

(2-3b) The exhibition *was* closed at 5 o'clock.

Die Beispiele zeigen, daß die Auxiliare mit einer infiniten Form eines Vollverbs auftreten, hier mit dem Partizip II.

Überträgt man die Perfekt-Sätze ins Präteritum und wandelt man den passiven in einen aktiven Satz, so entfallen die Auxiliare, ohne daß sich Wesentliches an der Bedeutung des Satzes ändert.<sup>9</sup> Dies bestätigt ihre abgeschwächte lexikalische Bedeutung:

(2-4a) Barry ging in die Kneipe.

(2-4b) Barry went to the pub.

(2-5a) Dick hörte eine Oper.

(2-5b) Dick listened to an opera.

(2-6a) X beendete die Ausstellung um 5 Uhr.

(2-6b) X closed the exhibition at 5 o'clock.

Wird jedoch das Vollverb aus den Sätzen getilgt, ergibt sich im Fall der Passivkonstruktion ein ungrammatischer Satz (2-9), ansonsten eine Ellipse (2-7a) oder eine Bedeutungsänderung (2-8):

(2-7a) Barry *ist* in die Kneipe (gegangen/verschwunden/etc).

(2-7b) Barry \**has* / *is* to the pub.

(2-8a) Dick *hat* Musik.

(2-8b) Dick *has* (got) music.

(2-9a) \*Die Ausstellung *wurde* um 5 Uhr.

(2-9b) \*The exhibition *was* at 5 o'clock.

(2-8) zeigt außerdem, daß *haben* (wie auch die anderen deutschen Hilfsverben) ohne infinites Vollverb auftreten und mithin selbst als Vollverb gebraucht werden kann. Der Vollverbgebrauch von *sein* beschränkt sich auf Prädikationen, in denen dem Subjekt spezifische Eigenschaften zugeschrieben werden.<sup>10</sup> *Sein* wird hier auch Kopulativverb genannt:

(2-10a) Der Mond *ist* klein.

(2-10b) The moon *is* small.

(2-11a) Das *ist* der Mann, der gestern ausgebrochen ist.

(2-11b) This *is* the man who broke out yesterday.

---

<sup>9</sup> Inwieweit sich die Aussagen doch ändern und v.a. weshalb in (2-6) ein weiteres Argument auftritt, wird in (2.4) dargelegt.

<sup>10</sup> Ausgenommen davon sind Konstruktionen, in denen *sein* als klassisches Vollverb gebraucht wird, wie etwa in „Ich denke, also bin ich.“

Nach dieser kurzen Darstellung der Auxiliарverwendung soll dieses sprachliche Phänomen strukturell untersucht werden. Der nächste Abschnitt (2.1) grenzt den Auxiliарbegriff inhaltlich ab und führt die Eigenschaften auf, welche den Auxiliaren zugerechnet werden. Auf drei dieser Eigenschaften wird in den darauf folgenden Abschnitten näher eingegangen: (2.2) setzt sich mit der Stellung der Auxiliare im Satz auseinander, (2.3) mit dem Auxiliar als Informationsträger, (2.4) schließlich untersucht die Auxiliarselektion und die mit ihr zusammenhängenden syntaktischen Phänomene.

## 2.1 Definition und Abgrenzung von Auxiliaren

Zunächst soll das Auxiliar von anderen lexikalischen Klassen abgegrenzt werden. Auxiliare weisen die gleichen morphosyntaktischen Eigenschaften wie Verben auf, so daß man versucht ist, sie als eine Unterklasse der Verben anzusehen (*Verb-Hypothese*, vgl. Heine 1993:8f). Dafür spricht auch, daß sie in vielen Sprachen als Hauptverb auftreten können:

(2-12) Der Mann hat viele Kinder.

(2-13) Allison ist unsere letzte Hoffnung.

Andererseits haben sie eine Fülle von Eigenschaften, die sie von Verben unterscheiden. So bilden sie keine Imperative, können nicht passiviert, nominalisiert oder direkt verneint werden; sie tragen keinen Satzakzent und können nicht das semantische Prädikat des Satzes stellen, da sie keine Bedeutung an sich allein tragen, sondern nur bei der Interpretation der Zusammenhänge helfen, in dem das Vollverb steht. Deshalb gehen viele Linguisten davon aus, daß Auxiliare keine Unterklasse der Verben bilden, sondern als eigenständige Kategorie zu betrachten sind (*Autonomitäts-Hypothese*, vgl. Heine 1993:8).

Auxiliare bilden im Gegensatz zu Verben eine geschlossene Klasse, im Deutschen und Englischen umfaßt diese eine relativ kleine Anzahl von Elementen:

(2-14) (A) Deutsch (nach DUDEN 1998:92):

(Hilfsverben:) haben, sein, werden

(Modalverben:) dürfen, können, mögen, müssen, sollen, wollen, (brauchen)

(B) Englisch (nach Steele 1999:50):

(Auxiliaries:) can, could, shall, should, will, would, may, might, must, ought (to),  
need (*jeweils mit verneinter Form*),  
be, have, do (*jeweils mit verneinten und flektierten Formen*),  
dare, used to, not.<sup>11</sup>

---

<sup>11</sup> Daß englische Modalverben nicht flektieren, spielt eine Rolle für die für sie angenommene Basisposition (vgl. 3.4). Auf *not* als Auxiliar geht Steele nicht näher ein, in dieser Arbeit wird *not* auch nicht als Auxiliar betrachtet.

Wie aus der Aufstellung (2-14) ersichtlich, besteht zwischensprachlich keine Einigkeit darüber, welche Elemente zu den Auxiliaren zu zählen sind. Weitgehend klar ist die Zuordnung von tempus- und passivbildenden Hilfsverben und den Modalverben (im DUDEN auch modale Hilfsverben genannt) zu dieser Klasse. Bisweilen wird auch der Infinitivmarker *zu* (*to* im Englischen) zur Klasse der Auxiliare gezählt (Gazdar, Klein, Pullam, Sag 1985); die radikalste Version der Auxiliardefinition ist wohl die von Hartmann und Stork (1972), die zu den Auxiliaren alle Wörter zählen, die keine eigenständige Funktion haben und ausschließlich in Kombination mit anderen Wörtern verwendet werden können. Heine (1993:22) faßt die Klasse der Auxiliare mit der vorläufigen Definition etwas enger:

A 'good auxiliary' is verblike to some extent and is used either to place the situation described in the sentence with reference to deictic time (tense), to ascribe a temporal contour to it (aspect), or to assess its reality (modality) (cf. Steele 1978:11)

Als zentrale Eigenschaft des Auxiliars sieht Heine, neben dessen Verbähnlichkeit, die Kontrolle der Domänen Zeit, Aspekt und Modus (englisch *tense*, *aspect*, *modus* - TAM). Heine fragt sich nun, wie es möglich ist, daß Auxiliare sich einerseits wie Verben verhalten können (vgl. 2-12 und 2-13), andererseits aber wie Funktoren, die keinerlei lexikalische Bedeutung mehr tragen. Aufbauend auf diesem Widerspruch betrachtet Heine die Auxiliare als Klasse, die einem Sprachwandelprozeß unterliegt, der mit dem von Meillet (1912) geprägten Begriff *Grammatikalisierung* bezeichnet wird. Im Verlauf dieses Prozesses erwirbt eine ehemals autonome lexikalische Einheit die Funktion einer abhängigen grammatischen Kategorie, was sich phonologisch, morphologisch, syntaktisch und semantisch ausdrückt. Die Grammatikalisierung hat einerseits eine diachronische Perspektive, da der Sprachwandel historisch nachzuweisen ist und immer seinen Ausgangspunkt in einer autonomen lexikalischen Einheit hat, andererseits aber auch einen synchronischen Aspekt, da der lexikalische Gebrauch (wie der Gebrauch des Auxiliars als Vollverb) neben der grammatischen Verwendung weiterbestehen kann. Der Prozeß vollzieht sich stufenweise; Heine nimmt zwischen den beiden Endpunkten des Grammatikalisierungsprozesses, also zwischen dem Punkt, an dem der Prozeß noch nicht eingesetzt hat und dem Punkt des völligen Vollzugs, noch fünf weitere Stufen an, auch um sprachkontrastive Verhaltensweisen von Auxiliaren erfassen zu können.<sup>12</sup> Die Stufen sind im Laufe des Sprachwandels unidirektional durchlässig. Ihre Verbindung faßt Heine als eine Kette auf, die *Verb-to-TAM-*

---

<sup>12</sup> So differiert z.B. das Verhalten englischer und deutscher Modalverben stark. Englische Modalverben weisen keinerlei Flexion auf und belegen eine fixe Position im Satz. Deutsche dagegen verhalten sich sehr viel mehr wie Vollverben. Heine (1993:73) erklärt sich dies, indem er die Position deutscher und englischer Auxiliare auf unterschiedlichen Stufen der Grammatikalisierung annimmt.

*Chain*. Wie oben angedeutet, vollzieht sich der Wandel auf mehreren sprachlichen Ebenen, für die jeweils Wandlungsparameter angegeben werden können. Auf semantischer Ebene vollzieht sich die *Desemantisierung*, morphosyntaktisch die *Dekategorisierung*, morphophonologisch die *Klitisierung* und phonologisch die *Erosion*. Für diese Parameter sind die Endpunkte der Verb-to-TAM-Chain in *Tab. 2-1* aufgeführt (Heine 1993:78).

**(Tab. 2-1) Verb-to-TAM-Chain**

	Verb	→	TAM
Desemantisierung	lexikalisches Konzept	...	grammatikalische Funktion
Dekategorisierung	volle Morphosyntax	...	Verlust aller verbalen Eigenschaften (Imperative, Nominalisierung, Passivierung, Negation)
Klitisierung	unabhängiges Wort	...	Klitik, später Affix
Erosion	volle phonologische Form	...	keine Betonung, kein Satzakzent möglich

Die Parameter sind voneinander unabhängig; im Heineschen Sieben-Stufen-Modell ist die Desemantisierung schon auf der dritten Stufe vollständig vollzogen, während Klitisierung und Erosion erst auf der siebten und damit letzten Stufe vollendet werden.<sup>13</sup>

Das Modell der Verb-to-TAM-Chain erfaßt den unterschiedlichen Gebrauch der Auxiliare, sowie die Entwicklung zum Auxiliar (die z.B. zur Zeit das deutsche Verb *brauchen* vollzieht) innerhalb einer Sprache, wie auch den zwischensprachlich differierenden Status von Auxiliaren. Zudem liefert es eine Definition des Begriffs Auxiliar (Heine 1993:131):

Auxiliaries may be defined as linguistic items located along the grammaticalization chain extending from full verb to grammatical inflection of tense, aspect, and modality, as well as a few other functional domains, and their behavior can be described with reference to their relative location along this chain, which is called the Verb-to-TAM chain.

Von den erwähnten funktionalen Domänen kann im Rahmen dieser Arbeit nur ein Teil behandelt werden. Es wurden die Domänen ausgewählt, die vom Autor als für die deutsche Sprache relevant angesehen werden. Dies sind zum einen die sich durch Flexion am deutschen Verb ausdrückenden Kategorien Kongruenz und Tempus, zum anderen die syntaktischen Auswirkungen der Auxiliarselektion.

<sup>13</sup> Dies ist u.a. nötig, um erklären zu können, weshalb deutsche Auxiliare Desemantisierung und Dekategorisierung bereits vollzogen haben, die Prozesse der Klitisierung und Erosion jedoch (noch) nicht.

## 2.2 Auxiliarfunktion: Träger der Finitheitsmerkmale

In den meisten Sprachen der Welt trägt das Auxiliar die Finitheitsmerkmale und findet sich an erster, zweiter oder letzter Position im Satz (Heine 1998:24). Im Deutschen kann es an jeder dieser Positionen vorkommen, im Entscheidungsfragesatz an erster, im Hauptsatz an zweiter, im durch eine Konjunktion eingeleiteten Nebensatz an letzter Stelle (Tab. 2-2).

(Tab. 2-2)	Vorfeld	LK	Mittelfeld	RK	Nachfeld
(a)		<i>Hat</i>	heute morgen der Postbote	geklingelt?	
(b)	Er	<i>ist</i>	heute nicht	gekommen...	
(c)	...,	weil	er	verschlafen	<i>hat</i> .

### 2.2.1 Das topologische Feldermodell

Nach dem topologischen Feldermodell (vgl. Grevendorff 1988) spannen das Hilfsverb bzw. die Konjunktion mit dem Vollverb des Satzes eine sogenannte Satzklammer auf, die den Satz in Felder unterteilt (LK und RK stehen für linke bzw. rechte Satzklammer). In der linken Satzklammer findet sich entweder eine Konjunktion oder das finite Verb. Infinite Verbalelemente stehen immer am Ende des Satzes.<sup>14</sup> In konjunkional eingeleiteten Nebensätzen steht das Finitum hinter den infiniten Verben, in Ausnahmefällen auch davor:

(d) ..., weil er das Unglück *hat* kommen sehen.<sup>15</sup>

Das topologische Feldermodell ist rein deskriptiv und weist einige Schwächen auf: Zwar werden alle grammatisch richtigen Sätze erfaßt, doch werden wegen der lediglich linearen Struktur des Modells keine hierarchischen Beziehungen der Konstituenten zueinander abgebildet. Weiterhin ist das Modell stark ideosynkratisch, da Klammerstrukturen in anderen Sprachen, wie z.B. dem Englischen und Französischen nicht auftreten.<sup>16</sup> Die generative Analyse bietet ein Modell, welches imstande ist, diese Probleme zu lösen und zugleich die Vorzüge des Topologischen Feldermodells zu nutzen.

<sup>14</sup> Es hängt von der Auslegung der Theorie ab, ob sich infinite Verbalelemente am rechten Rand des Mittelfeldes oder an der linken Peripherie der rechten Klammer befinden.

<sup>15</sup> Bech (1955/57) unterteilt das Verbalfeld, welches er Kohärenzfeld nennt, in ein Ober- und ein Unterfeld. Elemente, die normalerweise im rechten Unterfeld beheimatet sind, können ins Oberfeld wandern und damit links des Unterfeldes auftauchen. Hier geschieht aber noch etwas anderes: Das Auxiliar regiert nicht den dritten, sondern den ersten Status (den Infinitiv *sehen* statt des Partizips II *gesehen*). Zur weiteren Analyse vgl. Schmid (2000).

<sup>16</sup> vgl. Lenerz (1995:1266).

### 2.2.2 Das X-bar-Schema und die syntaktische Kategorie AUX

Die schon in der Einleitung erwähnte Generative Grammatik basierte in ihrer frühen Form, dem Standardmodell (Chomsky 1957), auf Regeln zur Erzeugung von Phrasenstrukturen. Es stellte sich jedoch heraus, daß weitgehend unrestringierte Regeln nicht in der Lage sind, das von der Generativen Grammatik geforderte Lernbarkeitskriterium zu erfüllen.<sup>17</sup> Das X-bar-Modell zur Beschreibung der Konstituentenstruktur von Jackendoff (1977) sollte in zweifacher Hinsicht ein höheres Maß an Adäquatheit erreichen (vgl. Ramers 2000:42):

- (I) Beschreibungsadäquatheit: Alle Phrasentypen aller natürlichen Sprachen sollen erfaßt werden (Universalitätskriterium).
- (II) Erklärungsadäquatheit: Die postulierten Phrasenstrukturen müssen im Spracherwerb gelernt werden können (Lernbarkeitskriterium).

Das X-bar-Schema beruht auf mehreren Prinzipien: dem Prinzip der Endozentrität, dem Phrasenprinzip und dem Ebenenprinzip. Das *Prinzip der Endozentrität* besagt, daß jede Phrase als obligatorisches Element einen Kopf der gleichen Kategorie besitzt. Anders ausgedrückt: Jedes syntaktische Wort  $X^{\circ}$  wird zu einer Phrase  $XP$  ( $X''$ ) projiziert. Ergänzungen syntaktischer Wörter oder ihrer Projektionen sind immer Phrasen (*Phrasenprinzip*), ihr Status als Komplement, Adjunkt oder Spezifizierer wird von der Ebene ihrer Mutter im Strukturbaum und der Projektion, deren Schwester sie sind, festgelegt. Im X-bar-Schema werden mehr als zwei Ebenen unterschieden (*Ebenenprinzip*). Jede Phrase kann durch die drei<sup>18</sup> folgenden allgemeinen X-bar-Regeln abgeleitet werden.

- (I)  $X'' \rightarrow (\text{Spezifizierer}) X' (\text{Spezifizierer})$
- (II)  $X' \rightarrow (\text{Adjunkt}) X' (\text{Adjunkt})$
- (III)  $X' \rightarrow (\text{Komplement}) X^{\circ} (\text{Komplement})$

Einzel sprachliche Parameter bestimmen, ob die Konstituenten in den Klammern rechts oder links des Kopfes bzw. seiner Projektionen auftreten. Komplementpositionen müssen notwendig gefüllt werden, wenn der Kopf ein Komplement fordert (selektiert oder subkategorisiert). Adjunkte dagegen haben fakultativen Status. Durch die rekursive Regel (II)

---

<sup>17</sup> Allein um alle möglichen deutschen Verbalphrasen zu erfassen, bedarf es mindestens sieben verschiedener Ableitungsregeln (vgl. Ramers 2000:41). Für ein Beispiel für die Steuerung eines Parsers durch nicht restringierte Phrasenstrukturregeln, vgl. Langer (2001), der allein zum Parsen des Deutschen ca. 2500 Regeln benötigt.

<sup>18</sup> In der neueren GB-Theorie wurde noch die zweite Adjunktregel [ $X'' \rightarrow (\text{Adjunkt}) X' (\text{Adjunkt})$ ] eingeführt, um bestimmte Phänomene, wie die wh-Bewegung, beschreiben zu können (vgl. Haegeman 1994:407). Im Rahmen des MP wird diskutiert, ob das X-bar-Schema konzeptuell notwendig ist (vgl. Ramers 2000:134).

können sie in beliebig großer Zahl innerhalb einer Phrase vorkommen (z.B. Adjektive in Nominalphrasen).

Ausgehend vom X-bar-Schema ließ sich das traditionelle Satzschema (Abb. 2-1) nicht mehr halten: Die Struktur verletzt das Prinzip der Endozentrität; sie enthält nur zwei Ebenen und widerspricht damit dem Ebenenprinzip. Des weiteren finden sich in Sätzen Auxiliärverben, die, wie sich durch Konstituententests, wie beispielsweise die Verschiebeprobe zeigen läßt, außerhalb der eigentlichen VP stehen. Nimmt man nun als Funktion des Auxiliars an, daß es Subjekt und Prädikat miteinander verknüpft und daß das Auxiliar gemäß dem X-bar-Schema projiziert, kann der Satz als Projektion einer Kategorie AUX interpretiert werden (Abb. 2-2, vgl. Chomsky 1965).<sup>19</sup>

Abb. 2-1

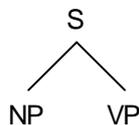
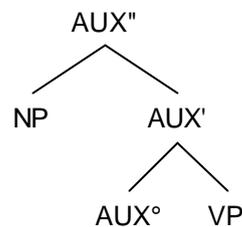


Abb. 2-2



Wie oben erwähnt, ist das Auxiliar Träger der Finitheitsmerkmale im Satz. Zum einen kongruiert es mit dem Subjekt (im Deutschen hinsichtlich der morphosyntaktischen Merkmale Person und Numerus), zum anderen trägt es die verbale Tempusmarkierung. Da auch in finiten Sätzen ohne Auxiliare diese Merkmale zwingend gefordert werden, genügt die Struktur in Abb. 2-2 nicht, um diese Sätze abzuleiten.

### 2.2.3 Funktionale Kategorien und Auxiliare

Statt AUX wird deshalb eine funktionale Kategorie INFL (Chomsky 1981) bzw. I (Chomsky 1986; beides für *Inflection*) angenommen, die die Attribute *Tense* und *AGR* (für *Agreement*/Kongruenz) enthält. Flexionsmerkmale werden von INFL regiert, unabhängig davon, ob sie in Auxiliaren (Abb. 2-3) oder in Vollverben (Abb. 2-4) inkorporiert<sup>20</sup> sind.

<sup>19</sup> Zur Begründung der Spezifiziererposition des Subjekts und der Komplementposition des Verbs siehe Ramers (2000:60ff).

<sup>20</sup> Zum Begriff der Inkorporation vgl. 3.2 und Baker (1988).

Abb. 2-3

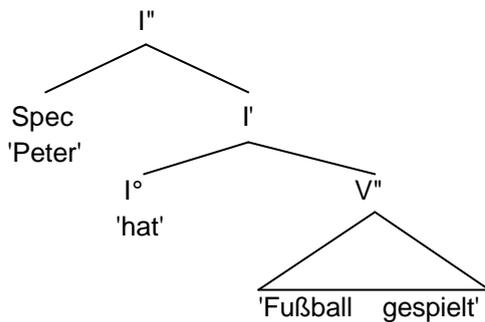
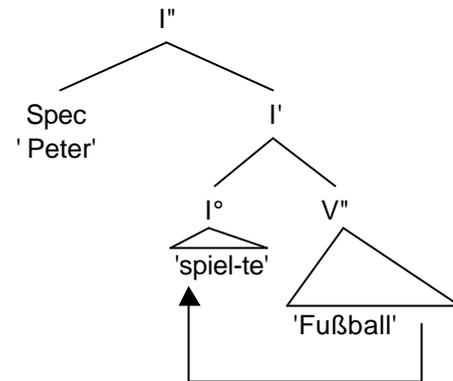


Abb. 2-4



Das Vollverb in *Abb. 2-4* muß sich unter I bewegen, damit es das Flexionsaffix inkorporieren kann.<sup>21</sup> Ein finites INFL vergibt den Kasus NOMINATIV<sup>22</sup> an seinen Spezifizierer (vgl. 2.4).

Die Annahme, der Satz sei eine Projektion des INFL-Knotens und damit eine erweiterte Projektion des Verbs (vgl. Haegeman 1994:108), reicht aus verschiedenen Gründen nicht aus, um die deutsche Satzklammer abzubilden. Zum einen findet sich kein Platz für eine Nebensatzleitende Konjunktion. Des weiteren wird angenommen, daß das deutsche Verb eine zugrundeliegende Verb-Endstellung hat, die in der Nebensatzstruktur sichtbar wird (2-15). Ebenso können keine Sätze abgebildet werden, in denen das finite Verb vor dem Subjekt steht (V1 in 2-16 und V2 in 2-17):

(2-15) ..., *weil* Peter Fußball gespielt *hat*.

(2-16) *Hat* das Wasser schon gekocht?

(2-17) Morgen *bin* ich eingeladen.

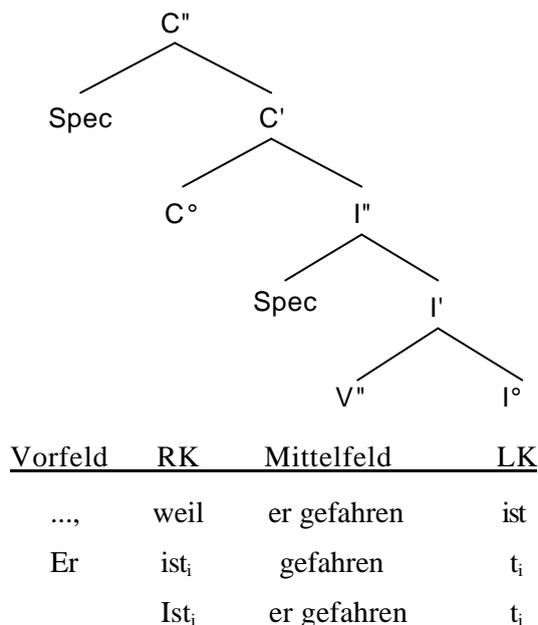
Nimmt man an, daß I° im Deutschen sein Komplement (die VP) nach links regiert, läßt sich die Position I° als rechte Satzklammer annehmen. Es fehlen dann allerdings noch Positionen für die linke Satzklammer und das Vorfeld. Richtet man über der IP eine Schale ein, die ihr Komplement (die IP) nach rechts regiert, werden diese zwei Positionen gewonnen. In der Kopfposition dieser Phrase kann sich eine Konjunktion (*Complementizer*) befinden, daher wird die Phrase CP genannt. Fehlt die satzleitende Konjunktion, so wird im Deutschen das finite Verb in die C°-Position bewegt. Nicht alle Sprachen benötigen zur Repräsentation ihrer

<sup>21</sup> Diese Analyse führt im Englischen zu Problemen. Wie sich an der Negationsprobe zeigen läßt, verbleibt das Vollverb im Englischen in seiner Basisposition, daher muß das Affix sich zum Vollverb bewegen (*Affix-Lowering*). Dies führt theoretisch zu Problemen, da das Affix eine Spur hinterläßt, die nicht korrekt regiert wird. Siehe auch 3.2 und Haegeman (1994:441).

<sup>22</sup> Durch Kapitälchen hervorgehobene Kasusbezeichnungen stehen für abstrakte Kasus, um sie gegenüber morphologischen Kasus kenntlich zu machen (vgl. Haegeman 1994:155f).

Sätze eine CP, ihre Projektion hängt von einem sprachspezifischem Parameter ab. *Abb. 2-5* überträgt das Topologische Feldermodell aus (2.2.1) auf das allgemeinere CP-IP-Schema, welches die Kriterien Lernbarkeit und Universalität zu erfüllen vermag. Zudem werden in diesem Modell hierarchische Beziehungen zwischen den Konstituenten erfaßt.

**Abb. 2-5**



Folgende Gründe stützen die Annahme, IP sei das Komplement des Satz Kopfes C° (nach Ramers 2000:69ff):

1. Der Komplementierer bestimmt das Finitheitsmerkmal des Verbs: Bestimmte Konjunktionen (bspw. *daß*) fordern ein finites, andere (bspw. *ohne*) ein infinites Verb.
2. Der Komplementierer kann den Modus des Verbs bestimmen (bekannt v.a. aus dem Lateinischen): nach *donec* steht bspw. der Indikativ, auf *num* folgt der Konjunktiv.
3. In der CP wird der Satzmodus festgelegt. C bildet eine funktionale Kategorie, die das grammatische Merkmal [+/- wh] enthält. Von diesem hängt ab, ob der Satz in deklarativem oder interrogativem Modus steht. Aufgrund ihrer Modussensitivität wird die CP bisweilen auch als Modusphrase bezeichnet (vgl. Lohnstein 2000).

Vor allem aber ermöglicht die CP-IP-Struktur, alle deutschen Sätze abzubilden. In finiten deutschen Sätzen muß C° immer gefüllt sein, entweder durch eine Konjunktion oder das finite Verb.<sup>23</sup> Enthält ein Satz ein Auxiliar, so ist dieses entweder - bei Anwesenheit einer Nebensatz einleitenden Konjunktion - unter I°, sonst unter C° inkorporiert.

<sup>23</sup> Davon ausgenommen ist die C°-Position in indirekten Frage- und Relativsätzen. Dies ist bekannt als das *Reissche Dilemma* (vgl. 3.1.1).

In neueren Ansätzen der Generativen Grammatik wird davon ausgegangen, daß INFL aus mehreren Knoten besteht (vgl. Pollock:1989). Mit dieser Entwicklung setzt sich Abschnitt (3.2) eingehend auseinander.

### 2.3 Auxiliare als Ausdruck verbaler Kategorien: Analytische Bildungen

Der Ausdruck verbaler Kategorien differiert zwischensprachlich stark. So stehen zur Grammatikalisierung zeitlicher Relationen in einigen Sprachen gleich zwei Möglichkeiten zur Verfügung: Tempora und Aspekte.

Geht es beim Tempus um die Relation zwischen der Ereigniszeit und einer Evaluationszeit - vorzüglich der Sprechzeit -, so geht es beim Aspekt um die interne zeitliche Gliederung des beschriebenen Ereignisses. (Vater 1991:65)

Das Englische z.B. besitzt einen durch die *-ing*-Form ausgedrückten progressiven Aspekt, der einerseits ein Ereignis, das andauert, während ein anderes Ereignis stattfindet (2-18), andererseits auch das Präsens ausdrücken kann. Die deutsche Standardsprache besitzt keine systematischen Mittel zum Aspektausdruck, denn das Perfekt trägt zwar den aspektuellen Zug der Abgeschlossenheit, ist aber im wesentlichen temporal zu deuten (vgl. Vater 1991:51ff). Abweichend von der Standardsprache findet sich in der westdeutschen Umgangssprache die sogenannte *rheinische Verlaufsform*, die morphologisch und semantisch dem englischen Progressiv entspricht (2-19). Das Standarddeutsche drückt den Aspekt nicht durch grammatische, sondern durch lexikalische Mittel aus, beispielsweise durch das Adverb *während* (2-20).<sup>24</sup>

(2-18) I was writing my MA thesis when I got ill.

(2-19) Ich war meine Magisterarbeit am schreiben, als ich krank wurde.

(2-20) Während ich an meiner Magisterarbeit schrieb, wurde ich krank.

Andere Sprachen wie das Finnische und Estnische aus der Familie der uralischen Sprachen unterscheiden nicht zwischen Vergangenem, Gegenwärtigem oder Zukünftigem, sondern lediglich zwischen den Zuständen vollzogen und nicht-vollzogen, was aspektuell als perfektiv vs. imperfektiv gedeutet wird (vgl. Vater 1991:38).

---

<sup>24</sup> Die Unterscheidung zwischen grammatischen und lexikalischen Mitteln zum Ausdruck für Zeitreferenz geht auf Comrie (1985:8) zurück. Grammatische Mittel (Tempus und Aspekt) sind grammatikalisierte und im Allgemeinen am Verb ausgedrückte Kategorien. Als lexikalische Mittel werden eigenständige Ausdrücke bezeichnet, wie etwa Temporaladverbien (*jetzt, gestern*) oder temporale Präpositional- oder Nominalphrasen (*vor Weihnachten, letzten Monat*).

Im Folgenden wird diese Analyse in Bezug auf die deutschen Verbalkategorien ausgearbeitet.

Schulgrammatisch flektiert das deutsche Verb nach fünf Kategorien, die zwischen zwei und sechs verschiedene Ausprägungen annehmen können (vgl. Duden 1998:206ff).

- Tempus (Präsens, Präteritum, Perfekt, Plusquamperfekt, Futur I + II)
- Genus Verbi (Aktiv, Passiv)
- Modus (Indikativ, Konjunktiv, Imperativ)
- Person (Lokutiv, Allokutiv, Delokutiv)
- Numerus (Singular, Plural)

Die Flexion hinsichtlich Person und Numerus ist abhängig vom Subjekt, mit dem das Verb kongruiert. Abgesehen von diesen *Agreement*-Kategorien verfügt das finite Verb über fünf synthetisch gebildete Paradigmen (vgl. Engel 1996:412):

- Präsens, Präteritum, Imperativ, Konjunktiv I + II

Alle weiteren Formen werden analytisch gebildet. Dabei steht ein Teil (das Auxiliar) in einem der fünf finiten Paradigmen, der andere, der auch aus mehreren Elementen bestehen kann, in einer der infiniten Formen Infinitiv oder Partizip II.<sup>25</sup> Analytische Formen bilden einen Verbalkomplex, der im einfachsten Fall aus zwei Elementen besteht, dem finiten Auxiliar und dem infiniten Vollverb. Dazu zählen die passiven Indikativformen des Präsens und Präteritum (2-21, 2-22), das Perfekt, das Plusquamperfekt (PQPerfekt) und das Futur I im Indikativ Aktiv (2-23 bis 2-25), sowie das Zustandspassiv (2-26):

(2-21) Das Buch *wird gestohlen*.

(2-22) Das Buch *wurde gestohlen*.

(2-23) David *hat* das Buch *gestohlen*.

(2-24) David *hatte* das Buch *gestohlen*.

(2-25) David *wird* das Buch *stehlen*.

(2-26) Das Buch *ist gestohlen*.

Werden (2-23 bis 2-25) ins Passiv gesetzt, entsteht ein dreigliedriger Verbalkomplex mit einem Vollverb und zwei Auxiliaren, von denen das eine die Kategorie Tempus, das andere die Kategorie *Genus Verbi* ausdrückt:

(2-27) Das Buch *ist gestohlen worden*.

(2-28) Das Buch *war gestohlen worden*.

<sup>25</sup> Diese infiniten Formen werden von Bech (1955/57) als *Supina* bezeichnet, die im Unterschied zu den *Partizipien* nicht deklinierbar sind. *Supina* wie *Partizipien* umfassen drei Status: den reinen Infinitiv, den Infinitiv mit *zu* und das Partizip Perfekt:

	<b>Supina</b> (+prädikativ)	<b>Partizipien</b> (+attributiv)
1. Status	lieben	liebende
2. Status	zu lieben	zu liebende
3. Status	geliebt	geliebte

(2-29) Das Buch *wird gestohlen werden*.

Der gleiche Satz, ausgedrückt im Futur II, benötigt gleich drei Auxiliare neben dem Vollverb:

(2-30) Das Buch *wird gestohlen worden sein*.

Insbesondere unter Einbeziehung von Modalverben und dem Konjunktiv sind auch Konstruktionen denkbar, die aus mehr als vier Verbalelementen bestehen. Sie sind jedoch schwer zu verarbeiten und werden im Sprachgebrauch so gut wie nie gebildet, weshalb der Leser hier auch von ihnen verschont bleiben dürfen soll.

Aufgrund der oben geschilderten hohen Produktivität, welche die Kombination von Verbalkategorien entfalten kann, wird hier eine Auswahl getroffen, die das Auftreten und Verhalten von Auxiliaren exemplarisch schildert.

### 2.3.1 Tempus

Tempus wird im allgemeinen als grammatikalisierte Lokalisierung von Ereignissen in der Zeit betrachtet (Comrie 1985:9).<sup>26</sup> Die Zahl der angenommenen Tempora im Deutschen schwankt in der sprachwissenschaftlichen Literatur zwischen zwei und zehn.<sup>27</sup> Schulgrammatisch bis hin zur Duden-Grammatik (1998) werden, wohl auch durch den Einfluß des Lateins, die sechs oben angeführten Tempora angenommen. Diese setzen sich aus zwei einfachen (synthetischen) und zwei Paaren von periphrastischen (analytischen) Typen zusammen. Eine Aufstellung der analytisch gebildeten Tempora findet sich in *Tab. 2-3*.

(Tab. 2.3)

		V o l l v e r b		
		Partizip II	Infinitiv Präsens	Infinitiv Perfekt
<b>A</b>	<i>sein/haben</i> Präsens	Perfekt	(Perfekt)	
<b>u</b>	<i>sein/haben</i> Perfekt	PQPerfekt	(PQPerfekt)	
<b>x</b>	<i>werden</i> Präsens	(Präsens Passiv)	Futur I	Futur II

#### 2.3.1.1 Die Verwendung der Tempora – Die Reichbachschen Zeitrelationen

Das Präsens wird normalerweise verwendet, um die Gleichzeitigkeit von Ereigniszeit (E) und Sprechzeit (S) auszudrücken. Vater (1983:107) zeigt, daß teilweise auch zukünftige Ereignisse und solche, die in der Vergangenheit liegen, aber noch nicht vollendet sind, durch

<sup>26</sup> Demgegenüber steht die Auffassung, daß es sich bei Tempora um den Ausdruck von Sprechhaltungen (besprechende vs. erzählende) handelt. Zu einer Übersicht vgl. Vater (1991:42).

<sup>27</sup> Hier konkurriert die Auffassung, nur die synthetischen Bildungen seien Tempora (Präsens und Präteritum), mit der Auffassung, Tempora umfaßten darüber hinaus noch analytische Bildungen (u.a. auch Doppelperfekt, Doppelplusquamperfekt). Zur Übersicht sei auch hier auf Vater (1991:43f) verwiesen.

das Präsens ausgedrückt werden können. Dies ist auch begründet durch die Natur des Präsens, das als zeitloser Standardfall mit lexikalischen Mitteln zum Ausdruck von Zeitrelationen kombiniert werden kann. Vater sieht die Grundbedeutung des Präsens deshalb als *Nicht-Vergangenheit*.

Präteritum und Perfekt drücken aus, daß E auf der Zeitachse vor S liegt, mithin in der Vergangenheit. Futur wird gewählt, um zukünftige Ereignisse auszudrücken (S liegt vor E).<sup>28</sup>

Zur Erläuterung der weiteren Tempora unterscheidet Reichenbach (1947) nicht mehr nur zwischen E und S, sondern führt noch eine dritte Zeit ein, die Referenzzeit (R): Komplexe Tempora beruhen darauf, daß auf einen relativ zum Ereigniszeitpunkt stehenden Referenzzeitpunkt Bezug genommen wird. So nimmt das PQPerfekt Bezug auf ein Ereignis, welches vor einem Referenzzeitpunkt stattgefunden hat, der seinerseits wiederum vor der Sprechzeit liegt (E liegt vor R liegt vor S). Mit dem Futur II wird ein Ereignis geschildert, das ebenfalls vor einem Referenzzeitpunkt stattgefunden hat, allerdings sind E und R beide zukünftig in Hinsicht auf den Sprechzeitpunkt (S liegt vor E liegt vor R).

Mit Hilfe der Referenzzeit ist es nun auch möglich, zwischen dem Perfekt und dem Präteritum zu unterscheiden. Das Präteritum ordnet einem Ereignis einen Referenzzeitpunkt zu, der in der Vergangenheit liegt. Das Perfekt hingegen versetzt das Ereignis zwar auch in die Vergangenheit, R fällt hier allerdings mit S zusammen, was ausdrückt, daß das Ereignis zum Referenz-/Sprechzeitpunkt vollzogen und abgeschlossen ist. Das Perfekt besitzt damit im Gegensatz zum Präteritum Gegenwartsbezug und mithin aspektuelle Züge.

Die zweite Spalte der *Tab. 24* gibt eine Übersicht über die Reichenbachschen Zeitrelationen. Spalte drei bildet die von Erich/Vater (1989) vorgelegte Analyse der Zeitrelationen ab, die sie *Komplexitätshypothese* genannt haben. Diese ist im Grunde deckungsgleich mit der Reichenbachschen, allerdings kompositionell angelegt: Es wird unterschieden zwischen der *intrinsic* Bedeutung - dem Verhältnis zwischen E und R - und der *kontextuellen* Bedeutung, dem Verhältnis zwischen S und R. Präsens und Präteritum stimmen in ihrer intrinsic Bedeutung überein (keine Zeitdifferenz zwischen E und R - ausgedrückt durch ein Komma), Präsens und Perfekt in ihrer kontextuellen Bedeutung (S,R). Die Vergangenheitstempora auf der einen und die Futurtempora auf der anderen Seite werden auf gleiche Weise deiktisch interpretiert (E<S bzw. S<E).<sup>29</sup>

---

<sup>28</sup> Vater (1991:55ff) geht davon aus, daß das Deutsche keine Futurtempora aufweist; die traditionell als Futur bezeichnete *werden*+Infinitiv-Konstruktion habe eher modale Züge, besonders „da ein modalitätsfreies Futur nur durch das reine Präsens ausgedrückt werden kann“ (mittels lexikalischer Mittel, z.B. Temporaladverbien).

<sup>29</sup> *Deiktisch* meint das Verhältnis von S und E.

(Tab.2-4): Zeitrelationen

Tempus	Reichenbach	Ehrich/Vater <sup>30</sup>	
Präsens	S,R,E	(S,R)	(R,E)
Präteritum	E,R<S	(R<S)	(R,E)
Futur	S,R<E	(S<R)	(R,E)
Perfekt	E<R,S	(S,R)	(E<R)
PQPerfekt	S<E<R	(S<R)	(R<E)
Futur II	E<R<S	(R<S)	(E<R)

### 2.3.2 Genus Verbi

Das "Geschlecht des Verbs" ist eine verbale Kategorie, die sich auf das Verhältnis zwischen semantischen Rollen (Agens, Patiens u.a.), sowie parallel dazu auf das Verhältnis zwischen syntaktischen Funktionen (Subjekt, Objekt) transitiver Verben bezieht. In Nominativsprachen wie den indoeuropäischen Sprachen wird zwischen Aktiv und Passiv unterschieden.<sup>31</sup> Der unmarkierte Fall ist das Aktiv: Dem Subjekt ist die Agens-Rolle zugeordnet, die Verbmorphologie entspricht der von Agreement und Tempus geforderten, das Aktiv wird im Deutschen am Verb folglich nicht eigens ausgedrückt. Im Passiv dagegen fehlt die Agens-Rolle des Aktivsatzes (sie kann nicht mehr als Argument, sondern nur noch als Adjunkt - mithilfe einer *von*-Präpositionalphrase - ausgedrückt werden), dafür stellt das direkte Objekt des Aktivsatzes nun das Subjekt (vgl. 2.4). Am Verb drückt sich das Passiv aus, indem im Vergleich zum äquivalenten aktiven Satz ein Auxiliar (meist *werden*, vgl. unten) hinzutritt und mit dem Partizip II des Vollverbs kombiniert wird. Das Tempusauxiliar der transitiven Verben, *haben*, wird im Passiv zu *sein*; dieses trägt weiterhin die Fintheitsmerkmale (vgl. den Satz (2-33c), der keine Vergangenheit ausdrücken kann, sondern entweder eine futurische oder eine modale Lesart hat), das Passivauxiliar *werden* steht in der Partizipform<sup>32</sup> (Perfekt, 2-33b) oder im Infinitiv (Futur I, 2-34b).

<sup>30</sup> Wie in Fußnote 28 erwähnt, geht Vater davon aus, daß das Deutsche keine Futurtempora aufweist. In der Tabelle wurde die kompositionelle Darstellung der Futurtempora deshalb hier hinzugefügt.

<sup>31</sup> In einigen wenigen Nominativsprachen (u.a. Altgriechisch) besteht das Genus Verbi aus Aktiv, Passiv und Medium. Letzteres ist in semantischer Hinsicht Reflexivkonstruktionen ähnlich (vgl. Bußmann 1990:476).

<sup>32</sup> Interessant ist die Tatsache, daß *werden* zwei Partizip II-Formen besitzt: Wird es als Vollverb gebraucht, so lautet die Form *geworden*, die Partizipialform des Auxiliars lautet *worden*:

Meine Oma ist 80 Jahre alt *geworden*. vs. Sie ist reich beschenkt *worden*.

Diese Tatsache legt die Annahme nahe, für Wörter, die sowohl als Auxiliar, als auch als Vollverb gebraucht werden, verschiedene Lexikoneinträge anzulegen (Vgl. 4.3).

- (2-31a) Hans kauft das Buch.  
 (2-31b) Das Buch *wird* gekauft.  
 (2-32a) Hans kaufte das Buch.  
 (2-32b) Das Buch *wurde* gekauft.  
 (2-33a) Hans hat das Buch gekauft.  
 (2-33b) Das Buch ist gekauft *worden*.  
 (2-33c) \*Das Buch *wird* gekauft sein. (keine Lesart für die Vergangenheit)  
 (2-34a) Hans *wird* das Buch kaufen.  
 (2-34b) Das Buch wird gekauft *werden*.

Tabellarisch dargestellt ergibt sich für das *werden*-Passiv Tab. 2-5. Horizontal sind die Tempusauxiliare eingetragen, vertikal die Paradigmen vom passivbildenden *werden*.

(Tab. 2-5)	∅	<i>sein</i> Präsens	<i>sein</i> Perfekt	<i>werden</i> Präsens
<i>werden</i> Präsens	Präsens			
<i>werden</i> Perfekt	Präteritum			
<i>werden</i> PartizipII		Perfekt	PQPerfekt	
<i>werden</i> Infinitiv Präsens				Futur I
<i>werden</i> Infinitiv Perfekt				Futur II

Neben den hier beschriebenen Formen des *werden*-Passivs, auch Vorgangs- oder Akkusativpassiv genannt, verfügt das Deutsche über weitere Konstruktionen, die das Passiv ausdrücken können. Dies sind unter anderem das *kriegen*-Passiv (auch Rezipienten- oder Dativpassiv genannt, bei dem das indirekte Objekt zum Subjekt wird; vgl. 2-35) und das *sein*-Passiv (Zustands- oder adjektivisches Passiv; bei dem eine *von*-Phrase nicht aktiviert werden kann; vgl. 2-36).<sup>33</sup>

- (2-35a) Gerd schenkte Doris (DAT) einen Blumenstrauß.  
 (2-35b) Doris (NOM) kriegte / bekam / erhielt einen Blumenstrauß geschenkt.  
 (2-36) Das Schaf ist (\*vom Schäfer) geschoren.

Der folgende Abschnitt setzt sich mit syntaktischen Auswirkungen der Passivkonstruktion auseinander und stellt diese in Zusammenhang mit der in (2.3.1) erwähnten Auxiliarselektion der Tempusbildung.

<sup>33</sup> Eine Übersicht über die Möglichkeiten, Passiv im Deutschen auszudrücken, findet sich in von Stechow/Nohl (1996:2ff); zum Status der Quasi-Auxiliare *kriegen/erhalten/bekommen* siehe Haider (1986:19ff).

## 2.4 Auxiliare und Theta-Rollen: Auxiliarselektion im Deutschen

Im Unterschied zum Englischen und Spanischen weisen das Deutsche und Italienische bei der Bildung analytischer Tempora eine Alternation zweier Auxiliare auf. Im Englischen und Spanischen wird das Perfekt ausschließlich mit *have* bzw. *habere* gebildet, im Deutschen und Italienischen teilweise mit *haben* bzw. *habere*, teilweise mit *sein* bzw. *essere*.

Haider (1984) zeigt, daß diese Alternation Ausdruck einer grammatischen Regularität ist. Die Selektion des Auxiliars scheint von der Argumentstruktur des selegierenden Verbes abhängig zu sein: Transitive und intransitive (oder unergative) Verben bilden ihr Perfekt mit *haben*:

(2-37) Lisa *hat* Dick geschlagen. [transitiv]

(2-38) Allison *hat* geschlafen. [intransitiv]

Verben, deren Subjekt sich so verhält, wie das Subjekt eines passiven Satzes, d.h. daß es in der Objekt-Basisposition generiert wird und erst durch Raising in die Subjektposition gelangt,<sup>34</sup> sogenannte ergative (oder unakkusative) Verben bilden ihr Perfekt mit *sein*:

(2-39) Lisa *ist* eingeschlafen. [ergativ]

Das Deutsche selegiert für die Passivbildung *werden*, für die Bildung eines Passivsatzes im Perfekt wiederum *sein*:

(2-40a) Dick *wurde* geschlagen.

(2-40b) Dick *ist* geschlagen *worden*.

Die passiven Sätze (2-40a und b) gehen gegenüber Satz (2-37) eines Arguments verlustig. Haider macht dafür die partizipiale Form "ge...n" verantwortlich.<sup>35</sup> Sie *blockiert* jenes Argument, welches im aktiven Satz das Subjekt stellt, die Agens- $\theta$ -Rolle zugewiesen bekommt und den Kasus NOMINATIV erhält. Dieses Argument wird in der generativen Grammatik als das externe Argument bezeichnet, da es seine Basisposition außerhalb der Verbalphrase (VP) hat.<sup>36</sup>

Subjekte ergativer oder passivischer Verben gelangen erst durch Raising in die Subjektposition und werden deshalb vom Blockierungsmechanismus des Partizips nicht erfaßt. Wie verhält es sich aber mit den externen Argumenten transitiver und unergativer Verben? Sie werden in (2-37) und (2-38) offensichtlich auch nicht blockiert. In diesen Sätzen taucht jedoch auch das Auxiliar *haben* auf, dem Haider eine deblockierende Wirkung zuschreibt. Die Blockierung ist also nur scheinbar nicht erfolgt, da sie durch die

---

<sup>34</sup> vgl. Haegeman (1994:306ff).

<sup>35</sup> Haider hebt die traditionelle Unterscheidung zwischen den Partizipien des Perfekts und dem des Passivs im Deutschen auf.

<sup>36</sup> Sportiche (1988) vertritt die Auffassung, daß auch Subjekte in der VP basisgeneriert werden. Diese Arbeit nimmt als Subjektbasisposition die Spezifiziererposition der Predication Phrase an (vgl. 3.1.3) und wird daher nicht näher auf die sogenannte *VP-intern-Subjekt-Hypothese* eingehen.

Deblockierung außer Kraft gesetzt wurde. Fehlt das deblockierende *haben*, so werden die Sätze ungrammatisch:

(2-41) \* Lisa ist Dick geschlagen.

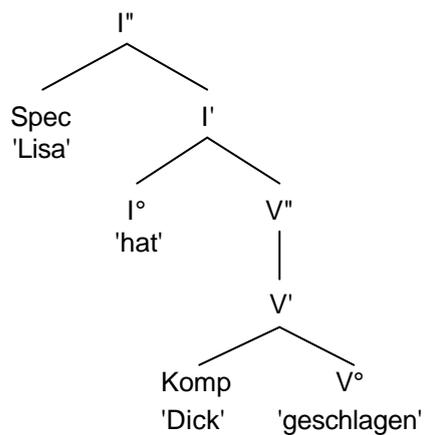
(2-42) \* Allison ist geschlafen.

Tilgt man beim transitiven Satz (2-41) das externe Argument, das nach Haider im Kontext von einer partizipialen Form ohne das Auxiliar *haben* nicht realisiert werden kann, erhält man wieder eine grammatische Konstruktion, das Zustandspassiv:

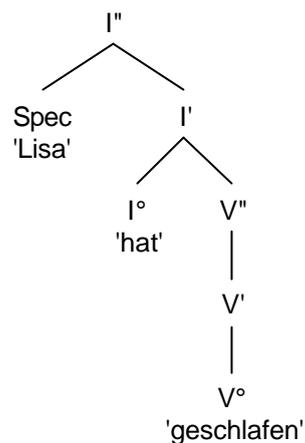
(2-43) Dick ist geschlagen.

Die Auxilarselektion des Verbes hängt also von der Präsenz eines lexikalischen externen Argumentes ab: Fehlt es, so wird *sein* bzw. *werden*, ist es vorhanden, wird *haben* selegiert. Zur Veranschaulichung werden die Sätze (2-37) bis (2-40a) hier in ihrer Struktur dargestellt, in der aus Gründen der Übersichtlichkeit auf die für das Deutsche sonst angenommene CP-IP-Struktur verzichtet wird:

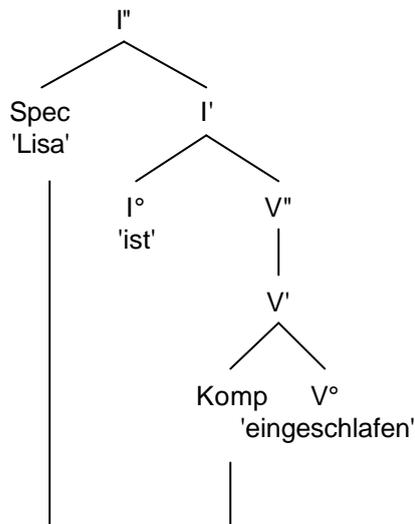
(Abb. 2-6)



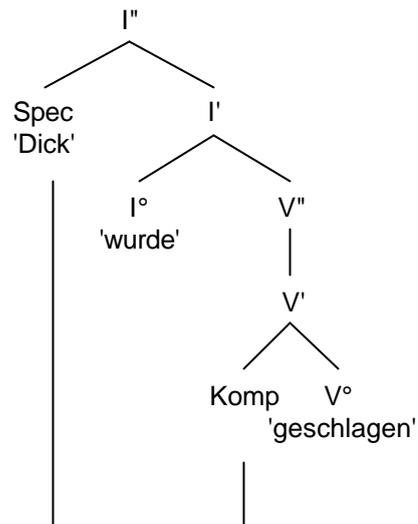
(Abb. 2-7)



(Abb. 2-8)



(Abb. 2-9)



Der Kasus NOMINATIV kann nur extern (von finitem INFL, vgl. 3.2) realisiert werden. Da beide mit einem Argument im Nominativ auftreten, wird der Unterschied zwischen intransitiven und ergativen Verben auf der Oberflächenstruktur verwischt (vgl. *Abb. 2-7* und *Abb. 2-8*). Der Unterschied findet sich im lexikalisch festgelegten Subkategorisierungsrahmen der Verben: Ergative Verben realisieren lexikalisch keinen externen Index, intransitive sehr wohl. Haider nennt diesen Index das *designierte* Argument, um es vom *externen* Argument, das auch strukturell belegt werden kann, zu unterscheiden. Ergative Verben realisieren ihr Argument intern, können ihm jedoch keinen Kasus AKKUSATIV zuweisen (weswegen sie auch unakkusative Verben genannt werden). Um dem Kasusfilter gerecht zu werden, muß das Argument in eine Position angehoben werden, an die Kasus vergeben wird. Dies ist im finiten Satz die [Spec, Infl]-Position (vgl. 3.2).

Wie dargelegt wurde, blockiert das Partizip das lexikalische externe (= designierte) Argument, nicht jedoch das strukturelle. Haider spricht deshalb von lexikalischer Blockierung und Absorption der  $\theta$ -Rolle. Haider (1984:25) stellt weiterhin fest, daß auch die Konstruktion *zu*+Infinitiv wie das Partizip II in attributiver und prädikativer Verwendung das designierte Argument blockiert:

(2-44a) Der *geschlagene* Mann

(2-44b) Der *zu schlagende* Mann

(2-45a) Der Mann ist *geschlagen*.

(2-45b) Der Mann ist *zu schlagen*.

Im prädikativen Kontext mit *haben* wird das designierte Argument wieder deblockiert:

(2-46a) Er *hat* den Mann *geschlagen*.

(2-46b) Er *hat* den Mann *zu schlagen*.

Allerdings ist im satzwertigen Infinitiv keine Blockierung zu beobachten<sup>37</sup>:

(2-47) Er versucht, [PRO die Arbeit fertig *zu stellen*]<sub>CP</sub>.

Da in (2-47) ein lexikalisch deblockierendes *haben* fehlt, muß das Argument auf andere Weise deblockiert werden. Haider nimmt statt dem lexikalischen Deblockierungsmechanismus durch *haben* einen syntaktischen Deblockierungsmechanismus durch das infinite INFL an. Folglich ist das Argument im Falle von (2-45b) auch nicht lexikalisch, sondern syntaktisch blockiert. Die syntaktische Blockierung bezieht sich nicht nur auf die designierten, also lexikalisch externen, sondern auf alle Elemente, die, ob nun lexikalisch oder syntaktisch, extern realisiert werden können. Der Unterschied zwischen lexikalischer und syntaktischer Blockierung läßt sich u.a. an der Alternation der Auxiliare bei ergativen Verben im Hinblick auf deren Konstruktionen im zweiten bzw. dritten Status beobachten. Transitive Verben weisen diese Alternation nicht auf. Scheinbar muß bei der Konstruktion mit dem zweiten Status das externe (nicht-designierte, vgl. *Abb. 2-8*) Argument deblockiert werden (2-48b), was bei der Partizipialkonstruktion (2-49b) nicht der Fall ist:

(2-48a) Lisa ist /\*hat eingeschlafen.

(2-48b) Lisa hat /\*ist einzuschlafen.

(2-49a) Lisa hat geschlafen.

(2-49b) Lisa hat zu schlafen.

Die *zu*+Infinitiv-Konstruktion blockiert somit in der gleichen Weise, in der INFL Kasus zuweist. Im satzwertigen Infinitiv ist das *zu* zwingend erforderlich, da es nach Haider (1986) als Marker für ein infinites INFL fungiert. Nach Haider (1984) befinden sich die INFL-Merkmale im Deutschen in COMP. Dieses COMP fehlt bei Sätzen, bei denen der Infinitiv ohne *zu* auftritt, wie etwa bei AcI-Konstruktionen:

(2-50) Er hört [ihn husten]<sub>IP</sub>.

*Tab. 2-6* stellt eine Übersicht über das Haidersche System der Blockierung und Deblockierung dar:

---

<sup>37</sup> Das designierte Argument besitzt zwar keine phonologische Form, erhält jedoch vom eingebetteten Infinitiv eine Thetarolle. Die generative Theorie nennt ein solches Element PRO. INFL kann in (2-47) keinen Kasus vergeben, da es infinit ist. In der generativen Theorie müssen aber alle overt NPs Kasus erhalten (Kasusfilter). Daher darf PRO nur covert auftreten, d.h. es darf keine phonologische Form haben. Zur Erläuterung siehe Haegeman (1994:253ff).

(Tab. 2-6)

Mechanismus	Blockiertes Element	Blockierendes E.	Deblockierendes E.
lexikalisch	designiertes Argument	Partizip II	<i>haben</i>
syntaktisch	externes Argument	<i>zu</i> + Infinitiv	INFL[-finit]

Wie oben angemerkt, nutzt auch das Italienische unterschiedliche Auxiliare, um zusammengesetzte Zeit-Paradigmen zu bilden. Es läßt sich zeigen, daß die Selektion dieser Auxiliare im Deutschen und Italienischen übereinstimmt, mit Ausnahme eines Kontextes: Italienische Konstruktionen mit klitisiertem Reflexiv, sogenannte *Middle Verbs*, selegieren *essere* (*sein*), die analogen deutschen selegieren *haben* (Beispiele aus Haider/Rindler-Schjerve 1987:1032):

(2-a) Questa macchina si è venduta bene.

(2-b) Dieser Wagen *hat* sich gut verkauft.

Dieser Unterschied wird auf die grundlegende Unterscheidung zwischen dem Italienischen als klitisierender Sprache und dem Deutschen als nicht-obligatorisch-klitisierender Sprache zurückgeführt. Mit der Klitisierung geht eine Veränderung der Argumentstruktur einher:

Middle verbs are ergative variants (see Burzio 1986:38) of transitive verbs, with the clitic *si* as a reflex of the 'loss' of the subject argument. The object of the transitive verb occurs as subject of the middle verb. (Haider/Rindler-Schjerve 1987: 1033)

Durch die Klitisierung kann das designierte Argument, das fortan keinen Argument-Charakter mehr hat, die Blockade des Partizips unterlaufen. Im Gegensatz zum deutschen Reflexiv muß es deshalb nicht von *haben* bzw. *habere* deblockiert werden. Die Unterschiede in der Auxiliarselektion der deutschen und italienischen *Middle Verbs* ist damit auf den Clitic-Parameter zurückzuführen.<sup>38</sup>

## 2.5 Zusammenfassung

Dieses Kapitel setzte sich in erster Linie deskriptiv mit der Verwendung und dem Verhalten von Auxiliaren auseinander. Es wurde dabei schon auf Teiltheorien der Generativen Grammatik Bezug genommen, um die beobachteten Phänomene adäquat beschreiben zu können. Das nächste Kapitel wird die hier beobachteten Gegebenheiten im umfaßenden theoretischen Gebilde der Generativen Grammatik darstellen, um sie schließlich im letzten

<sup>38</sup> Haider und Rindler-Schjerve spekulieren zudem darüber, ob *essere* + Clitic ein starkes Auxiliar wie *haben* ist, das zur Deblocierung fähig ist. Auch Kayne (1993) schlägt eine kompositionelle Analyse von *habere* als *essere* + klitisiertes Possessivpronomen vor. Darauf kann hier allerdings nicht weiter eingegangen werden.

Abschnitt in der formalen Sprache LPS-Prolog widerspruchsfrei formulieren zu können.

Aus den in diesem Kapitel gewonnenen Erkenntnissen erwachsen folgende zu beantwortende Fragen:

1. Wie läßt sich der Unterschied zwischen Vollverben und Auxiliaren (die durch ihre Position auf der Verb-to-TAM-Chain definiert sind) strukturell erfassen?
2. In welcher Position werden Auxiliare basisgeneriert, wohin bewegen sie sich und was sind die Gründe für diese Bewegungen?
3. In welchen Positionen und auf welche Art und Weise werden ihre Merkmale abgeglichen?
4. Wie kann der Blockierungs-/Deblockierungsmechanismus beschrieben werden?

### 3. Auxiliare und generative Grammatik

Grundlage der traditionellen generativen Grammatik, der sogenannten Transformationsgrammatik (vgl. Chomsky 1965), bildet die Annahme, alle Sätze einer Sprache ließen sich auf zugrundeliegende, aus Phrasenstrukturregeln abgeleitete Datenstrukturen zurückführen, auf die Algorithmen in Form von Transformationsregeln angewendet werden, die diese zugrundeliegenden Tiefenstrukturen in abgeleitete Oberflächenstrukturen überführen. Innerhalb der seit Ende der siebziger Jahre entwickelten Rektions- und Bindungstheorie (*Theory of Government & Binding* – GB; vgl. Chomsky 1981) wurde die regelorientierte Transformationsgrammatik durch ein prinzipienorientiertes Grammatikmodell ersetzt, das aus einem System von interagierenden Wohlgeformtheitsbedingungen besteht, die auf beliebige strukturelle Repräsentationen anwendbar sind. Die neueste Ausprägung der generativen Sprachtheorie, das Minimalistische Programm (*Minimalist Program* – MP, vgl. Chomsky 1995) verzichtet auf die Unterscheidung zwischen Oberflächen- und Tiefenstruktur. Stattdessen ist die Sichtweise des MP eine dynamische, da nicht von einer vollständig aufgebauten Struktur ausgegangen wird, innerhalb derer sich durch verschiedene Subtheorien gestützte Bewegungen vollziehen können. Der Strukturaufbau wird stattdessen in einem bottom-up-Verfahren durch die Interaktion zweier Prozesse gewährleistet: einerseits durch den Prozeß der *lexical insertion*, die über einen sogenannten *Merger* erfolgt und für den Aufbau von lexikalischen Phrasen verantwortlich ist, andererseits durch einen Prozeß, der

Bewegungen steuert, durch die *Checking-Theorie* (vgl. 3.3) motiviert ist und Strukturen funktionaler Phrasen erzeugt.

Den verschiedenen Ausprägungen der Generativen Grammatik ist die Annahme gemeinsam, daß die Sätze einer Sprache auf grammatische Strukturen abgebildet werden können, die Positionen bereitstellen, in denen durch Bewegung (z.B. morphosyntaktische) Merkmale (Features) abgeglichen oder zugewiesen werden können. Die am Ende des vorigen Kapitels aufgeworfenen Fragen sind daher zunächst so zu modifizieren, daß sie diese zugrundeliegenden Überlegungen der Generativen Grammatik mit einbeziehen. Dies ergibt folgende neue Fragestellung:

- Welche potentiellen Auxiliarpositionen werden in der Generativen Grammatik angenommen?
- Wann und auf welche Art und Weise vollzieht sich eine Bewegung des Auxiliars von einer dieser Positionen in eine andere?
- Wo und wie werden welche Auxiliarmederkmale gecheckt?

Jeder dieser Fragen wird im Folgenden ein Abschnitt gewidmet. Daran anschließend folgen Überlegungen zur Auxiliarbasisposition, bevor die Ergebnisse zusammengefaßt werden.

### 3.1 Verbpositionen

Schon die Valenzgrammatik (vgl. Tesnière 1959) sah das Verb als den zentralen Knoten (*Zentralnexus*) des gesamten Satzes an, da der Satz zumindest aus dem Verb und den von ihm geforderten (und damit von ihm abhängigen) Aktanten bestehen muß.<sup>39</sup> Diese Vorstellung läßt sich auch auf die Generative Grammatik übertragen, da hier in der Kopfposition des Satzes immer ein Verb oder zumindest ein mit ihm korrespondierendes Element - Auxiliar oder Komplementierer - zu finden ist. Von seiner Basis-Kopfposition kann sich das Verb nur in andere (und zwar funktionale) Kopf- und nicht etwa auf Spezifiziererpositionen bewegen (vgl. 3.2).<sup>40</sup> Diese funktionalen Köpfe expandieren gemäß dem  $\bar{X}$ -Schema zu Phrasen, die neben der Kopfposition einen Spezifizierer und ein Komplement enthalten. Der Spezifizierer kann durch bestimmte lexikalische XPs belegt werden; das Komplement ist wiederum eine Phrase einer funktionalen Kategorie oder die Verbalphrase. Zwischen dem Kopf und dem Spezifizierer einer Phrase werden die Features der Phrasen-Kategorie mittels einer Spec-Head-Beziehung gecheckt bzw. zugewiesen. Als Veranschaulichung dieser abstrakten

---

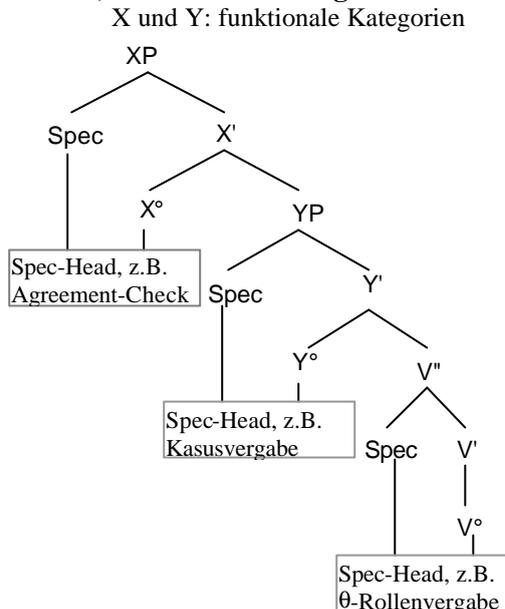
<sup>39</sup> Ausgenommen von dieser Regel sind Bildungen wie *„Mir nach!“* oder *„Au, Du Blödmann!“* (vgl. Vater 1994:114).

<sup>40</sup> Es ist möglich, einen infiniten Verbalteil zu topikalieren oder zu extraponieren. Dies ist dann allerdings keine Kopfbewegung, sondern eine XP-Bewegung der gesamten Verbalphrase, vgl.

[Gefrühstück<sub>i</sub> habe [ich heute noch nicht [<sub>t<sub>i</sub></sub>]<sub>VP</sub>]<sub>IP</sub>]<sub>CP</sub>.

Vorüberlegung soll (Abb. 3-1) dienen:

**(Abb. 3-1) Funktionale Kategorien und VP**



### 3.1.1 Die traditionellen Positionen

Um die Stellung des Auxiliars im Satz zu erfassen, sind in Abschnitt (2.2) schon die lexikalische Verbbasiskategorie V und die funktionalen Kategorien I und C eingeführt worden. Die Motivation dieser Kategorien sei hier noch einmal kurz geschildert. Das lexikalische Verb wird als V°, als Kopf einer Verbalphrase VP, basisgeneriert. In dieser Position ist das Verb jedoch nicht in der Lage, seine Tempus- und Agreementmerkmale direkt abzugleichen, sei es, weil die Flexionsmerkmale an anderer Position eingefügt werden (wie in der GB-Theorie), sei es, weil die im Lexikon eingefügten Merkmale nicht gegen die Merkmale funktionaler Kategorien abgeglichen werden können (wie im MP). Es bedarf also einer Position, in der die Flexionsmerkmale eingefügt bzw. gecheckt werden können. Diese stellt spätestens seit Chomsky (1981) die Kategorie INFL oder später I (Chomsky 1986, beides für *Inflection*) dar. Für die strukturelle Repräsentation der deutschen Satzklammer wurde in (2.2) die CP eingeführt, die als hierarchisch höchste Satzklammer angenommen wurde.

Einzelsprachlich variiert die Verbstellung hinsichtlich der oben genannten Kategorien relativ stark: Während man beim Deutschen bei allen finiten Sätzen und satzwertigen Infinitiven von der CP-Struktur ausgeht, ist diese bei englischen Aussagesätzen nicht notwendig. Im Französischen wird sie nur bei direkten Fragesätzen, Konditionalsätzen und wh-Konstruktionen, in denen bestimmte Adverbien, wie *peut-être* 'vielleicht', *sans doute*

'wahrscheinlich' u.a. topikalisiert sind, angenommen. In allen anderen Fällen kommt das Französische ohne die CP-Schale aus (vgl. Lalande 1997:97). Darüber hinaus ist das finite deutsche Verb nie in seiner Basisposition anzutreffen, das englische - sofern es sich nicht um ein Auxiliar handelt (vgl. 3-2) - jedoch immer. Zu erkennen ist dies an der sogenannten *do-Insertion* bei der Verneinung eines englischen Satzes ohne Tempusauxiliar (3-1b vs. 3-1c) sowie bei Fragen (3-3 und 3-4):

(3-1a) [Er ißt<sub>t<sub>i</sub></sub> [nicht [t<sub>i</sub> ]<sub>VP</sub> t<sub>i</sub>]<sub>IP</sub>]<sub>CP</sub>.

(3-1b) [He doesn't [eat ]<sub>VP</sub>]<sub>IP</sub>.

(3-1c) \*[He eats<sub>i</sub> not [t<sub>i</sub> ]<sub>VP</sub>]<sub>IP</sub>.

(3-2a) [Er hat<sub>t<sub>i</sub></sub> [nicht [gegessen ]<sub>VP</sub> t<sub>i</sub>]<sub>IP</sub>]<sub>CP</sub>.

(3-2b) [He hasn't [been eating ]<sub>VP</sub>]<sub>IP</sub>.

(3-3) [Does<sub>i</sub> [he t<sub>i</sub> [eat ]<sub>VP</sub>]<sub>IP</sub>]<sub>CP</sub> ?

(3-4) [What does<sub>i</sub> [he t<sub>i</sub> [eat ]<sub>VP</sub> t<sub>i</sub>]<sub>IP</sub>]<sub>CP</sub> ?

Die einzelsprachlichen Unterschiede und die Basisposition des Auxiliars werden in den weiteren Abschnitten (3.2 und 3.4) behandelt. Hier soll nur eine kurze Zusammenfassung der klassisch angenommenen Positionen finiter Verben<sup>41</sup> im Deutschen (SOV, hier regieren I° und V° ihr Komplement nach links), und im Englischen und Französischen (beide SVO, hier regieren I° und V° wie C° ihr Komplement nach rechts) folgen:

- Deutsch: Das finite Verb steht in der C°-Position, es sei denn, diese ist belegt durch eine Nebensatzeinleitende Konjunktion. Die C°-Position stellt bis auf Entscheidungsfragen und Nebensätze die zweite Position im Satz, weswegen das Deutsche zu den Verbzweitsprachen gezählt wird (C° muß immer gefüllt sein). Ausgenommen von diesem Effekt sind indirekte Frage- und Relativsätze, bei denen das finite Verb nicht an zweiter, sondern an letzter Position erscheint; C° bleibt leer:

(3-3) Das ist der Mann, der mir das Buch verkaufen *wollte*.

(3-4) Ich weiß nicht, was er mir verkaufen *wollte*. (vs. 'Was *wollte* er dir verkaufen')

Dieses Asymmetriephänomen des Deutschen ist auch als *Reissches Dilemma*<sup>42</sup> bekannt. Im nächsten Abschnitt (3.2) wird eine mögliche Lösung für dieses Problem vorgestellt. In den Fällen, in denen das Finitum nicht unter C° steht, belegt es die Position I°, jedoch niemals die Basisposition V°.

<sup>41</sup> Aufgrund der Übersichtlichkeit der Darstellung werden hier infinite Verben und damit auch infinite Sätze ausgeklammert.

<sup>42</sup> vgl. Lalande (1997:87).

- Englisch: Vollverben verbleiben auch im finiten Zustand in ihrer Basisposition. Das führt innerhalb der GB-Analyse theorieintern zu Problemen, da die Affixe in I° basisgeneriert werden und mittels des sogenannten *Affix lowering* zu dem Verb in V° nach unten bewegt werden müssen. Die Spuren, die diese Affixe zurücklassen, werden ob ihrer höheren strukturellen Position gegenüber dem Landeplatz der Affixe nicht korrekt regiert (vgl. 3.2). Die neuere Analyse im Rahmen des MP löst dieses Problem, da hier nicht mehr angenommen wird, daß Verbstamm und Affix an verschiedenen Positionen in den Satz eingefügt (*gemergt*) werden. Stattdessen werden sprachliche Einheiten aus dem Lexikon vollspezifiziert in die *language faculty* gebracht, wo sie mit Hilfe von *Mergern* wohlgeformte Sätze bilden.<sup>43</sup> Die Verbbewegungen, bei der die Merkmale des Verbs gegen funktionale Kategorien abgeglichen werden, finden erst später (nach dem sogenannten *spell out*) statt (vgl. 3.2.1).

Die Tatsache, daß englische Vollverben in ihrer Basisposition verbleiben und mithin strukturell nicht vor dem Subjekt stehen können, führte zu ihrer Unterscheidung von Hilfsverben, die diese *Subject-Aux-Inversion* erlauben (vgl. Chomsky 1965). Das Subjekt steht im Spezifizierer der IP und wird umrahmt von den Kopfpositionen C° und I°. Bei durchgeführter Inversion in Fragesätzen befindet sich das Auxiliar in C° (vgl. 3-3/3-4), in allen anderen Fällen nimmt es die I°-Position ein.

- Französisch: Das Französische unterscheidet sich

- a) vom Englischen dadurch, daß die Verbbewegung von V° zu I° obligatorisch ist,
- b) vom Deutschen in der Hinsicht, daß es keine systematische V2-Stellung kennt, mithin nicht obligatorisch eine CP-Schale aufweist.<sup>44</sup>

Das französische Verb findet sich also - sieht man von den oben aufgeführten Kontexten ab, in denen es unter C° steht - unter I°. Allerdings läßt sich an infiniten französischen Konstruktionen nachweisen, daß das Verb bisweilen eine Zwischenposition zwischen Negation und VP-initialem Adverb einnimmt. Die Negation wird jedoch nach Pollock (1989) von Infl dominiert, weshalb I° nicht als Verbposition infrage kommt. Welche Kategorie die benötigte Position stellt, soll im nächsten Abschnitt erläutert werden.

### 3.1.2 Pollocks Split-Infl-Hypothese und ihre Erweiterung durch Chomsky

Pollock (1989) untersucht, basierend auf einer umfangreichen Datensammlung, die Kontraste in der Verbstellung zwischen dem Englischen und Französischen. Um die Verbstellung finiter Sätze strukturell zu fassen, genügen die im letzten Abschnitt angeführten Kategorien. Finite Auxiliare werden wie französische finite Vollverben jeweils nach I° bewegt, englische finite

---

<sup>43</sup> vgl. Radford (1998:106ff).

<sup>44</sup> Diachronisch läßt sich die CP-Schale im Altfranzösischen nachweisen, sie ist im Neufranzösischen jedoch weitgehend nicht vorhanden (vgl. Adams 1987).

Vollverben verbleiben in V°. Bei kontrastiver Analyse infinitiver Sätze ergibt sich jedoch zwischen den beiden untersuchten Sprachen ein Unterschied, der eine Zwischenposition annehmen läßt, die zwischen I° und V° interveniert. Zunächst sieht es so aus, als ob der Unterschied hinsichtlich der Vollverbbewegung in infiniten Kontexten aufgehoben wäre:

(3-5a) [Not [to seem happy]<sub>VP</sub>]<sub>IP</sub> ...

(3-5b) [Ne pas [sembler heureux]<sub>VP</sub>]<sub>IP</sub> ...

(3-5c) \* [To seem<sub>i</sub> not [t<sub>i</sub> happy]<sub>VP</sub>]<sub>IP</sub>...

(3-5d) \* [Ne sembler<sub>i</sub> pas[t<sub>i</sub> heureux]<sub>VP</sub>]<sub>IP</sub>...

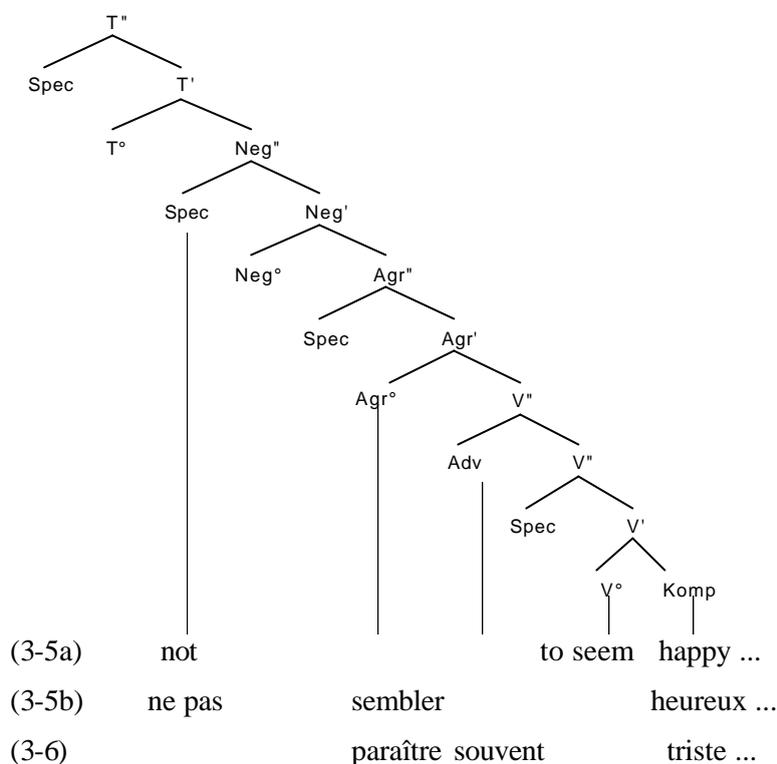
Die Annahme, daß die aus dem Englischen bekannte Beschränkung, keine lexikalischen Verben bewegen zu können, auch für französische infinite Vollverben gelte, läßt sich aufgrund vieler durch Pollock angeführte Gegenbeispiele nicht halten. Das Französische weist sehr wohl Konstruktionen auf, in denen das infinite Vollverb einem - als VP-initial angenommenen Adverb - vorausgeht:

(3-6) Parler à peine l'italien ...

(3-7) Paraître souvent triste ...

Diese Datenlage erfordert eine Position, die zwischen der traditionellen IP und der VP interveniert. Sie wird von französischen infiniten Vollverben genutzt, von englischen hingegen nicht. Pollock nimmt an, daß diese Position der Kopf einer *Agreement*-Phrase (AgrP) ist. Dies ist eine der Kategorien, die der traditionellen IP zugesprochen wurde. Die ursprüngliche IP wird von Pollock fortan als *Tense*-Phrase (TP) bezeichnet. Auch die Negation wird zu einer maximalen Projektion expandiert und interveniert zwischen TP und AgrP. Zur Veranschaulichung der Pollockschen These dient *Abb. 3-2*, in der die oben erwähnten Beispielsätze dargestellt werden.

(Abb. 3-2)



Die Pollocksche Split-InfI-Analyse ermöglicht es, vormalig als ideosynkratisch angesehene, kontrastive einzelsprachliche Effekte auf universalgrammatische Prinzipien zurückzuführen. Diese werden eigens noch erläutert, wenn die Prinzipien der Verbbewegung eingeführt werden (3.2). Die von Pollock vertretene Abfolge von TP und AgrP wurde allerdings inzwischen revidiert, vor allem auch aufgrund der Reihenfolge der Affixe für Tempus und Agreement: Das Tempusaffix steht im Allgemeinen näher zum Verbstamm als das Affix, welches das Agreement ausdrückt (vgl. im Deutschen *lern-e*<sub>1.Pers.</sub> vs. *lern-t*<sub>Prät-e</sub><sub>1.Pers.</sub>, im Lateinischen *stude-t*<sub>3.Pers.</sub> vs. *stude-ba*<sub>Prät-t</sub><sub>3.Pers.</sub>), was zur Annahme führt, daß das Verb das Tempusmorphem vor dem Kongruenzmorphem inkorporiert.<sup>45</sup> Der Grund, die gegenteilige Abfolge anzunehmen, war für Pollock, daß sich eine kurze Verbbewegung zu T° für infinite französische Verben nicht rechtfertigen läßt (zur Diskussion vgl. Lalande 1997:127 und 3.2). Chomsky (1989) löst das Dilemma, indem er die AgrP wiederum in zwei Phrasen aufteilt: eine ist für die Subjekt-, die andere für die Objektkongruenz zuständig. Aus dieser Annahme ergaben sich gleich mehrere Vorteile: Zum einen wurde, wie bereits erwähnt, das Dilemma der Abfolge von TP und AgrP gelöst, da nun sowohl über, als auch unter der TP eine AgrP

<sup>45</sup> Inkorporation ist ein Effekt, der aus der Verbbewegung resultiert (vgl. 3.2).

angenommen werden kann. Weiterhin kann nun das Phänomen der Objekt-Partizip-Kongruenz, die in einigen romanischen Sprachen wie dem Französischen beobachtet werden kann, strukturell erfaßt werden (vgl. Kayne 1989). Außerdem beschränkt man mit eigenen Kongruenzphrasen für Subjekt und (vorerst nur direktes) Objekt die strukturelle Kasuszuweisung auf eine Spec-Head-Beziehung (vgl. 3.3).

Unter den Vertretern der generativen Grammatik ist diese bisweilen als 'Inflation funktionaler Kategorien' bezeichnete Aufsplitterung der IP mehrfach diskutiert worden. Eine schöne Aufstellung der Argumente für und wider die Annahme der Split-Infl-Hypothese findet sich in Schoofs (1994), der nur hinzuzufügen ist, daß Chomsky (1995) inzwischen den Bestand der von ihm selbst mitgeprägten Agreement-Phrasen anzweifelt. Es ist wohl unsinnig, jede neu postulierte funktionale Kategorie als universal anzunehmen,<sup>46</sup> doch soll hier noch eine weitere vorgestellt werden, die v.a. in Hinsicht auf die hier zu behandelnden Auxiliare einige Vorteile mit sich bringt.

### 3.1.3 Bowers Predication Phrase

Bowers' (1993) Überlegungen gehen von dem Problem der Basisstellung des Subjekts aus: Als externes Argument (vgl. 2.4) c-kommandiert es den Rest der Phrase, so daß ursprünglich [Spec, I] als dessen Basisposition angenommen wurde. Ausgehend von dem Ebenenprinzip der X-bar-Theorie stellte daraufhin Kuroda (1988) die Frage, was denn dann die [Spec, V]-Position für ein Element beheimate. Seine Lösung, als *internal subject hypothesis* bekannt geworden, war, das Subjekt als in der Spezifiziererposition der VP basisgeneriert anzunehmen, von der aus es sich aufgrund der Kasuszuweisung in die Spezifiziererposition der IP bewegen muß (Kasuszuweiser für NOMINATIV ist ein finites I°, vgl. 2.4). Der Nachteil beider Annahmen ist, daß sich die Symmetrie der Prädikation von (3-8) und (3-9) nicht erfassen läßt:

(3-8) Ich finde [W. blöd]<sub>SC</sub>.

(3-9) Ich finde, [W. ist blöd]<sub>S (CP/IP/VP)</sub>.

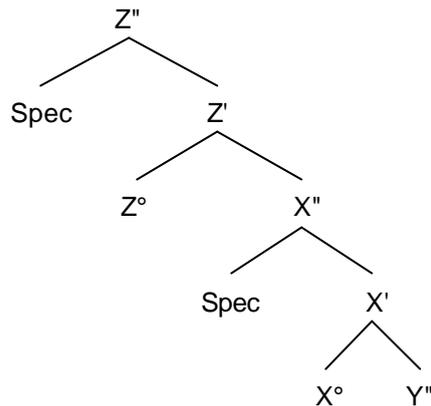
In (3-8), einer sogenannten *Small Clause* (SC), wird das Subjekt ebenso mit dem Adjektiv in Verbindung gebracht wie im Satz (3-9), weshalb Bowers es für nötig hält, daß für beide Subjekte die gleiche strukturelle Position angenommen wird. Doch welche Kategorie sollte dies sein? V und T scheiden aus, da in der SC weder lexikalische Verben noch Tempusmerkmale vorhanden sind. AgrS subkategorisiert T, weshalb auch diese Kategorie

---

<sup>46</sup> vgl. die beiden noch namenlosen Kategorien, die Haegeman (1998) für das untere Mittelfeld des Westflämischen annimmt.

nicht die gesuchte sein kann, ebenso wie AgrO, da es unsinnig wäre, diese als Basisposition des Subjektes anzunehmen. Statt der oben aufgeführten wird eine Kategorie gesucht, deren maximale Projektion entweder zwischen I und V interveniert (wie in Satz 3-9) oder als Komplement von V selegiert wird (wie in der SC 3-8). Strukturell ergibt sich damit folgendes Bild ( $Z=\{I, V\}$ ;  $Y=\{V, A, N, P\}$ ;  $X$ =gesuchte Kategorie):

**(Abb. 3-3): Position der neuen Kategorie**



Wie oben gezeigt, kann die gesuchte Kategorie keine der in den letzten Abschnitten beschriebenen sein. Bowers geht deswegen von einer neuen funktionalen Kategorie aus und nennt sie PrP für *Predikation*-Phrase. Sie hat folgende Eigenschaften:

- (a) Sie stellt mit ihrem Spezifizierer die Position für externe Argumente auf der Tiefenstruktur.
- (b) Sie selegiert eine beliebige lexikalische Phrase als Komplement, diese stellt das Prädikat.
- (c) Sie selbst wird entweder von I° (bzw. T° oder AgrO°) selegiert oder als Komplement des Verbs subkategorisiert.
- (d) Ihre semantische Funktion ist die Prädikation, m.a.W. das in-Beziehung-setzen von Subjekt und Prädikat (was vorher AUX bzw. Infl. zugesprochen wurde).

Bowers führt eine ganze Reihe empirischer Argumente vor allem auch semantischer Art an, die für die Existenz einer PrP sprechen, hier allerdings nicht in ihrer Gesamtheit dargestellt werden können. Die traditionelle Analyse der SC (vgl. Haegeman 1994:123ff) nimmt für sie als maximale Projektion eine AgrP an, die jedoch nicht die AgrP eines Satzes sein kann, da diese eine TP subkategorisiert. Es würde also eines neuen Typs von Agr bedürfen, um die SC strukturell unterzubringen. Stattdessen schlägt Bowers einen neuen Phrasentyp vor, der erfreulicherweise auch die Symmetrie zwischen Sätzen und SCs abbilden kann. Des weiteren ähnelt seine PrP der im Rahmen des Minimalistischen Programms angenommenen Erweiterung der VP, der sogenannten *kleinen v* oder *vP-Shell*. Diese wurde zuerst von Larson

(1988) als eine erweiterte Projektion der VP eingeführt. Die Annahme einer solchen Kategorie hatte vor allem den Vorteil, daß beide Argumente, die strukturellen Kasus erhalten, das Subjekt (in [Spec, v]) und das direkte Objekt (in [Spec, V]) sich stark ähnelnde Positionen einnehmen, so daß die strukturelle Kasuszuweisung in beiden Fällen auf ein Spec-Head-Verhältnis zurückgeführt werden konnte, ohne eine AgrOP annehmen zu müssen. Chomsky (1995) zeigte, daß die vP-Shell aus der Argumentstruktur des Verbs ableitbar ist: Verben, die ein externes (designiertes, siehe 2.4) Argument besitzen, projizieren zur vP-Shell, ergative Verben tun dies nicht. Hier liegt der fundamentale Unterschied zwischen der Chomskyschen vP und der Bowersschen PrP: letztere muß immer projiziert werden, da sie die Semantik der Prädikation syntaktisch ausdrückt.

Für den Autor gleichen die Vorteile der Annahme einer neuen Kategorie Pr bei weitem den Nachteil der wiederum steigenden Anzahl der funktionalen Kategorien aus. Zum einen würde sie für Small Clauses ohnehin benötigt, zum anderen ist die einheitliche Zuweisung von strukturellem Kasus ein - auch für computerlinguistisches Arbeiten, s.u. - angenehmer Effekt. Zudem ist sie bei der Suche nach der Basisposition für Auxiliare hilfreich (vg. 3.4).

### **3.2 Verbbewegung**

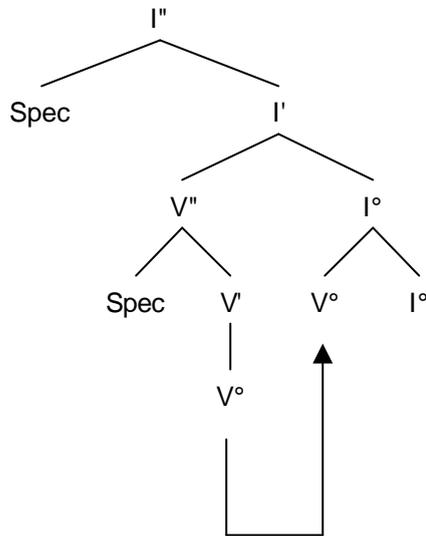
Der enge Zusammenhang zwischen den in (3.1) vorgestellten Positionen und der Verbbewegung führte dazu, daß bisweilen mit dem Phänomen der Bewegung argumentiert werden mußte, um bestimmte Positionen und deren Abfolge zu rechtfertigen. Die der Verbbewegung zugrundeliegenden Prinzipien sollen in diesem Abschnitt behandelt werden. Zuerst wird der Mechanismus der Bewegung dargestellt (3.2.1), um dann auf Bewegung in den Einzelsprachen einzugehen (3.2.2).

#### *3.2.1 Verbbewegung - prinzipiell*

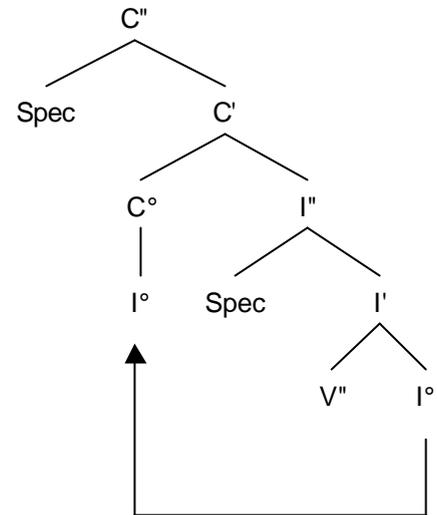
Bewegung innerhalb der Generativen Grammatik ist letztlich nichts anderes als die Überführung einer Struktur in eine andere. Es wird unterschieden zwischen der Bewegung ganzer Phrasen (*A- und A'-Movement*, siehe Haegemann 1994:293ff und 369ff) und der Kopfbewegung (*Head-Movement*), der ausschließlich verbale Elemente und evtl. Klitika unterliegen (vgl. Kayne 1989). Die Kopfbewegung hat immer eine *Inkorporation* des bewegten Kopfes in die empfangende Kategorie zur Folge. Inkorporation ist nach Baker (1988) immer eine Art der Substitution; entweder wird der bewegte Kopf vom empfangenden morphologisch subkategorisiert, wie es bei der Bewegung von V<sup>o</sup> zu I<sup>o</sup> der Fall ist (Abb. 3-4),

oder es findet eine direkte Subkategorisierung statt. In diesem Fall ist der empfangende Kopf, z.B. C°, morphologisch leer (Abb. 3-5).

(Abb.3-4): Morphologische Subkategorisierung



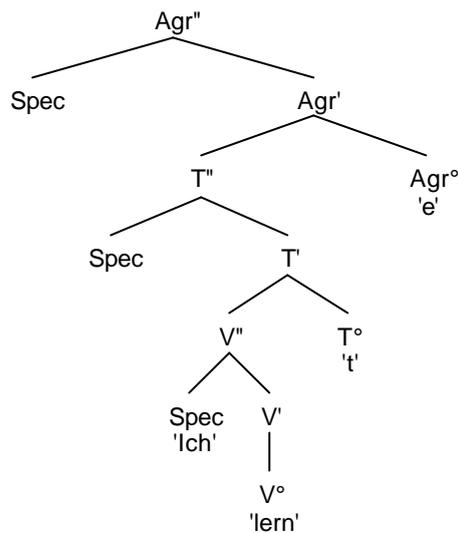
(Abb. 3-5): Direkte Inkorporation



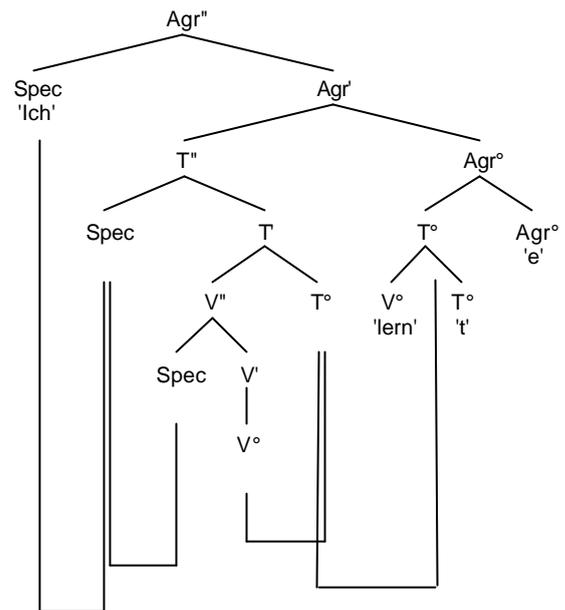
Die aus der Verbbewegung resultierende Inkorporation erzeugt komplexe Köpfe. Durch zyklische Wiederholung wird diese Komplexität noch gesteigert. Abb. 36 zeigt einen Satz in der Derivation von der Tiefenstruktur, in der die Affixe noch unter ihren kategorialen Köpfen stehen, in die Oberflächenstruktur, wo die Affixe durch die Bewegung des Verbs quasi 'eingefangen' wurden<sup>47</sup>:

<sup>47</sup> In einer flektierenden Sprache wie dem Deutschen lassen sich nur wenige Fälle finden, in denen das Tempusmorphem sich eindeutig vom Agreementaffix trennen läßt. In agglutinierenden Sprachen ist die Trennung sehr viel klarer nachzuvollziehen.

(Abb. 3-6) Vor...



...und nach der Derivation



Die Bewegung eines Elementes hinterläßt eine Spur in ihrer Ausgangsposition. Diese Spur muß lizenziert (gebunden) sein; die Bedingungen für diese Lizenzierung benennt das *Empty Category Principle* (Haegeman 1994:442, Übersetzung von mir):

**(3-10) Empty category principle (ECP)**

Spuren müssen streng regiert sein.

A regiert B streng genau dann, wenn A B theta-regiert oder wenn A B antezedens-regiert.

A theta-regiert B genau dann, wenn A B regiert und A B theta-markiert.

A antezedenz-regiert B genau dann, wenn A B regiert und A mit B koindiziert ist.

Der Begriff *Rektion* hat in der GB-Theorie eine zentrale Stellung inne und soll hier kurz definiert werden. Haegeman tut dies folgendermaßen (1994:442; Übersetzung und *Kommentare* von mir):

**(3-11) Rektion**

A regiert B genau dann, wenn

- (i) A ein Regens ist [*zu den Regenten zählen Köpfe und koindizierte XPs*]
- (ii) A B m-kommandiert [*das m-Kommando bezieht sich auf eine Relation der Konstituenten im Strukturbaum. Ein Knoten m-kommandiert seinen Schwesterknoten und dessen Nachkommen, ein Phrasenkopf außerdem alle Elemente innerhalb seiner maximalen Projektion*]
- (iii) sich keine Barriere zwischen A und B befindet [*zu Barrieren zählen bestimmte Kombinationen von XPs, vgl. Haegeman (1994:554ff)*]
- (iv) die Minimalität eingehalten ist [*Es darf keinen Knoten C geben, der ein potentielles Regens für B ist und zwischen A und B interveniert*]

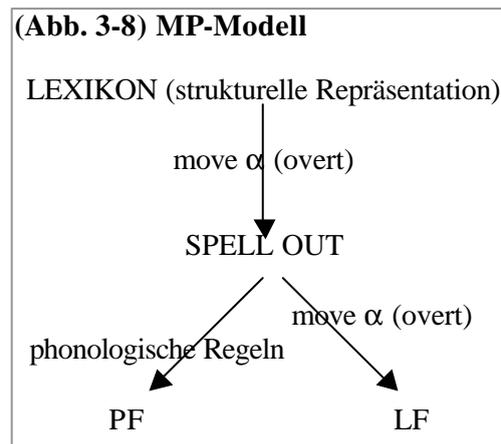
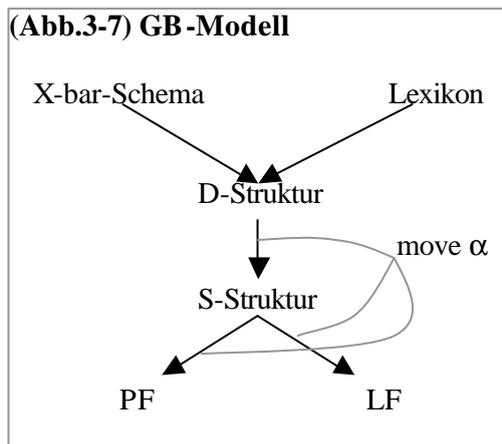
Aus dem ECP ergeben sich für die Kopfbewegung folgende Konsequenzen:

1. Bewegung darf immer nur zur nächsten freien Kopfposition erfolgen (folgt aus 3-11(iii)).
2. Ein Kopf kann sich nicht über einen anderen lexikalisch gefüllten Kopf hinwegbewegen (folgt aus 3-11(iv)).

Konsequenz (1) erklärt, weshalb die AgrP die TP dominieren muß: Wie in (3.1.2) gezeigt, inkorporiert das Verb das Tempusaffix vor dem Agreement-Affix. Das Verb muß also zuerst nach  $T^{\circ}$  bewegt werden, bevor es sich nach  $Agr^{\circ}$  weiterbewegen kann. Das ECP scheidet vor, daß die VP von der TP und diese wiederum von der AgrP dominiert werden muß. Aus Konsequenz (2) ergibt sich der *Head Movement Constraint* (HMC, Travis 1984), eine Spur kann nicht über einen intervenierenden Kopf gebunden werden, vgl.:

(3-12) \*[You have<sub>i</sub> [should [t<sub>i</sub> read this book]<sub>VP</sub>]<sub>IP</sub>]<sub>CP</sub>

Im Unterschied zur oben dargestellten GB-Analyse erfolgt aus der Sicht des MP die Verbbewegung nicht, weil das Verb durch sie erst die Finitheitsmerkmale erlangt, sondern sie ist eine Folge der einzig verbliebenen Wohlgeformtheitsbedingung für die syntaktische Repräsentation von Sätzen, dem *Principle of Full Interpretation* (vgl. Chomsky 1995): Strukturaufbau und Bewegung sind Resultat eines einzigen Derivationsmechanismus, der sogenannten *Generalized Transformation*. Dieser durchläuft verschiedene Stadien und liefert einen Output zum einen auf der Ebene der Phonetischen Form (PF), der Schnittstelle zum artikulatorisch-perzeptuellen Interpretationssystem, zum anderen auf der Ebene der Logischen Form (LF), der Schnittstelle zum konzeptuell-intentionalen Interpretationssystem. Die Eingabe in die PF erfolgt im *Spell out*, vor dem overte Bewegungen stattgefunden haben und nach dem keine lexikalischen Einheiten mehr in die Repräsentation eingefügt werden dürfen. Nach dem Spell out können noch covert Bewegungen erfolgen, die dann die Eingaberepräsentation für die LF erzeugen. Dem *Principle of Full Interpretation* zufolge dürfen die lexikalischen Einheiten bei der Eingabe in die LF keine abstrakten funktionalen Merkmale mehr mit sich führen, da diese auf LF nicht interpretierbar sind. Zur Veranschaulichung des Unterschieds zwischen GB- und MP-Modell der Sprachproduktion sollen *Abb. 3-7* und *Abb. 3-8* dienen (aus Ramers 2000:132f).



Die lexikalischen Einheiten werden durch den Prozeß der *lexikal insertion* vollspezifiziert und somit angereichert mit funktionalen Merkmalen in die strukturelle Repräsentation gebracht. Diese Merkmale können gelöscht werden, indem sie mit den korrespondierenden Merkmalen funktionaler Kategorien abgeglichen werden. Der Abgleich (*Checking*) kann mittels einer Spec-Head- oder einer Head-Head-Beziehung erfolgen. Dafür müssen die lexikalischen Einheiten in die Checking-Positionen der funktionalen Kategorien bewegt werden. Der Derivationsprozeß unterliegt allerdings dem Ökonomieprinzip, das möglichst sparsame Bewegungen fordert. Kurze Bewegungen sind ökonomischer als lange, covert sparsamer als overt. Overt Bewegungen finden daher nur unter speziellen Bedingungen statt.<sup>48</sup> Die Existenz dieser Bedingungen ist einzelsprachlich parametrisiert und mithin für die Stellungsvarietäten der Sprachen verantwortlich. Auf LF hingegen gelten für alle Sprachen wieder die gleichen Bedingungen: Dem Prinzip der vollständigen Interpretierbarkeit folgend müssen alle Merkmale gecheckt sein, mithin alle Bewegungen - ob nun overt oder covert - vollzogen sein.

Die einzelsprachliche Parametrisierung und das daraus folgende Stellungsverhalten von Verben wird im nächsten Abschnitt behandelt, dem Überprüfen der Merkmale (*Features*) ist der übernächste Abschnitt vorbehalten.

### 3.2.2 Verbbewegung - parametrisiert

In (3.1) wurde gezeigt, daß Auxiliare und Vollverben in den Sprachen Deutsch, Englisch und Französisch unter verschiedenen Köpfen funktionaler Kategorien zu finden sind. Dies wird gedeutet als eine Folge unterschiedlich langer Verbbewegung, genauer: einer unterschiedlichen Anzahl gleich langer (kurzer) zyklischer Verbbewegungen. Zwischen der

<sup>48</sup> Die Bewegung wird genauegenommen durch zwei Bedingungen gesteuert: *Greed* (bzw. *Last Resort*) und *Procrastinate*. Während *Procrastinate* die spätmöglichste Bewegung fordert, meint *Greed* die 'Gier' funktionaler Kategorien, ihre Features so früh wie möglich checken zu wollen.

Anfangs- und der Zielposition des Verbs muß jeder Kopf einer funktionalen Kategorie durchlaufen werden, sei es aufgrund des ECP (in der GB-Theorie), sei es, weil alle funktionalen Features gecheckt und damit gelöscht werden müssen (im MP).

Den Unterschied zwischen englischen und französischen Vollverben führt Pollock (1989) auf eine abweichende Stärke der funktionalen Kategorie Agr in diesen beiden Sprachen zurück. Er argumentiert, daß Agr im (Neu-)Englischen defektiv<sup>49</sup> ist und daher nicht die Kraft besitzt, das Vollverb unter Agr<sup>o</sup> zu inkorporieren. Es ist nur fähig, verbale Bestandteile ohne Theta-Raster anzuziehen, d.h. Auxiliare. Das liegt daran, daß ein schwaches Agr undurchsichtig (*opaque*) zur Thetarollenzuweisung ist, so daß ein Verb mit Theta-Raster dieses nicht in Agr kopieren und damit diese Rollen nicht zuweisen kann. Dies bedeutet eine Verletzung des Projektionsprinzips (vgl. Haegeman 1994:55). Das französische Agr ist dagegen reich und damit *transparent* zur Thetarollenzuweisung, so daß französische Verben unter Agr<sup>o</sup> und weiter unter T<sup>o</sup> bewegt werden können, ohne daß das Projektionsprinzip verletzt würde. Englische lexikalische Verben dagegen können sich nicht unter Agr<sup>o</sup> bewegen, weswegen ihnen aufgrund des ECP auch die Bewegung nach T<sup>o</sup> nicht möglich ist.

Das MP erklärt die unterschiedlich weite Bewegung damit, daß englische finite Auxiliare ein starkes funktionales Merkmal Agr haben,<sup>50</sup> das unbedingt vor *Spell out* gecheckt werden muß und damit dem Prinzip *Greedy* Folge leistet. Die englischen Vollverben haben ein solches starkes Merkmal nicht, weshalb sie dem Prinzip *Procrastinate* gehorchen und erst nach *Spell out* bewegt werden.

Das deutsche COMP besitzt im Gegensatz zum englischen und französischen COMP - sofern es nicht durch eine Konjunktion gefüllt und nicht selbst subkategorisiert ist<sup>51</sup> - im finiten Satz immer ein starkes Merkmal [+I], welches gecheckt werden muß, so daß sich verbale Elemente unter diesem Umstand immer in die C<sup>o</sup>-Position bewegen müssen.

Allen drei Sprachen gemeinsam ist die Bewegung des Verbs nach Pr<sup>o</sup>: Die Prädikation muß in jedem Fall vor *Spell out* ausgedrückt sein. Eine Übersicht über die angenommenen Bewegungen der finiten Verben der drei behandelten Sprachen findet sich in *Tab. 3-1*.

---

<sup>49</sup> Ein Indiz für die Defektivität ist die Tatsache, daß sich im Neuenglischen nur die Verbalform in der 3. Pers Singular von den anderen Formen unterscheiden läßt. Das Altenglische weist sehr viel mehr Kongruenzformen auf, weshalb sich dort auch eine Bewegung des Vollverbs nachweisen läßt (vgl. Pollock 1989:418ff).

<sup>50</sup> vgl. (3-13) *I am - you are - he is - we/you/they are*

<sup>51</sup> Das ist die Lösung des Reisschen Dilemmas: indirekte Fragesätze sind vom Matrixverb subkategorisiert, weshalb das eingebettete Verb – trotz Finitheit - nicht die C<sup>o</sup>-Position okkupieren kann (vgl. Lalande 1997:107).

**(Tab.3-1): Kontrastive Aufstellung der finiten Auxiliär- und Vollverbewegung**

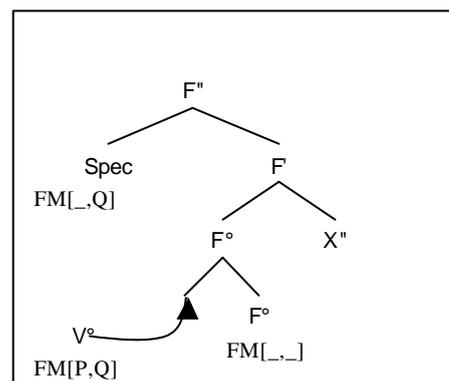
Finitum	Sprache	V° → Pr°	Pr° → T°	T° → Agr°	Agr° → C°
Auxiliär	Deutsch	immer	immer	immer	wenn C [+I]
	Englisch	immer	immer	immer	bei Fragen
	Französisch	immer	immer	immer	u. best. Bed.
Vollverb	Deutsch	immer	immer	immer	wenn C [+I]
	Englisch	immer	nie	nie (ECP)	nie (ECP)
	Französisch	immer	immer	immer	u. best. Bed.

### 3.3 Feature-Checking

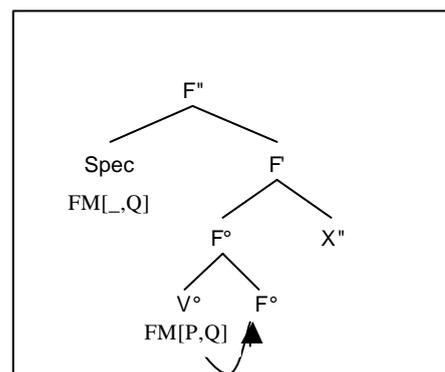
Wie im Vorfeld bereits erwähnt, erfolgt das Checken der Features über Head-Head bzw. Spec-Head-Beziehungen. Dabei werden die Feature-Matrizes der Beteiligten miteinander unifiziert.<sup>52</sup> Dies geschieht entweder durch Substitution von Leerstellen mit instanziierten Werten oder durch Vergleich der Attributwerte: Sind die Attribute auf den gleichen Wert instanziiert oder leer, gelingt die Unifikation. Weichen die Attributwerte voneinander ab, so scheitert sie. (Abb. 3-9) bildet die schrittweise Vorgehensweise der Unifikation in einer funktionalen Kategorie F ab. Das Verb schreibt seine Feature-Matrix in die Kopfposition der funktionalen Kategorie, diese unifiziert danach ihre kopierte Feature-Matrix mit der des Spezifizierers:

**(Abb. 39) Mechanismus des Feature-checkings**

1. Schritt: Das Verb wird in die funktionale Kategorie inkorporiert.

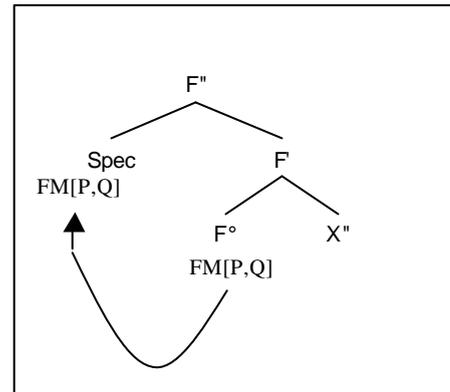


2. Schritt: V° und F° werden im Head-Head-Verhältnis gegeneinander gecheckt, F° erhält die V°-Featurematrix.



<sup>52</sup> Die hier dargestellte Sicht bezieht sich schon auf die Verfahrensweise im LPS-System (vgl. Rolshoven 2001).

3. Schritt: Die Features von  $F^\circ$  werden mit dem Spezifizierer der FP abgeglichen.



Der Unterschied auf theoretischer Ebene zwischen Zuweisung (GB) und Checking (MP) hat letztlich keine praktischen Auswirkungen auf die Unifikation in LPS (vgl. 4.1), wohl aber auf die Komplexität der Lexikoneinträge: nimmt man im Lexikon vollspezifizierte Formen an, so steigert dies seinen Umfang beträchtlich. Eine Lösung dieses Problems für die maschinelle Sprachverarbeitung wird unter Kapitel (4.1) geboten. Zunächst sollen jedoch die Phrasen für die Unifikation der einzelnen im zweiten Kapitel ausgearbeiteten Features benannt werden.

Subjekt und Objekt eines Satzes wird jeweils ein struktureller Kasus und eine Thetarolle zugewiesen. Das geschieht für beide in unterschiedlichen Positionen: Das Subjekt erhält in [Spec, Agr] NOMINATIV und in [Spec, Pr] die externe Thetarolle (meist AGENS), das Objekt in [Spec, V] AKKUSATIV und die weitere vom Verb geforderte Thetarolle. Für die Kongruenz werden für beide nominale Satzteile verschiedene AgrPs angenommen. Es ist zu überlegen, ob diese durch die Präsenz der PrP überflüssig werden,<sup>53</sup> man verlöre allerdings dann die Position für die von Pollock beschriebene kurze Verbbewegung französischer infinitiver Vollverben (vgl. 3.1.2). In der TP erfolgt der Abgleich der verbalen Tempusmerkmale, aus denen die vom Satz ausgedrückten Zeitrelationen abgeleitet werden sollen. Dies birgt einige Schwierigkeiten, da die Zeitrelationen durch das Zusammenspiel

<sup>53</sup> Die Kasuszuweisung des NOMINATIV müßte dann in die PrP verlegt werden. Hierbei muß einerseits beachtet werden, daß nur ein  $Pr^\circ$ , das eine finite VP selektiert, NOMINATIV zuweisen kann. In einer SC erhält [Spec, Pr] den Kasus AKKUSATIV (vgl. 3-14).

(3-14) Ich finde [den Mann dumm]<sub>SC</sub>.

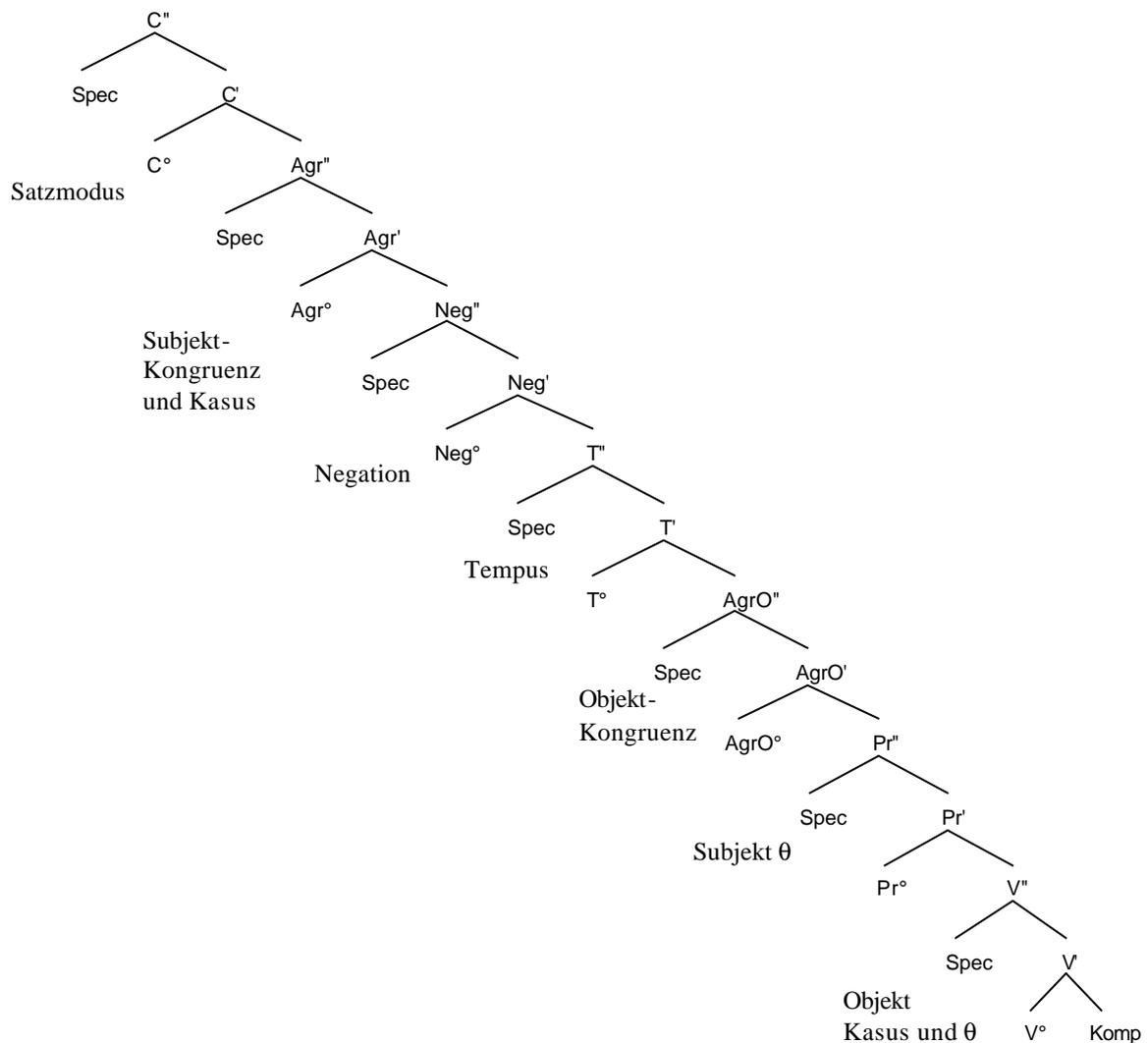
Andererseits wäre eine Bewegung des Subjekts in  $T^\circ$  nicht zu rechtfertigen, da diese Bewegung bisher durch die Notwendigkeit der Kasuszuweisung in Agr gegeben war.

Weiterhin erhält das interne Argument nur den Kasus AKKUSATIV, wenn es ein externes Argument gibt. Von der Akkusativzuweisung sind damit ergative und passivische Verben ausgenommen. Beides wird in (4.2) formalisiert werden müssen.

zweier verbaler Kopfelemente ausgedrückt wird (vgl. 2.3), von denen der infinite Teil in der VP zurückbleibt. Auch wird von einigen Forschern angenommen, daß intrinsische und kontextuelle Bedeutung des Satzes in zwei verschiedenen funktionalen Kategorien ermittelt werden (vgl. Giorgi & Pianesi 1991). In dieser Arbeit wird dieser Schritt nicht vollzogen, die Zeitrelationen werden einzig in der TP ermittelt. Eine detaillierte Beschreibung des dafür nötigen Mechanismus findet sich in (4.2.2). Für die Ermittlung des Satzmodus steht die CP zur Verfügung, worauf diese Arbeit jedoch aufgrund ihres begrenzten Umfangs nur hinweisen kann.

Obige Überlegungen werden in *Abb. 3-10* in die aus (3.1) bekannte Satzstruktur übertragen:

**(Abb. 3-10): Featureverteilung innerhalb der funktionalen Kategorien**

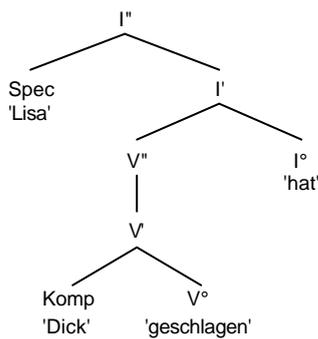


### 3.4 Zur Basisposition des Auxiliars

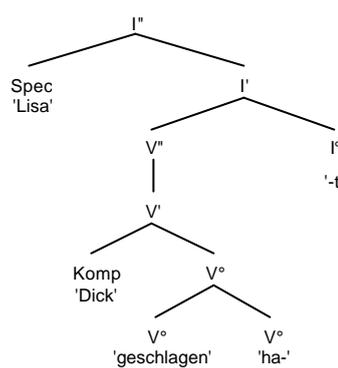
Zur Basisposition der Auxiliare finden sich in der Literatur recht unterschiedliche Angaben. In der traditionellen generativistischen Theorie wurde das Auxiliar als in der IP basisgeneriert angenommen (vgl. Haegeman 1994:114). Dem gegenüber steht die These, das Auxiliar expandiere nicht zu einer eigenen Phrase, sondern stehe mit dem Vollverb in dessen Grundposition unter V°. Diesen Ansatz vertritt auch Bierwisch (1989). Er hält zumindest die futur-, perfekt- und passivbildenden Auxiliare *werden*, *haben* und *sein* für affixoide Verben, die den 1. oder 3. Bechsen Status regieren. Eine weitere Lösung bietet Schmid (2000) an: Sie projiziert das Auxiliar zu einer VP, die als Komplement die VP des infiniten Verbalbestandteils subkategorisiert. Abb. 3-11 bildet diese drei konträren Ansichten auf die angenommenen Tiefenstrukturen ab:

(Abb. 3-11) Tiefenstrukturelle Position der Auxiliare nach...

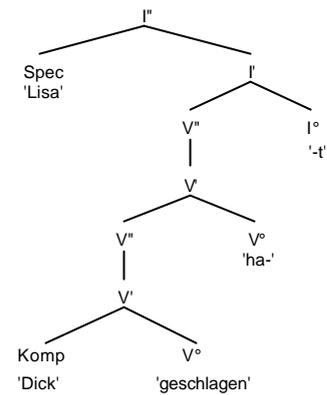
(a) Haegeman (1994)



(b) Bierwisch (1989)



(c) Schmid (2000)



Die Annahme von Struktur (b) bedürfte einer umfassenden Änderung der Ansicht über Inkorporationsstrukturen. Diese entstehen nach Baker (1988) nur durch Bewegung und sind nicht bereits in der Tiefenstruktur vorhanden. Deshalb soll diese Ansicht hier nicht weiter verfolgt werden.

Schmid bezeichnet die aus ihrer Auffassung der Auxiliar-Basisposition resultierende Struktur (c) als *Verbkette*. In einer Verbkette bestimmt das übergeordnete Element den Status des e-kommandierten untergeordneten: Die klassischen Hilfsverben *sein*, *haben*, *werden* fordern ein Komplement im dritten Status, Modalverben den ersten Status und Vollverben den zweiten.<sup>54</sup>

<sup>54</sup> U.a sind folgende Ausnahmen sind zu beobachten:

- *haben* kann alle drei Status regieren: Ich habe *gegessen*. vs. Ich habe *zu essen*. - Ich habe *essen wollen*.
- *helfen* kann den ersten und zweiten Status regieren: ...weil er mir den Esel *beladen* / *zu beladen* hilft.

Allerdings birgt diese Auffassung das Problem, daß das Vollverb strukturell als ein Argument des Auxiliars fungiert und folglich von diesem eine Thetarolle zugewiesen bekommen muß.<sup>55</sup>

In dieser Arbeit wird daher angenommen, daß die Basisposition der Auxiliare unter  $I^\circ$ , genauer unter  $Pr^\circ$  zu finden ist. Dadurch ergeben sich folgende Vorteile:

- a) Die  $Pr^\circ$ -Position regiert die eingebettete VP unmittelbar. Über diese Kopf-Komplement-Beziehung kann der Status des eingebetteten Verbs überprüft werden.
- b) Da die [Spec, Pr]-Position als Basisposition für das externe Argument angenommen wird (vgl. 3.1.3), ist der Deblockierungsmechanismus von *haben* (vgl. 2.4) auf eine Spec-Head-Relation zurückführbar.
- c) Bisherige Annahmen über die Inkorporation und die Subkategorisierung von Verben müssen nicht geändert werden.
- d) Der in Abschnitt (2.1) eingeführten Verb-to-TAM-Chain wird Rechnung getragen: Als Auxiliare verwendete Elemente können keine VP mehr projizieren. Sie haben ihren Platz stattdessen in einer in der Struktur höher stehenden funktionalen Kategorie. Für englische Modalverben, denen jegliche Flexion fehlt, kann evtl. auch eine höhere Basisposition als  $Pr^\circ$  angenommen werden.

### 3.5 Zusammenfassung

Ziel dieses Kapitels war die Einbindung der aus Kapitel (2) gewonnenen Erkenntnisse in die Generative Grammatik. Dafür wurde ein Weg gewählt, der zugleich den Umgang dieser Theorie mit anderen verbalen Elementen schildert. Dies führte zu folgenden Ergebnissen:

- Die Basisposition des Verbs ist die Kopfposition der  $V^\circ$ . Gibt es ein externes Argument, wird dem Spezifizierer der VP - dem Objekt - der Kasus AKKUSATIV und die vom Verb vergebene interne Thetarolle zugewiesen. Fehlt ein externer Index, so wird nur die Thetarolle vergeben, das Objekt muß in eine andere Position geraist werden, in der es Kasus erhält. Nimmt man keine AgrO an, so wird in der VP außerdem eine eventuelle Objektkongruenz mit dem Partizip abgeglichen.
- Als Basisposition von Auxiliaren wird  $Pr^\circ$  angenommen. Von hier aus wird der Status des eingebetteten Verbs unter  $V^\circ$  regiert, eventuell ein externes Argument deblockiert und ihm die Agens-Thetarolle zugewiesen. Kann das externe Argument nicht deblockiert werden (was der Fall ist, wenn das Auxiliar nicht *haben* ist), so wird keine Thetarolle vergeben, sondern das

---

<sup>55</sup> Die Analyse von Schmid (2001) hat allerdings den Vorteil, daß sich mit ihr auch Verbketten, die vom hier behandelten Muster Auxiliar-Vollverb abweichen, analysieren lassen. Sie sollte daher in Betracht gezogen werden, wenn diese Konstruktionen in die maschinelle Sprachverarbeitung einbezogen werden.

interne Argument in die [Spec, Pr]-Position angehoben. Tilgt man die Subjekt-Agreement-Phrase aus der Struktur, so wird in der PrP in einem [+finit]-Kontext zusätzlich noch der Kasus NOMINATIV zugewiesen, sowie die Kongruenz zwischen Subjekt und Finitum gecheckt.

- Das Auxiliar wird in die Kopfposition der jeweils einbettenden Phrase bewegt und geht mit dem dort beheimateten funktionalen Kopf eine Inkorporation ein. Features werden zuerst innerhalb dieser Inkorporationsdomäne gecheckt, später dann mit dem Element, das sich in der Spezifiziererposition der Phrase befindet, abgeglichen. In der TP wird die Zeitrelation, in der CP der Satzmodus festgelegt. In der AgrP wird im finiten Kontext der Kasus NOMINATIV an den Spezifizierer zugewiesen und die Kongruenz zwischen Subjekt und Finitum überprüft. Die Anzahl der zyklischen overt Bewegungen, denen das Auxiliar unterliegt, ist einzelsprachlich parametrisiert. Sie hängt davon ab, wie stark die funktionalen Kategorien das Auxiliar anziehen. Ein Auxiliar kann immer nur aus der direkt eingebetteten Phrase angezogen werden.
- Nach *Spell out* werden alle noch nicht vollzogenen Bewegungen des Verbs durchgeführt und die Merkmale der durch diese coverten Bewegung durchlaufenen funktionalen Kategorien überprüft.

Das nächste Kapitel hat nun die Aufgabe, diese Erkenntnisse soweit zu formalisieren, daß sie maschinell verarbeitet werden können.

#### **4. Auxiliare und LPS-Prolog**

In Kapitel (2) wurden einige Eigenschaften von Auxiliaren dargestellt und in Kapitel (3) in den Rahmen der Generativen Grammatik eingebunden. Erklärtes Ziel dieser Arbeit ist es, mit dem natürlichsprachlichen Phänomen Auxiliar maschinell umgehen zu können. Dazu bedarf es einer Formalisierung, die von der Maschine interpretiert werden kann. Die hier gewählte Formalisierung ist die am Institut für Sprachliche Informationsverarbeitung der Universität zu Köln von Rolshoven seit 1986 entwickelte objektorientierte deklarative Programmiersprache LPS-Prolog. Diese wird im folgenden Abschnitt vorgestellt. Daran anschließend werden die in den vorangehenden Kapiteln gewonnenen Erkenntnisse in LPS-Prolog umgesetzt. Den Abschluß bilden Überlegungen, die den Eintrag von Auxiliaren in das maschinenlesbare Lexikon des LPS-Systems betreffen.

#### 4.1 LPS-Prolog und linguistische Objektorientierung

Prolog<sup>56</sup> ist eine rein deklarative logische Programmiersprache; seine Syntax besteht aus drei Arten von Ausdrücken: *Fakten*, *Regeln* und *Anfragen*. *Fakten* drücken Eigenschaften von Objekten und ihre Beziehungen untereinander aus. Sie bestehen aus einem Prädikat und den diesem zugeordneten Argumenten. Die Zahl der Argumente legt die Wertigkeit des Prädikates fest. Regeln bestehen aus einem zweistelligen Prädikat, dessen Argumente der Regelkopf und der Regelkörper sind. Die Argumente dieses Prädikats können wiederum Prädikate sein, in denen Beziehungen zwischen Argumenten ausgedrückt werden, die nur unter bestimmten Bedingungen gelten. Als Beispiel für eine Regel wird die Kongruenzbeziehung zwischen dem Finitum und dem Subjekt prädikatenlogisch ausgedrückt:

```
agreement(Subject, Finitum):-  
    UnifyValues(Subject, Finitum, 'Person'),  
    UnifyValues(Subject, Finitum, 'Numerus').
```

Die Regel ist wie folgt zu interpretieren: Im *Regelkopf* - der ersten Zeile - steht das zu beweisende Prädikat. Die Kongruenz zwischen Subjekt und Finitum ist vorhanden genau dann, wenn (*gdw* wird durch das Zeichen :- ausgedrückt) der *Regelkörper* - die folgenden beiden Zeilen - bewiesen werden kann. Im Regelkörper finden sich wiederum zwei Prädikate, die durch ein logisches *UND* - das Komma - miteinander verknüpft sind. Sie beginnen mit einer Majuskel, da sie sogenannte *Built-In-Prädikate*<sup>57</sup> sind, die nur im Regelkörper auftreten können und damit nie selbst durch einen Regelkörper beschrieben werden. Stattdessen werden sie vom Prolog-System direkt interpretiert. Die beiden Built-in-Prädikate werden bewiesen, wenn es gelingt, Subjekt und Finitum hinsichtlich der Werte für die Attribute *Person* und *Numerus* zu unifizieren. Das Prädikat kann also folgendermaßen umschrieben werden: Zwischen Subjekt und Finitum besteht eine Kongruenzbeziehung *gdw* für sie in *Person* und *Numerus* übereinstimmende Werte gefunden werden können.

Argumente, die mit einer Majuskel beginnen, sind Variablen. Sie müssen instanziiert werden, damit die Regel ausgeführt werden kann. In seiner neueren Entwicklung (vgl. Rolshoven 1996, 2001) hat LPS-Prolog eine objektorientierte Sichtweise angenommen. Dies hat zur Folge, daß Methoden, wie sie das oben formulierte Prädikat darstellt, an Klassen gebunden

---

<sup>56</sup> Gemeint ist hier das sogenannte *Edinburgh Prolog*. Als eine grundlegende Einführung in dessen Syntax sei hier auf Clocksin & Mellish (1987) verwiesen.

<sup>57</sup> Dies bezieht sich, wie die folgende Beschreibung, auf LPS-Prolog.

werden. Eine Klasse ist ein Bauplan für Objekte, sie beschreibt deren Verhalten durch statische Attribute (Datenstrukturen) und dynamische Methoden (Algorithmen).<sup>58</sup> Ein Objekt ist immer Instanz einer Klasse und übernimmt damit den in der jeweiligen Klasse formulierten Bauplan. Die an der Kongruenz beteiligten Klassen werden folgendermaßen formuliert:

```
CLASS SubjectAgr;  
    agreement(Self, Other);  
END;  
CLASS FinitumAgr;  
    agreement(Other, Self);  
END
```

Jede Klasse beginnt mit dem Schlüsselwort `CLASS`. Darauf folgen der Name der Klasse sowie eventuell vorhandene Attribute und Methoden. Abgeschlossen wird eine Klasse durch das Schlüsselwort `END`. Die beiden oben aufgeführten Klassen haben jeweils nur eine zweitstellige Methode `agreement(X,Y)`. Sie rufen diese Methode jedoch mit komplementär verteilten Argumenten auf: Die Klasse `Subjekt` instanziiert das in ihr angelegte Objekt in der ersten Argumentstelle, die Klasse `Finitum` das ihre in der zweiten.<sup>59</sup> Der Beweis der Methode wird ausgeführt, sobald beide Variablen instanziiert sind.

Mit den bisher aufgeführten Klassen kann die Kongruenz zwischen `Subjekt` und `Finitum` jedoch noch nicht überprüft werden, da zum Beweis der Methode die Unifikation zweier Merkmale (Attribute) verlangt wird, die noch nicht an die Objekte gebunden sind. Diese Bindung erfolgt nicht in den oben genannten Klassen, sondern in spezialisierteren, aus jenen ererbenden. Die spezialisiertesten Klassen bei der hier geschilderten objektorientierten Umsetzung stellen die Knoten im von der sprachlichen Äußerung abgeleiteten Strukturbaum. Als `Subjekt` fungiert ein  $D''$ -<sup>60</sup>, als `Finitum` ein  $Agr^o$ -Knoten, unter dem das inkorporierte finite Verb zu finden ist. Ein Objekt der Klasse  $D''$  erbt aus einer abstrakteren Klasse  $D$ , die wiederum aus der Klasse `NomFeatures` u.a. die Attribute `Numerus` und `Person` erbt. Die Werte dieser Attribute werden aus der lexikalischen Information der in der zu analysierenden sprachlichen Äußerung beteiligten Lexeme entnommen.<sup>61</sup> Da diese Information einzelner

---

<sup>58</sup> vgl. Mössenböck (1998).

<sup>59</sup> Der *Self*-Operator entspricht den in objektorientierten Programmiersprachen verwendeten *this*-Operator. Er hat zur Folge, daß sich das Objekt auf dieser Argumentposition selbst instanziiert.

<sup>60</sup> Nach Abney (1987) wird jede NP von der funktionalen Kategorie  $D$  (für Determinierer) subkategorisiert.

<sup>61</sup> Die Systeme LPS und GBX verfügen über die Möglichkeit, auf ein maschinenlesbares Lexikon zugreifen zu können (vgl. 4.3).

Lexeme in den meisten Fällen mehrdeutig ist, wird in der Struktur unterhalb des D"-Knotens das sogenannte *message passing*, ein Kommunikationsprozeß der Knoten untereinander, genutzt, um die Informationsmenge zu reduzieren. Dieser Kommunikationsprozeß beruht ebenfalls auf einem Unifikationsmechanismus.<sup>62</sup> Das Finitum erhält seine Attribute vom inkorporierten Verb. Auch diese können teilweise schon mit Werten belegt sein, die sich durch die Flexionsendung des Verbs ergeben, die aber noch mehrdeutig sein können (vgl. *sie laufen – wir laufen; ich lief – er lief*). Der Kongruenzabgleich ist so auch als Reduktion der Informationsmenge anzusehen, da er diese Mehrdeutigkeiten auflöst.

Die Vererbung wird in LPS-Prolog mittels des Zeichens < ausgedrückt. A<B bedeutet, das A aus B erbt. A ist dabei die spezialisiertere Klasse, B die allgemeinere.<sup>63</sup> Das Subjekt ist ein Objekt der spezialisierten Klasse D", das zusätzlich noch aus der Klasse SubjectAgr erbt. Die beteiligten Klassen werden untenstehend aufgeführt:

```

CLASS NomFeatures | [Genus, Numerus, Kasus, Person];
    nomFeatures(Self); (*gewährleistet den Abgleich der
                        Attribute innerhalb der DP *)
END;
CLASS D < NomFeatures;
END;
CLASS D" < D;
END;
CLASS D"[Subject] < D", SubjectAgr;
END;

```

Die Vererbungshierarchie wirkt für diesen hier dargestellten Ausschnitt etwas umständlich, sie eignet sich jedoch hervorragend für komplexere linguistische Beschreibungen, da sie ob ihrer hohen Modularität imstande ist, das linguistische Wissen kompakt und transparent abzubilden, wie im Folgenden noch gezeigt werden soll.

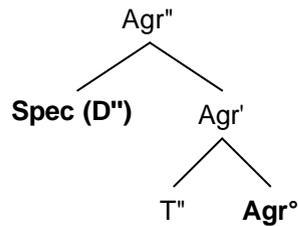
Zunächst aber muß festgelegt werden, in welchem Knoten die oben vorgestellte Kongruenzmethode ausgeführt werden soll. Vgl. dazu den folgenden Strukturbaum:

---

<sup>62</sup> Eine genaue Schilderung des Merkmalabgleichs in der DP findet sich in Hühn (2001).

<sup>63</sup> In LPS-Prolog ist, im Gegensatz zur Programmiersprache Java, Mehrfachvererbung gestattet. Sie ist insofern problematisch, da dadurch eine Klasse die gleiche Vorfahrklasse auf unterschiedlichen Wegen ererben kann. Dies muß durch Vererbungspfadkontrolle ausgeschlossen werden (vgl. Rolshoven 2001:7).

(Abb. 4-1)



In einem Spec-Head-Verhältnis dominiert der X"-Knoten den Spezifizierer unmittelbar, den Kopf allerdings nur mittelbar. Gilt es, eine Methode, die auf einem Spec-Head-Verhältnis beruht, zu beweisen, so muß sie vom Spezifizierer also um einen Knoten, vom Kopf um mindestens zwei Knoten hochgereicht werden, damit beide Variablen instanziiert werden können. Das Hochreichen von Methoden wird vom System maximal eingeschränkt: Im Default-Fall wird die Methode um nur einen Knoten - zur Mutter - hochgereicht; der Zielknoten kann aber auch eigens angegeben werden. Als Zielknoten fungiert der Knoten, in dessen Mutter die Methode instanziiert wird.<sup>64</sup> Im Fall des Spezifizierers tritt der Default-Fall ein, eine explizite Angabe des Zielknotens ist also nicht nötig. Die Klasse des Kopfes (hier: `FinitumAgr`) aber benötigt diese explizite Angabe und muß dahingehend geändert werden:

```
CLASS FinitumAgr;  
    agreement(Other, Self) Agr';  
END;
```

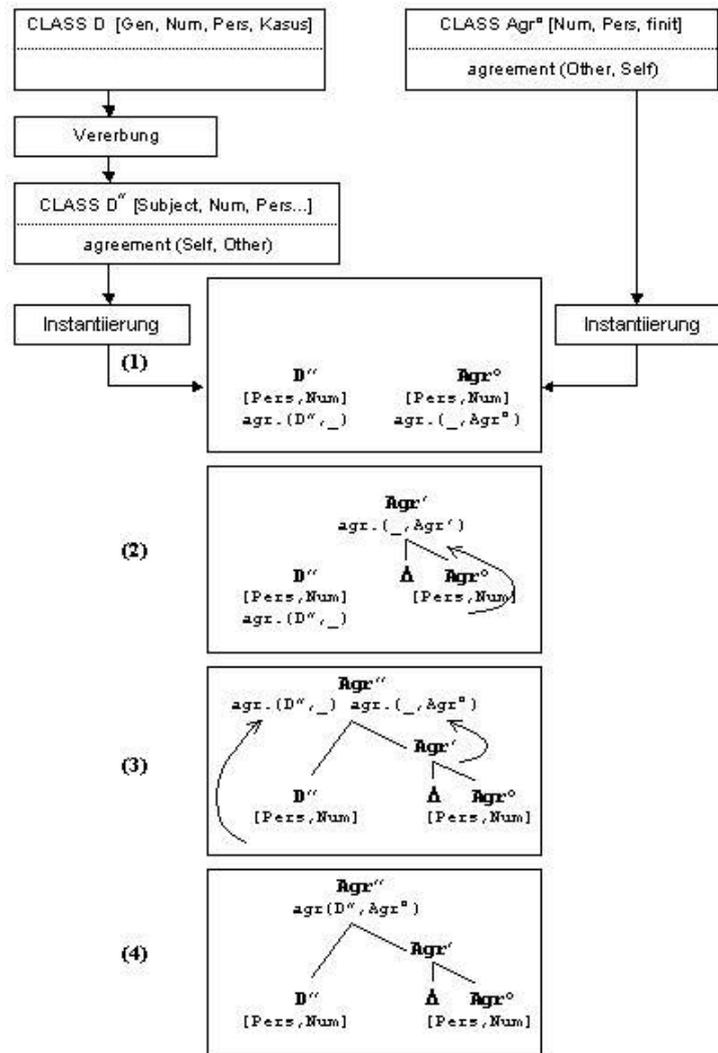
Der Mechanismus wird in *Abb. 4-2* schrittweise beschrieben.<sup>65</sup>

---

<sup>64</sup> Daß die Mutter nicht direkt als Zielknoten angegeben werden kann, liegt daran, daß das LPS-System die Struktur dynamisch aufbaut. So steht zur Zeit des Methodenbeweises die Barzahl der Mutter noch nicht fest. Sie kann somit nicht direkt angesprochen werden und eignet sich daher nicht zur Zielknotenangabe (vgl. Rolshoven 2001:11).

<sup>65</sup> Graphik entnommen aus Rolshoven (2001) und hinsichtlich des Inhalts leicht modifiziert.

(Abb. 4-2)



Zu den Attributen und ihren potentiellen Werten ist noch zu bemerken, daß diese in einer *Attribute-Value*-Datei mit der Dateiendung `.atv` definiert werden müssen. Dies verhindert einerseits Inkonsistenzen, da der linguistische Programmierer Rechenschaft über die von ihm verwendeten Attributwerte in einer zentralen Datei ablegen muß, andererseits hat es den Vorteil, daß Attribut und Wert nicht beide angegeben werden müssen, da das System ihre Zuordnung selbständig vornehmen kann. Dies erfordert eine Überprüfung der Werte, da sie jeweils nur einem Attribut zugeordnet sein dürfen. Untenstehend findet sich ein Ausschnitt aus der deutschen Attributwertliste (`dts.atv`). Die `atv`-Dateien müssen einzelsprachlich festgelegt werden, da jede Sprache spezifische Attribute fordert (s.a. 2.3).

(\* Ausschnitt aus `dts.atv` \*)

```
Person: 1-pers, 2-pers, 3-pers;  
Numerus: sing, plur;
```

Das System muß dafür sorgen, daß die Attributwertliste eingelesen wird, damit die Attribute identifiziert werden können. Daneben müssen das Lexikon und das in Klassen organisierte linguistische Wissen der zu analysierenden Sprache vorliegen. Gewährleistet wird dies, indem diese Komponenten in einer `.gff`-Datei (für *General File of Files*) aufgeführt werden, die zu jeder Anwendung aufgerufen werden muß. Der Inhalt einer solchen `.gff`-Datei (hier: der für das Linguistische Wissen des Deutschen) hat folgende Form:

```
(* dts.gff *)
    builtIn.blt
    dts.atv
    dts.lex
    dts.cff
```

In `builtIn.blt` werden die oben schon erwähnten Built-In-Prädikate definiert. `dts.lex` stellt das deutschsprachige Lexikon, `dts.cff` die Datei *Class File of Files*, in der die für das deutsche benötigten Klassen linguistischen Wissens aufgeführt sind, die ebenfalls eingelesen werden müssen. Diese Klassen werden in Modulen organisiert. Die Modularisierung garantiert eine transparente und konsistente Organisation, sie ist ein zentrales Gütekriterium der objektorientierten Programmierung (vgl. Mössenböck 1998). Einzelne Module haben spezifische Aufgaben, die durch wohldefinierte Schnittstellen abgerufen werden können. Die beliebige Kombinierbarkeit von Modulen hat den großen Vorteil, daß bestimmte sprachliche Phänomene einzelsprachlich, andere wiederum sprachübergreifend modelliert werden können. Die neueren Arbeiten, die in LPS-Prolog ausgeführt wurden (vgl. Lalande 1997 und Rolshoven 2001), unterscheiden zwischen drei verschiedenen Ebenen:

1. Die Ebene der Universalgrammatik: Hier finden sich Module, deren Klassen das sprachliche Wissen, das allen Sprachen gemeinsam ist, darstellen. Es handelt sich hierbei um sehr abstrakte Prinzipien wie z.B. die Rektion (vgl. Rolshoven 1996).
2. Die einzelsprachliche Ebene: Auf dieser Ebene werden die sprachspezifischen Parameter modelliert. Dazu zählt z.B. die Eigenart des deutschen Adjektivs, in seiner Deklinationsart vom Typ des Determinierers abhängig zu sein (vgl. Hühn 2001).
3. Die Ebene sogenannter europäischer Sprachen: Dies ist eine Metaebene, welche zwischen den europäischen Einzelsprachen und UG interveniert. Ihre Einführung hatte nicht den Anspruch, daß die dort beschriebenen Phänomene auf alle europäischen Sprachen übertragen werden können. Vielmehr wird sie als eine anwendungsorientierte Zwischenstufe angenommen, die es ermöglicht, sprachübergreifend zu modellieren. Wird im Laufe der Arbeit von dieser "Europäischen Ebene" gesprochen, so bezieht sich diese auf jenes aus der

Praxis geborene Konstrukt, welches aber zumindest die großen germanischen (Englisch, Deutsch) und romanischen Sprachen (Französisch, Spanisch, Italienisch) umfassen soll.

Der folgende Abschnitt wird einzelne Module aufführen, die aus den vorangehenden Kapiteln erarbeitet wurden. Dabei soll überprüft werden, auf welcher Ebene diese Module angesiedelt werden können.

## 4.2 LPS-Module

Hier sollen insgesamt vier Module vorgestellt werden, die sich auf den Merkmalabgleich bzw. die Merkmalszuweisung, an der Auxiliare beteiligt sein können, beziehen (vgl. 3.3). Außenvor bleibt zunächst die Behandlung des dynamischen Strukturaufbaus mit LPS. Dieser wird momentan noch vom System direkt geregelt, soll aber in naher Zukunft für den linguistischen Programmierer einsehbar sein und damit für ihn modifizierbar werden (Jürgen Rolshoven, persönliche Kommunikation). Der Strukturaufbau beinhaltet auch den dynamischen Prozeß der Verbbewegung, dessen statische Auswirkungen (das Hinterlassen von Spuren und deren Lizenzierung) z.T. schon in Lalande (1997) beschrieben und deshalb hier nicht noch einmal eigens aufgeführt werden.

### 4.2.1 Das Agreement-Modul

Die zum Kongruenzabgleich benötigte Methode und die beiden Klassen, an die diese Methode gebunden ist, wurden bereits bei der Darstellung der Erzeugung linguistischen Wissens in LPS-Prolog in (4.1) aufgeführt. Die Klassen `SubjectAgr` und `FinitumAgr` bilden zusammen das Modul `Agreement`. Die Kongruenz zwischen Subjekt und finitem Verb läßt sich zumindest bei allen Sprachen der oben definierten europäischen Ebene nachweisen. Deshalb wird das Agreement-Modul auf der europäischen Ebene als `EurAgreement.cls`<sup>66</sup> eingebunden. Objekte der Klasse `EurFK.D[subject]` erben aus der Klasse `EurAgreement.SubjectAgr`; Objekte der Klasse `EurFK.Agro` aus der Klasse `EurAgreement.FinitumAgr`.

Für die Objekt-Partizip-Kongruenz werden zwei weitere Klassen in das Modul `EurAgreement` eingeführt: `ObjectAgr` und `PartizipAgr`. Sie besitzen die Methode `objAgreement(X,Y)`, die sie komplementär instanzieren. Sie können allerdings nicht aus den `EurFK`-Klassen geerbt werden, da nicht jede europäische Sprache Objektkongruenz

---

<sup>66</sup> Eine Datei mit der Endung `.cls` (für Classes) enthält eine Zusammenstellung von Einzelklassen, deren Syntax aus Abschnitt (4.1) bekannt ist. Zu jeder `.cls`-Datei gehört eine `.msp`- (für Methodenspezifikation)-Datei, in der die an die Klassen gebundenen Methoden ausformuliert sind (vgl. Rolshoven 1996).

aufweist. Stattdessen sind diese Klassen Erblasser für einzelsprachliche Objekte bzw. Partizipien. Diese sind z.B. im Französischen als Objekte der Klassen `FrzFK.D` [object] und `FrzFK.V` [partizip] instanziiert. Im Deutschen fehlt diese Vererbung. Eine Übersicht über das Modul `EurAgreement` findet sich in *Anhang A*.

#### 4.2.2 Das Tempus-Modul

Abschnitt (2.3) hat gezeigt, daß die grammatischen Mittel, Tempus auszudrücken, einzelsprachlich stark variieren. Daraus ergibt sich, daß die Modellierung der Tempusmodule einzelsprachlich erfolgen sollte; als Beispiel soll hier die Ermittlung der deutschen Tempora aus den beteiligten verbalen Elementen dienen.

Hier wird für das Deutsche die schulgrammatische Anzahl der Tempora übernommen, d.h. es wird unterschieden zwischen sechs Tempora: Präsens, Imperfekt (Präteritum), Perfekt, Plusquamperfekt, Futur I und II. Die Schwierigkeit der Modellierung liegt vor allem in den zwei unterschiedlichen Bildungen des deutschen Tempus. Bei den synthetischen Bildungen erfolgt die Tempusermittlung aus nur einem Element, bei den analytischen aus mindestens zweien (vgl. 2.3):

- Präsens:    Verbstamm + Personalmorphem
- Imperfekt: Verbstamm + 't' + Personalmorphem  
          oder: Präteritumverbstamm + Personalmorphem
- Perfekt:    Partizip II + hab-/sei- + Personalmorphem
- PQPerfekt: Partizip II + hab-/sei-(Präteritum) + Personalmorphem
- Futur I:    Infinitiv + werd- + Personalmorphem
- Futur II:   Infinitiv Perfekt + werd- + Personalmorphem

Die synthetischen Bildungen (Verbstamm + Tempusmorphem + Personalmorphem) werden im Morphologiemodul `DtsMorph` behandelt (vgl. Hühn 2000:12ff) und können so aus `DtsTempus` herausgehalten werden. Das Morphologiemodul ist vorgeschaltet, was es möglich macht, bei der syntaktischen Analyse von Beginn an mit vollspezifizierten Formen arbeiten zu können. Abgesehen von den für die Tempusanalyse unnötigen Personalmorphemen kann das Verb in insgesamt acht Paradigmen auftreten, in fünf finiten (Präsens, Präteritum, Imperativ, beide Konjunktive) und drei infiniten (Infinitiv, beide Partizipien). Analytische Bildungen setzen sich aus diesen Paradigmen zusammen:

**(Tab 4-1)**

	Auxiliar (1= sein/haben, 2= werden)	Vollverb
Präsens	∅	Präsens
Imperfekt	∅	Präteritum
Perfekt	1 - Präsens	Partizip II
PQPerfekt	1 - Präteritum	Partizip II
Futur I	2 - Präsens	Infinitiv
Futur II	2 - Präsens	Infinitiv Perfekt

Bis auf den Infinitiv Perfekt, der wiederum eine analytische Bildung ist, wird die Tempusinformation aus höchstens zwei Elementen gewonnen: dem Finitum und evtl. dem infiniten Vollverbparadigma. In der Attribut-Wert-Liste müssen hier zwei Attribute unterschieden werden: Das immer synthetisch gebildete und mithin im Morphologiemodul ermittelte Paradigma und das Tempus.

```
(* Dts.atv *)
```

```
Verbparadigma: Praes, Praet, Imp, Konj1, Konj2, Inf, Part1, Part2;
```

```
Tempus: Praesens, Imperfekt, Perfekt, PQPerfekt, Futur1, Futur2;
```

Das Tempusattribut wird im Modul `DtsTempus` gesetzt. Betrachten wir zuerst die analytisch gebildeten Zeiten, bei denen zwei Elemente beteiligt sind: Das infinite Vollverb unter  $V^0$  erbt aus der Klasse `DtsTempus.Vollverbparadigma`, das finite Auxiliar unter  $T^0$  aus der Klasse `DtsTempus.Finitum`.<sup>67</sup>

```
(* DtsTempus.cls *)
```

```
CLASS Finitum;
```

```
    setTense(Self, Other);
```

```
END;
```

```
CLASS Vollverbparadigma;
```

```
    setTense(Other, Self) Pr" |Agr0";
```

```
END.
```

<sup>67</sup> Für die Methode des Vollverbparadigmas sind hier zwei Zielknoten angegeben. Dies ist nötig, damit das System in Hinsicht auf eine Annahme der Kategorie `Agr0` flexibel bleibt (vgl. Abb. 3-10).

In der Methode `setTense` muß nach *Tab. 4.1* zunächst zwischen vier verschiedenen Fällen unterschieden werden. Durch ein logisches *ODER* (ausgedrückt durch ein Semikolon) ist diese Fallunterscheidung in LPS-Prolog umzusetzen.<sup>68</sup>

```
(* DtsTempus.Msp *)
setTense(Finitum, VePara):-
    HasValue(Finitum, 'praes'),
    HasValue(Finitum, 'aux1'),
    HasValue(VePara, 'part2'),
    SetValue(Finitum, 'perfekt');
    HasValue(Finitum, 'praet'),
    HasValue(Finitum, 'aux1'),
    HasValue(VePara, 'part2'),
    SetValue(Finitum, 'ppperfekt');
    HasValue(Finitum, 'praes'),
    HasValue(Finitum, 'aux2'),
    HasValue(VePara, 'inf'),
    SetValue(Finitum, 'futur1');
    HasValue(Finitum, 'praes'),
    HasValue(Finitum, 'aux2'),
    HasValue(VePara, 'infperf'),
    SetValue(Finitum, 'futur2').
```

Wie aber können auch das Präsens und das Imperfekt erfaßt werden, ohne dafür eigene Klassen generieren zu müssen? Die Lösung gestaltet sich relativ simpel, da im Deutschen auch das Vollverb in die  $T^\circ$ -Position angehoben wird, zugleich aber eine Spur von sich in  $V^\circ$  zurückläßt. Diese Spur hat zwar keinen phonetischen Ausdruck mehr, ist aber ansonsten eine genaue Kopie des Knotens, der unter  $T^\circ$  steht. So werden `Finitum` und `Paradigma` mit dem gleichen Objekt instanziiert, das demnach auch den gleichen Wert des `Paradigma`-Attributes haben muß. Ist dieser Wert `praes`, wird das `Tempus`-Attribut auf den Wert `praesens` gesetzt; aus dem `Paradigma`-wert `praet` folgt der `Tempus`-wert `imperfekt`:

```
setTense(Finitum, VePara):-
    HasValue(Finitum, 'praes'),
    HasValue(VePara, 'praes'),
    SetValue(Finitum, 'praesens');
    HasValue(Finitum, 'praet'),
```

---

<sup>68</sup> Unter `aux1` sind *sein* und *haben* subsummiert, unter `aux2` vorerst nur *werden*.

```

    HasValue(VePara, 'praet'),
    SetValue(Finitum, 'imperfekt');
...

```

Sprachen, die ihr Vollverb overt nicht unter  $T^{\circ}$  bewegen, verlangen hier nach einer anderen Strategie. Entweder ermittelt man das Tempus in diesen Fällen durch covert Bewegungen oder man verlegt die Ermittlung in die  $Pr^{\circ}$  (theoretisch unsauber).

#### 4.2.3 Das Kasus-Modul

Die Vergabe von strukturellem Kasus ist bei den sogenannten europäischen Sprachen (s.o.) einheitlich geregelt. Auf UG gibt es zumindest zwischen Akkusativ- und Ergativsprachen Unterschiede,<sup>69</sup> so daß hier beschlossen wurde, die strukturelle Kasusvergabe in einem Modul `EurCase` zu modellieren. Sie ist wie folgt geregelt: Gibt es ein finites  $Agr^{\circ}$ , so weist dies NOMINATIV an seinen Spezifizierer zu. AKKUSATIV wird von  $V^{\circ}$  nur dann an [Spec, V] zugewiesen, wenn  $V^{\circ}$  eines der Merkmale `transitiv` oder `ditransitiv` hat und das interne Argument nicht auf die externe Position geraist wurde, um dort NOMINATIV zu erhalten, d.h., daß sich in [Spec, V] keine NP-Spur befinden darf.<sup>70</sup> In LPS-Prolog ergibt sich für die Zuweisungsmethoden so folgendes Bild:

```

nominativ(AgrNull, SpecAgrP):-
    HasAttribute(AgrNull, 'FinVerbPara'),
    SetValue(SpecAgrP, 'nom').

akkusativ(VNull, SpecVP):-
    (HasValue(VNull, 'trans'); HasValue(VNull, 'ditrans')),
    not(HasAttribute(SpecVP, 'Index')),
    SetValue(SpecVP, 'akk').

```

Die Methoden sind so formuliert, daß sie nur dann nach wahr evaluieren, wenn die Kasuszuweisung tatsächlich durchgeführt wurde. Wird auch nur einer der strukturellen Kasus nicht vergeben, scheitert die gesamte Struktur. Dies ist nicht wünschenswert, da in infiniten Kontexten kein NOMINATIV und von intransitiven oder ergativen Verben kein AKKUSATIV vergeben werden darf. Man behilft sich hier damit, daß die Ausführung der Methoden als fakultativ gesetzt wird. Ein Scheitern der Methode führt so nicht zwangsläufig

<sup>69</sup> In Ergativsprachen werden ABSOLTIV und ERGATIV strukturell vergeben, vgl. Dixon (1979).

<sup>70</sup> NP-Spuren folgen aus einer A-Bewegung und unterscheiden sich von wh-Spuren dadurch, daß sie einen anaphorischen Charakter haben (vgl. Haegeman 1994:440).

zum Scheitern der Struktur. Ausgedrückt wird diese Fakultativität in der die Methode aufrufenden Klasse durch ein kleines  $f$ :<sup>71</sup>

```
(* EurKasus.Cls *)
CLASS NomAsser;
    nominativ(Self, Other) Agr',f;
END;
CLASS AkkAsser;
    akkusativ(Self, Other) V',f;
END;
CLASS NomAssed;
    nominativ(Other, Self),f;
END;
CLASS AkkAssed;
    akkusativ(Other, Self),f;
END.
```

Um zu gewährleisten, daß alle overtten D"-Knoten Kasus erhalten, muß aufgrund der Fakultativität der Kasusvergabe noch eine weitere Klasse `CaseCheck` eingeführt werden, an die eine Methode gebunden ist, welche die Kasus-Vergabe überprüft.<sup>72</sup>

```
CLASS CaseCheck;
    checkCase(Self),f;
END.

checkCase(DZwei):-
    HasAttribute(DZwei,'Overt'),
    HasAttribute(DZwei,'Kasus');
    HasValue(DZwei, 'PRO'),          (* An PRO darf kein Kasus vergeben *)
    not(HasAttribute(DZwei, 'Kasus')).(* werden (vgl. Haegeman 1994:269ff*)
```

Welche Objekte aber erben aus den oben vorgestellten Klassen?

- a) Aus der Klasse `CaseCheck` erben alle Objekte der Klasse `D` .
- b) Aus der Klasse `NomAsser` erben alle Objekte der Klasse `Agr`° .

---

<sup>71</sup> *Asser* soll soviel wie *Assigner*, *Assed* soviel wie *Assigned* bedeuten. Aus Kompatibilitätsgründen werden vom LPS-System Bezeichnungen nach dem achten Zeichen abgeschnitten. Es ist also darauf zu achten, daß sich Bezeichner innerhalb der ersten acht Zeichen unterscheiden.

<sup>72</sup> Bei der Modellierung wird zwischen morphologischem und abstraktem Kasus unterschieden. Ihre Übereinstimmung wird ebenfalls in `CheckCase` gecheckt (s. Anhang).

- c) Aus der Klasse `NomAssed` erben alle Objekte der Klasse `D" [ subject ]`.
- d) Aus der Klasse `AkkAsser` erben alle Objekte der Klasse `V°`.
- e) Aus der Klasse `AkkAssed` erben alle Objekte der Klasse `D" [ do ]`.

Lexikalischer Kasus wird in den Klassen `GetDativ`, `SetDativ` etc. abgeglichen (s. *Anhang A*). Das Ererben dieser Klassen wird von lexikalischer Information gesteuert und interagiert mit der `checkCase`-Methode.

#### 4.2.4 Das Theta-Modul

Aus Gründen der Übersichtlichkeit soll hier lediglich die Vergabe der externen Thetarolle an `[Spec, Pr]` und die der internen an `[Spec, V]` behandelt werden. Dabei muß beachtet werden, daß eine externe Rolle nur vergeben wird, wenn sie nicht von der Verbmorphologie (durch *zu* oder *ge-*) blockiert ist (es sei denn, die Blockade ist aufgehoben; vgl. 2.4). Ergative Verben vergeben prinzipiell keine externe Rolle. Eine interne Thetarolle wird außer von intransitiven Verben immer vergeben. Diese Überlegungen führen zu folgender Modellierung der beteiligten Methoden, die an die Klassen `GetInternRole`, `SetInternRole`, `GetExternRole` und `SetExternRole` gebunden sind:

```
setInternRole(Verb,Empfaenger):-
    UnifyValues(Verb, Empfaenger, 'ThetaIntern').

setExternRole(Verb, Empfaenger):-
    not(HasValue(Verb, 'blockON')),
    UnifyValues(Verb, Empfaenger, 'ThetaExtern').
```

Der Attributwert `blockON` wird durch die Verbmorphologie gesetzt. Er kann durch das Auxiliar *haben* wieder gelöscht bzw. überschrieben werden.<sup>73</sup> Dieser Mechanismus wird durch die fakultative Methode `deblock(Auxiliar,Vollverb)` beschrieben, die aus den Klassen `Deblocker` bzw. `Blocked` mit komplementär instanziierten Argumenten aufgerufen wird:

```
CLASS Deblocker;
    deblock(Self, Other),f;
END;
```

---

<sup>73</sup> Hier wird zunächst nur der lexikalische Blockierungs-/Deblockierungsmechanismus bearbeitet, da der syntaktische Mechanismus (vgl. 2.4) nicht auf der Auxiliarselektion beruht und damit nicht Thema dieser Arbeit ist.

```

CLASS Blocked;
    deblock(Other, Self),f;
END;
deblock(Auxiliar, Verb):-
    HasValue(Verb, 'blockON'),
    HasValue(Auxiliar, 'haben'),
    SetValue(Verb, 'blockOFF').

```

blockON und blockOFF sind Werte des Attributs ThetaBlock. Dieses kann nur einen oder keinen der beiden Werte annehmen. Hat es keinen Wert oder den Wert blockOFF, so wird die externe Thetarolle vergeben, hat es den Wert blockON, so geschieht dies nicht.

Während die Zuweisung der Thetarollen im Modul EurTheta abgehandelt werden kann, muß der Deblockierungsmechanismus einzelsprachlich modelliert werden (vgl. 2.4). Dies bedarf einer Schnittstelle zwischen den einzelsprachlichen und den "europäischen" Modulen. Um die Vererbungswege transparent zu halten, muß diese Schnittstelle wohldefiniert sein. Sie wurde zwischen den einzelnen Bar-Ebenen der lexikalischen bzw. funktionalen Klassen festgelegt (vgl. Hühn 2001). Ein Objekt der Klasse DtsLK.V[0-bar] erbt beispielsweise aus EurLK.V[0-bar]. Daraus erwächst folgendes Problem: Auf der europäischen Ebene muß sich ein Objekt der Klasse EurLK.V[0-bar] in eine der Klassen EurV.V[0-bar, intransitiv], EurV.V[0-bar, transitiv], EurV.V[0-bar, ditransitiv] oder EurV.V[0-bar, ergativ] spezialisieren, um die geeigneten Methoden setInternRole bzw. setExternRole bzw. beide an sich zu binden. Erbt nun ein Objekt der Klasse DtsLK.V[0-bar] der definierten Schnittstelle gemäß aus EurLK.V[0-bar], so ist ihm diese Spezialisierung auf dem normalen Vererbungsweg nicht zugänglich. LPS-Prolog hält für solche Fälle den Mechanismus der *späten Bindung* bereit, der ermöglicht, daß Objekte sich während der Laufzeit spezialisieren können. Die Schnittstelle zwischen den Ebenen kann so gewahrt werden. Späte Bindung wird durch einen Stern hinter der zu erbenden, spezialisierten Klasse ausgedrückt und in einer Methode mittels des BuiltIn-Prädikates Extend festgelegt:

```

(* DtsLK.cls *)
CLASS V[0-bar] < EurLK.V[0-bar], DtsV.VIntrans*, DtsV.VTrans*,
    DtsV.VErg*, DtsV.VDitrans*;
    valence(Self);
END;

```

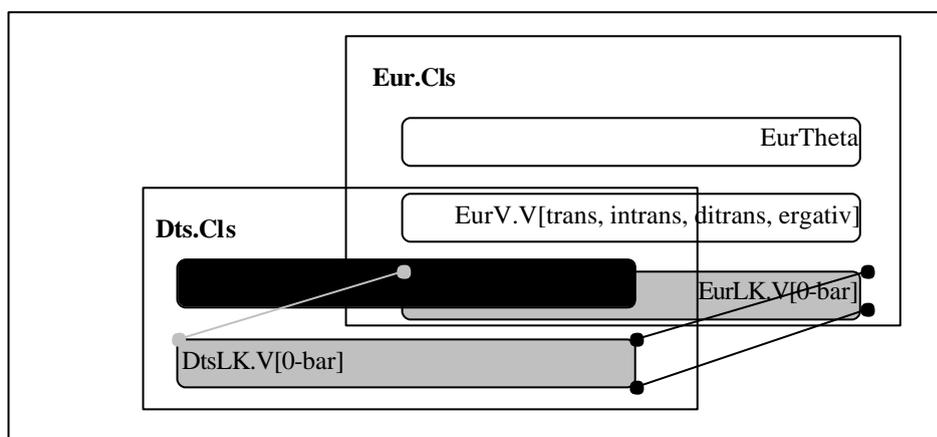
```

valence(VNull):-
  HasValue(VNull, 'intrans'),
  Extend(VNull, 'DtsV', 'VIntrans');
  HasValue(VNull, 'trans'),
  Extend(VNull, 'DtsV', 'VTrans'); ...

```

Ein Objekt der deutschen  $V^\circ$ -Klasse erbt aus der europäischen Klasse  $V^\circ$ , diese wiederum erbt mittels später Bindung aus den europäischen Valenzklassen. Schematisch ergibt sich damit folgendes Bild:

(Abb.4.3) Vererbungsschnittstelle zwischen Einzelsprachen und angenommener europäischer Metaebene



Intransitive Verben erben aus der Klasse `SetExternRole`, transitive außerdem noch aus `SetInternRole` usw. Eine Aufstellung der Vererbungshierarchie findet sich in *Anhang B*.

#### 4.3 Auxiliare und das LPS-Lexikon

Die Klassen aus Abschnitt (4.2) drücken allgemeines linguistisches Wissen aus. Idiosynkratisches linguistisches Wissen wird im Lexikon codiert. Das System LPS stützt sich dabei auf maschinenlesbare Lexika, die sogenannten LPS-Lexika. Diese enthalten Einträge, in denen dem *Signifiant*, der Ausdrucksseite des sprachlichen Zeichens, Merkmale (in Form von Attributwerten) zugeordnet werden, die einen Teil des *Signifié* enthalten. Das LPS-Lexikon nimmt im System LPS eine zentrale Rolle ein, da es beim Vorgang der maschinellen Übersetzung die einzige Schnittstelle zwischen den Einzelsprachen bildet (vgl. Rolshoven 1996). Dieser Abschnitt stellt in kurzer Form die Behandlung der Auxiliare in einem solchen Lexikon dar.

Abschnitt (2.1) definierte das Auxiliar als verbähnliche Kategorie, die auf der Verb-to-TAM-Chain angesiedelt ist. In (3.4) wurde dies als eine andere Basisposition der Auxiliare in der

Struktur interpretiert. Wie ist dem nun in der maschinellen Verarbeitung Rechnung zu tragen? Auxiliaren wird im LPS-Lexikon die Kategorie AUX zugewiesen. Diese Kategorie kann allerdings nicht wie andere Kategorien (bspw. V oder N) zu einer maximalen Projektion expandieren, sondern spezialisiert sich mittels später Bindung zur Laufzeit in eine der Kategorien  $V^\circ$  bzw.  $Pr^\circ$ . Welche der beiden Spezialisierungen gewählt wird, entscheidet sich während des Parsingprozesses des zu analysierenden Satzes: Ist das Auxiliar das einzige verbale Element im Satz, wird es als Vollverb gebraucht und spezialisiert sich aus der Klasse  $V[0\text{-bar}]$ . Tritt neben dem Auxiliar noch ein weiteres verbales Element in der Form des Partizips II oder des Infinitivs auf, so spezialisiert sich die Klasse AUX aus  $Pr[0\text{-bar}]$ .  $V^\circ$  und  $Pr^\circ$  sind in der Lage, maximale Projektionen auszubilden und damit den Strukturbaum aufzubauen. Der Lexikoneintrag eines Auxiliars hat z.B. folgende Form.<sup>74</sup>

```
ist, AUX[3-pers, sing, präs, sein]
```

Die Klasse AUX erbt durch späte Bindung aus den Klassen  $V^\circ$  oder  $Pr^\circ$ .<sup>75</sup>

```
CLASS AUX < V[0-bar]*, Pr[0-bar]*;
    checkKategorie(Self);
END;
```

#### 4.4 Zusammenfassung

Die in (4.2) beschriebenen Module wurden in das System GBX von Lalande (1997) integriert und an diesem evaluiert. Ihre Integration in das System LPS zu Zwecken der maschinellen Übersetzung ist für die nahe Zukunft geplant und dürfte ohne größere Probleme durchführbar sein.

Zum Abschluss soll die Interaktion der Module gezeigt werden. Zuerst wird in dem System GBX der transitive Satz *Er hat ihn getroffen* mittels des erzeugten linguistischen Wissens analysiert. Ein Screenshot (Abb. 4-4) zeigt die erzeugte Baumstruktur und eine Auswahl der Knoten, in denen Methoden bewiesen wurden. In  $V''$  wurde die interne Thetarolle und der Kasus AKKUSATIV vergeben, in  $Pr'$  die von der Verbmorphologie des Partizips blockierte externe Thetarolle deblockiert. In  $Pr''$  wurde diese externe Rolle dann an  $[Spec, Pr]$  vergeben.

<sup>74</sup> Aufgrund der relativ inkonsistenten Flexion und der begrenzten Anzahl der deutschen Auxiliare werden diese als Vollformen ins LPS-Lexikon eingetragen (vgl. Hühn 2000). Die Angabe der Merkmale in den eckigen Klammern hat keinen Anspruch auf Vollständigkeit.

<sup>75</sup> Denkbar ist auch, daß z.B. englische Modalverben sich zu einer im Strukturbaum höher stehenden Kategorie spezialisieren können. Dies müsste dann in der englischen Klasse *EngFK.AUX* festgelegt werden.







1. Wie läßt sich der Unterschied zwischen Vollverben und Auxiliaren (die durch ihre Position auf der Verb-to-TAM-Chain definiert sind) strukturell erfassen?
2. In welcher Position werden Auxiliare basisgeneriert, wohin bewegen sie sich und was sind die Gründe für diese Bewegungen?
3. In welchen Positionen und auf welche Art und Weise werden ihre Merkmale abgeglichen?
4. Wie kann der Blockierungs-/Deblockierungsmechanismus beschrieben werden?

Der Unterschied zwischen Vollverben und Auxiliaren wird so interpretiert, daß für diese Kategorien unterschiedliche Basispositionen angenommen werden. Die Basisposition des Vollverbs ist die Kopfposition der VP, die Basisposition des Auxiliars die Kopfposition der PrP. Von dieser Basisposition aus bewegen sich die Auxiliare im Strukturbaum aufwärts, um ihre funktionalen Merkmale mit denen der funktionalen Kategorien, die sie mittels ihrer Bewegung durchlaufen, abzugleichen. In der TP werden die Tempusmerkmale, in der AgrP die Agreementmerkmale abgeglichen. Zudem wird in einem finiten Kontext der Kasus NOMINATIV durch Agr<sup>o</sup> vergeben. Fordert der Subkategorisierungsrahmen des Vollverbs ein designiertes Argument, so wird dessen Status in der PrP überprüft: Wenn es blockiert ist und nicht durch das Auxiliar *haben* deblockiert wird, kann keine externe Thetarolle vergeben werden.

In LPS-Prolog werden Knoten im Strukturbaum durch Objekte von Knotenklassen repräsentiert. Diese Objekte erben aus allgemeineren Klassen, welche Methoden für den Merkmalabgleich oder die Merkmalszuweisung bereitstellen. Die Klassen sind in Modulen organisiert; diese Module können auf einzelsprachlicher und/oder sprachübergreifender Ebene angesiedelt sein. Vier der Module wurden in dieser Arbeit vorgestellt. Sie sind zuständig für den Abgleich von Agreement und Tempus, sowie für die Zuweisung von Kasus und Thetarollen.

Die ausgearbeiteten Module können in das LPS-System integriert werden. Damit sind sie nutzbar für computerlinguistische Anwendung, wie etwa die maschinelle Übersetzung. Module, die in dieser Arbeit einzelsprachlich für das Deutsche modelliert wurden, müssten dafür noch für andere Sprachen entworfen werden.

## Bibliographie

- Abney, S. (1987): "The english noun phrase in its sentential aspects." Unpublished diss., MIT, Cambridge Mass.
- Adams, M. (1987): "From Old French to the Theory of Pro-Drop." In: *Natural Language and Linguistic Theory 1*, 1-32.
- Baker, J. (1988): *Incorporation. A Theory of Grammatical Function Changing*. Chicago: University Press of Chicago.
- Bech, G. (1955/57): "Studien über das deutsche verbum infinitum." In: *Historisk-filologiske Meddelelser udgivet af Det kongelige Danske Videnskabernes Selskab*. Band 35,2 (1955) und 36,6 (1957). [<sup>2</sup>1983 Tübingen: Niemeyer].
- Bierwisch, M. (1989): "Verbkomplexe im Deutschen." Handout zum Vortrag.
- Bowers, J. (1993): "The Syntax of Predication." In: *Linguistic Inquiry 24*: 591-656.
- Bußmann, H. (<sup>2</sup>1990): *Lexikon der Sprachwissenschaft*. Stuttgart: Kröner.
- Burzio, L. (1986): *Italian Syntax - A Government-Binding Approach*. Dordrecht: Reidel.
- Chomsky, N. (1957): *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1965): *Aspects of the Theory of Syntax*. Cambridge, Mass: MIT Press.
- Chomsky, N. (1970): "Remarks on Nominalization." In: R. Jacobs & P. Rosenbaum: *English Transformational Grammar*. Waltham, Mass.: Ginn.
- Chomsky, N. (1981): *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. (1986): *Barriers*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1989): "Some Notes on Economy of Derivation and Representation." In: R. Freidin (ed.) (1991): *Principles and Parameters in Comparative Grammar*, Cambridge, Mass.: MIT Press: 417-454.
- Chomsky, N. (1995): *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Clocksins, W. F. & C. S. Mellish (<sup>3</sup>1987): *Programming in Prolog*. Berlin: Springer.
- Comrie, B. (1985): *Tense*. Cambridge: University Press.
- Dixon, R. (1979): "Ergativity." In: *Language 55*, 59-138.
- Dorffner, G. (1991): *Konnektionismus*. Stuttgart: Niemeyer.
- Dudenredaktion (<sup>4</sup>1998): DUDEN. *Grammatik der deutschen Gegenwartssprache*. Mannheim, Leipzig, Wien, Zürich: Duden-Verlag.

- Ehrich, V. & H. Vater (1989): "Das Perfekt im Dänischen und im Deutschen." In: W. Abraham & T. Janssen (eds.): *Tempus - Aspekt - Modus. Die lexikalischen und grammatischen Formen in den germanischen Sprachen*. Tübingen: Niemeyer (= Linguistische Arbeiten 237).
- Engel, U. (1996): *Deutsche Grammatik*. Heidelberg: Groos.
- Fodor, J. (1975): *The Language of Thought*. Harvard: University Press.
- Gazdar, G., E. Klein, G. Pullum & I. Sag (1985): *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell.
- Giorgi, A. & F. Pianesi (1991): "Toward a syntax of temporal representation." In: *Probus* 3: 187-213.
- Grevendorf, G. (1988): *Aspekte der deutschen Syntax. Eine Rektions-Bindungs-Analyse*. Tübingen: Narr (= Studien zur deutschen Grammatik 33).
- Haegeman, L. (1994): *Introduction to the Government & Binding Theory*. Oxford: Blackwell.
- Haegeman, L. (1998): "Verb Movement in Embedded Clauses in West Flemish." In: *Linguistic Inquiry* 29: 631-656.
- Haider, H. (1984): "Was zu haben ist und was zu sein hat." In: Abraham, W. (ed.) (1985): *Erklärende Syntax des Deutschen*. Tübingen: Narr.
- Haider, H. (1986): "Fehlende Argumente: Vom Passiv zur kohärenten Konstruktion." In: *Linguistische Berichte* 101: 3-33.
- Haider, H. & R. Rindler-Schjerve (1987): "The parameter of auxiliary selection: Italian-German contrasts." In: *Linguistics* 25, 1029-1055.
- Hartmann, R. & F. Stork (1972): *Dictionary of language and linguistics*. London: Applied Sciences Publishers Ltd.
- Heine, B. (1993): *Auxiliaries: cognitive forces and grammaticalization*. New York: Oxford University Press.
- Hühn, J. (2000): "DtsMorph.lex. Ein Versuch, die Flexion deutscher Substantive, Adjektive und Verben im Lexikon adäquat zu behandeln." Ms. Sprachliche Informationsverarbeitung, Universität zu Köln. URL: <http://www.spinfo.uni-koeln.de/~jhuehn/DtsMorph.pdf> (25.09.2001).
- Hühn, J. (2001): "Merkmalabgleich innerhalb der DP." Ms. Sprachliche Informationsverarbeitung, Universität zu Köln. URL: <http://www.spinfo.uni-koeln.de/~jhuehn/DP.pdf> (25.09.2001).
- Jackendoff, R. (1977): *X'-Syntax: A Study of Phrase Structure*. Cambridge, Mass: MIT Press.

- Jain A. & A. Waibel (1991): "Parsing with Connectionist Networks." In: Tomita (ed.): *Current Issues in Parsing Technologie*. Boston/Dordrecht/London: Kluwer Academic Publishers.
- Kayne, R. S. (1989): "Facets of Romance Past Participle Agreement." In: P. Benincá (ed.): *Dialect Variation and the Theory of Grammar*. Dordrecht: Foris, 85-103.
- Kayne, R. S. (1993): "Toward a Modular Theory of Auxiliary Selection." In: *Studia Linguistica* 47: 3-31.
- Kuroda, S.-Y. (1988). "Whether we agree or not: Rough ideas about the comparative syntax of English and Japanese." In: *Linguisticae Investigationes* 12: 1-47
- Lalande, J.-Y. (1997): *Verbstellung im Deutschen und Französischen. Unter Anwendung eines CAD-basierten Expertensystems*. Tübingen: Niemeyer (= Linguistische Arbeiten 365).
- Langer, H. (2001): *Parsing-Experimente. Praxisorientierte Untersuchungen zur automatischen Analyse des Deutschen*. Frankfurt a. M.: Lang.
- Larson, R. (1988): "On the double object construction." *Linguistic Inquiry* 19:355-391.
- Lenerz, J. (1995): "Klammerkonstruktionen." In: J. Jacobs, A. v. Stechow, W. Sternefeld, T. Vennemann (eds.): *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*. 2. Halbband; Berlin, New York: de Gruyter: 1266-1276.
- Lohnstein, H. (2000): *Satzmodus - kompositionell. Zur Parametrisierung der Modusphrase im Deutschen*. Berlin: Akademie Verlag (= studia grammatica 49).
- Meillet, A. (1912): "L'evolution des formes grammaticales." In: A. Meillet: *Linguistique historique et linguistique générale*. (2<sup>e</sup> 1921) Paris: Champion.
- Mössenböck, H. (1998): *Objektorientierte Programmierung in Oberon*. Berlin, Heidelberg: Springer.
- Newell, A. & H.A. Simon (1976): "Computer Science as Empirical Inquiry: Symbols and Search." In: *Communications of the Association for Computing Machinery* 19: 113-126.
- Pollock, J.-Y. (1989): "Verb movement, UG and the Structure of IP." In: *Linguistic Inquiry* 20: 365-424.
- Radford, A. (1997): *Syntax. A Minimalist Introduction*. Cambridge: University Press.
- Ramers, K.-H. (2000): *Einführung in die Syntax*. München: Fink.
- Reichenbach, H. (1947): *Elements of Symbolic Logic*. New York: Macmillan.
- Roberts, I. (1991): "Excorporation and Minimality." In: *Linguistic Inquiry* 22: 209-218.

- Rolshoven, J. (1987): "LPS. Eine linguistische Programmiersprache." In: U. Klenk, P. Scherber & M. Thaller (eds.): *Computerlinguistik und philologische Datenverarbeitung*. Hildesheim: Olms, 115-129.
- Rolshoven, J. (1991): "GB und sprachliche Informationsverarbeitung mit LPS." In: J. Rolshoven & D. Seelbach (eds.): *Romanistische Computerlinguistik. Theorien und Implementationen*. Tübingen: Niemeyer (= Linguistische Arbeiten 266), 133-158.
- Rolshoven, J. (1996): "Lexikalisches Wissen in der maschinellen Übersetzung." In: P. Blumenthal, G. Rovere & C. Schwarze (eds.): *Lexikalische Analyse romanischer Sprachen*. Tübingen: Niemeyer (=Linguistische Arbeiten 353), 85-100.
- Rolshoven, J. (2001): "Linguistische Objektorientierung." Ms. Sprachliche Informationsverarbeitung, Universität zu Köln.
- Rummelhard D. E. & J. L. McClelland (1986a): *Parallel Distributed Processing, Explorations in the Microstructure of Cognition, Vol 1: Foundations*. Cambridge, Mass.: MIT Press.
- Rummelhard D. E. & J. L. McClelland (1986b): "On Learning the Past Tenses of English Verbs." In: McClelland J. L. & Rummelhard D. E. (eds.): *Parallel Distributed Processing, Explorations in the Microstructure of Cognition, Vol 2: Psychological and Biological Models*. Cambridge, Mass.: MIT Press.
- Schmid, T. (2000): "Die Ersatzinfinitivkonstruktion im Deutschen." In: *Linguistische Berichte* 26: 325-351.
- Schoofs, T. (1994): "Kleine Geschichte der IP (mit besonderer Berücksichtigung der Diskussion um ihre Aufteilung)." Ms. Sprachliche Informationsverarbeitung, Universität zu Köln.
- Sportiche, D. (1988): "A Theory of floating Quantifiers and its Corrolaries for Constituent Structure." In: *Linguistic Inquiry* 19: 425-449.
- Stabler, E. (1992): *The Logical Approach to Syntax. Foundations, Specifications, and Implementations of Theories of Government and Binding*. Cambridge, Mass.: MIT Press.
- Stechow, A. v. & H. Nohl (1996): "Passiv." URL: <http://www2.sfs.nphil.uni-tuebingen.de/arnim/stechow-nohl.pdf> (12.5.2001)
- Steele, S. (1978): "The category AUX as a language universal." In: Greenberg (ed.): *Universals of human language*. Volume 3. Stanford: Stanford University Press: 7-45.
- Steele, S. (1999): "Auxiliaries." In: CK Brown & J. Miller (eds.): *Concise Encyclopedia of Grammatical Categories*. Cambridge: University Press.
- Vater, H. (1991b): *Einführung in die Zeit-Linguistik*. Hürth: Gabel (= KLAGE 25).
- Winograd, T. (1983): *Language as a cognitive Process. Vol. I: Syntax*. Reading, Mass: Addison-Wesley.

## Anhang A: Module in LPS-Prolog

**(\* EurAgreement.cls \*)**

```
(* 1. Kongruenz zwischen Subjekt und Finitum *)
CLASS SubjectAgr;
    agreement(Self, Other);
END;
CLASS FinitumAgr;
    agreement(Other, Self) Agr';
END;

(* 2. Kongruenz zwischen Objekt und Partizip *)
(* Klassen werden nur aus bestimmten europäischen Sprachen
geerbt (z.B. Französisch) *)
CLASS ObjectAgr;
    objAgreement(Self, Other);
END;
CLASS PartizipAgr;
    objAgreement(Other, Self) AgrO';
END.
```

**(\* EurAgreement.msp \*)**

```
(* Kongruenz besteht, wenn Subjekt und Finitum in Person und
Numerus unifizierbar sind. *)

agreement(Subj, Fin):-
    UnifyValues(Subj, Fin, 'Pers'),
    UnifyValues(Subj, Fin, 'Num').
```

**(\* DtsTempus.cls \*)**

```
(* Klassen zur Bestimmung des Satztempus *)
(* Verbparadigma kann Infinitum oder Spur des finiten Vollverb
unter V° sein, Finitum ist immer Finitum unter T°. So ist es
möglich, daß Self und Other mit dem gleichen Objekt
instanziiert sind, was allerdings die getrennte Behandlung von
finitem Vollverb auf der einen Seite und infinitem Vollverb +
finitem Auxiliar auf der anderen Seite überflüssig macht.*)

CLASS FinitumTense;
    setTense(Self, Other);
END;

CLASS VParaTense;
    setTense(Other, Self) Pr";
END.
```

**(\* DtsTempus.msp \*)**

setTense(Finitum, VePara):-

```
    HasValue(Finitum, 'praes'),
    HasValue(VePara, 'praes'),
    SetValue(Finitum, 'praesens');
```

```
    HasValue(Finitum, 'praet'),
    HasValue(VePara, 'praet'),
    SetValue(Finitum, 'imperfekt');
```

```
    HasValue(Finitum, 'praes'),
    HasValue(Finitum, 'aux1'),
    HasValue(VePara, 'part2'),
    SetValue(Finitum, 'perfekt');
```

```
    HasValue(Finitum, 'praet'),
    HasValue(Finitum, 'aux1'),
    HasValue(VePara, 'part2'),
    SetValue(Finitum, 'pqperfekt');
```

```
    HasValue(Finitum, 'praes'),
    HasValue(Finitum, 'aux2'),
    HasValue(VePara, 'inf'),
    SetValue(Finitum, 'futur1');
```

```
    HasValue(Finitum, 'praes'),
    HasValue(Finitum, 'aux2'),
    HasValue(VePara, 'infperf'),
    SetValue(Finitum, 'futur2');
```

**(\* EurCase.Cls \*)**

(\* CaseCheck überprüft, ob D" Kasus trägt, Asser bedeutet 'Assigner', Assed 'Assigned' \*)

```
CLASS CaseCheck;
    checkCase(Self);
END;
```

```
CLASS NomAsser;
    nominativ (Self, Other) Agr';
END;
```

```
CLASS NomAssed;
    nominativ (Other, Self);
END;
```

```
CLASS AkkAsser;  
    akkusativ (Self, Other) V';  
END;
```

```
CLASS AkkAssed;  
    akkusativ (Other, Self);  
END;
```

(\* Dativ und Genitiv werden lexikalisch zugewiesen, Genitiv evtl. auch strukturell an SpecD. Akkusativ kann auch z.T. lexikalisch zugewiesen werden \*)

```
CLASS SetDativ;  
    dativ (Self, Other);  
END;
```

```
CLASS Genitiv;  
    genitiv (Self, Other);  
END;
```

```
CLASS akkusativ;  
    lexAkkusativ (Self, Other);  
END;
```

#### (\* EurCase.msp \*)

(\* CaseCheck überprüft, ob D" Kasus aufweist oder PRO ist und damit keinen Kasus zugewiesen bekommen darf. Im Default-Fall wird versucht, DZwei lexikalischen Kasus zuzuweisen \*)

```
checkCase(DZwei):-  
    not(HasValue(DZwei, 'PRO')),  
    HasAttribute(DZwei, 'Kasus');  
    HasValue(DZwei, 'PRO'),  
    not(HasAttribute(DZwei, 'Kasus'));  
    dativ (Other, DZwei);  
    genitiv (Other, DZwei);  
    akkLex (Other, DZwei).
```

(\* Nominativ wird von finitem Agr<sup>o</sup> an SpecAgr zugewiesen. Im infiniten Kontext darf sich in SpecAgr nur eine Spur oder PRO befinden \*)

```
nominativ (AgrNull, DZwei) :-  
    HasAttribute(AgrNull, 'FinVerbPara'),  
    SetValue(DZwei, 'nom');  
    not(HasValue(AgrNull, 'FinVerbPara')),  
    (HasAttribute(DZwei, 'Index'); HasValue(DZwei, 'PRO')).
```

(\* Akkusativ wird von transitiven V° an SpecV zugewiesen, es sei denn, dort befindet sich eine NP-Spur \*)

```
akkusativ (VNull, DZwei) :-  
    (HasValue(VNull, 'trans');HasValue(VNull, 'ditrans')),  
    not(HasAttribute(DZwei, 'Index')),  
    SetValue(DZwei, 'akk').
```

```
dativ (Alpha, Beta) :- True.
```

```
genitiv (Alpha, Beta) :- True.
```

```
akkLex (Alpha, Beta) :- True.
```

**(\* DtsTheta.Cls \*)**

```
CLASS Deblocker;  
    deblock(Self, Other),f;  
END;
```

```
CLASS Blocked;  
    deblock(Other, Self),f;  
END.
```

**(\* DtsTheta.msp \*)**

(\* Attribut ThetaBlock wird im Kontext von 'haben' überschrieben \*)

```
deblock(Auxiliar,Verb):-  
    HasValue(Verb, 'blockON'),  
    HasValue(Auxiliar, 'haben'),  
    SetValue(Verb, 'blockOFF').
```

**(\* EurTheta.cls \*)**

(\* Vergabe der externen Rolle ist fakultativ, da sie evtl. blockiert ist. interne Rolle muß immer vergeben werden, wenn von Verb gefordert. CheckTheta überprüft, ob an Argumentpositionen Thetarollen vergeben wurden \*)

```
CLASS SetExternRole;  
    externRole(Self,Other) Pr',f;  
END;
```

```
CLASS SetInternRole;  
    internRole(Self,Other) Pr';  
END;
```

```

CLASS SetBoth;
    externRole(Self,Other) Pr',f;
    internRole(Self,Other) V'.

CLASS GetExternRole |[ThetaExt];
    externRole(Other,Self);
    checkTheta(Self);
END;

CLASS GetInternRole |[ThetaInt, ThetaExt];
    internRole(Other,Self);
    checkTheta(Self);
END;

(* EurTheta.msp *)

internRole(Verb,Argument):-
    UnifyValues(Verb, Argument, 'ThetaInt').

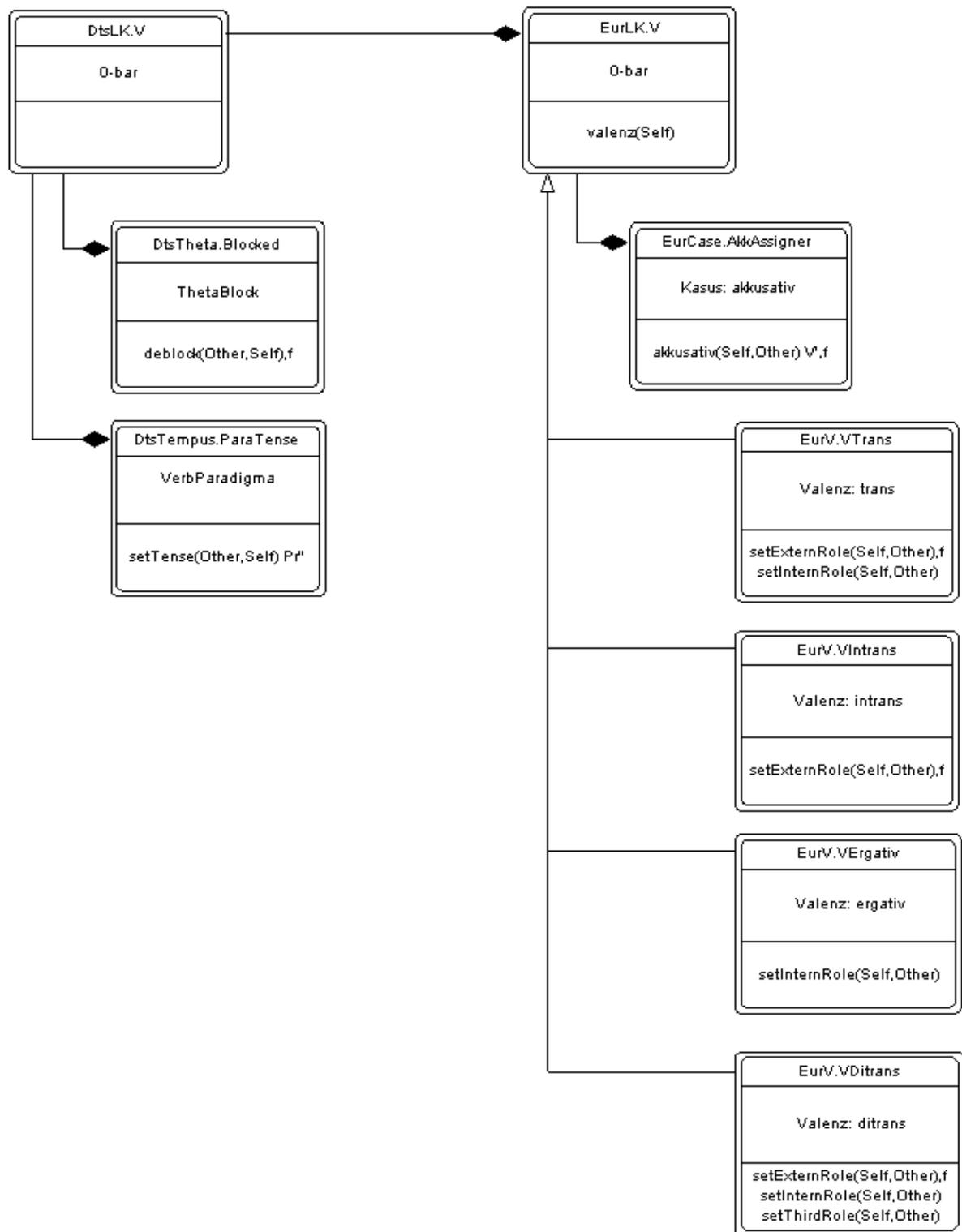
externRole(Verb, Argument):-
    not(HasValue(Verb,'blockON'));
    UnifyValues(Verb, Argument, 'ThetaExt').

checkTheta(Empfaenger):-
    HasAttribute(Empfaenger, 'Theta').

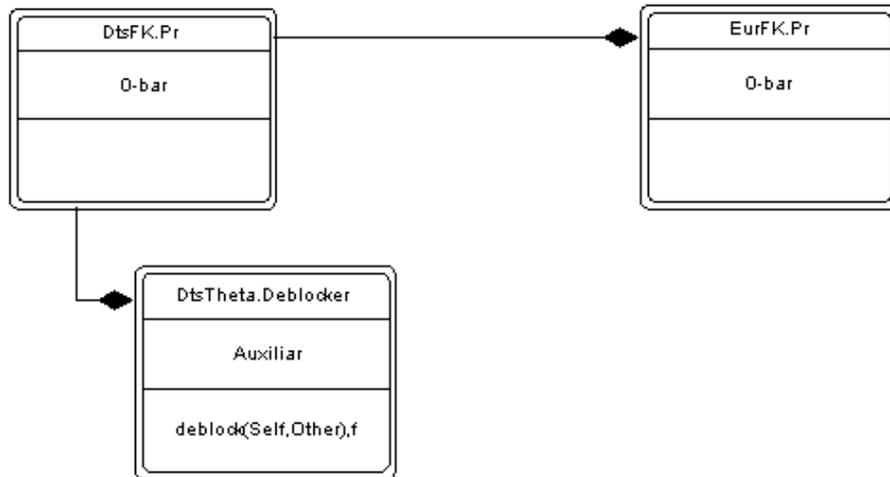
```

## Anhang B: Vererbungshierarchie der lexikalischen und funktionalen Verbalklassen

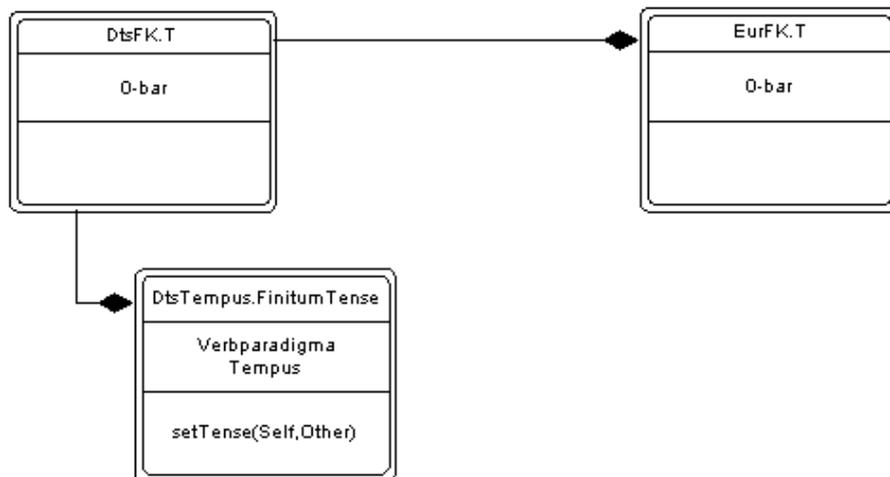
Abb. B1: Die V<sup>o</sup>-Klassen



**Abb. B2: Die Pr<sup>o</sup>-Klassen**



**Abb. B3: Die T<sup>o</sup>-Klassen**



**Abb. B4: Die Agr<sup>o</sup>-Klassen**

