Kapitel 7

Das System TECHDOC

7.1 Zur Architektur des TECHDOC–Systems

In der Gesamtarchitektur des Systems (vgl. Abb. 7.1) lassen sich grob drei Typen von Systemkomponenten unterscheiden:

- die verschiedenen Wissensbasen (terminologische Wissensbasis, assertionale Wissensbasis) als Kern,
- Grammatiken und Lexika für die jeweiligen Zielsprachen (derzeit: Deutsch, Englisch, Französisch),
- Hilfsprogramme (Benutzeroberfläche, Formatter, u.ä.)

Im einzelnen:

Derzeit existieren drei Grammatiken. Diese sind allerdings nicht völlig getrennt voneinander, sondern benutzen solange Systeme (im Sinne der systemisch–funktionalen Grammatik) gemeinsam, solange keine Unterscheidungen gemacht werden müssen.

Beim Durchlaufen der Systemnetzwerke für jeweils eine Sprache wird u.a. auf Information aus dem sogenannten 'upper model' zugegriffen [Bateman 1990]. Dies ist eine Taxonomie allgemeiner Konzepte, die sprachliche Unterschiede reflektieren. Unter die Konzepte des 'upper model' werden die Konzepte des jeweiligen Gebietsmodells subsumiert.

Weiter wird ein Lexikon konsultiert, das Konzepten im Gebietsmodell und 'upper model' Lexikalisierungen zuordnet. Wo dieser Zugriff, der eine direkte Entsprechung von Konzepten und Lexemen voraussetzt, nicht flexibel genug ist, werden ergänzend auch Lexikalisierungsregeln (vgl. Abschnitt 7.4.3) herangezogen.

In der Sprechweise der KL–ONE–Tradition sind das 'upper model' und das Gebietsmodell die terminologische Wissensbasis (oder T–Box) und die Instanzen, die zusätzlich gebildet werden, bilden die sogenannte assertionale Wissensbasis oder A–Box. Zentrale Instanzen sind die Planinstanzen, die jeweils einen möglichen War-

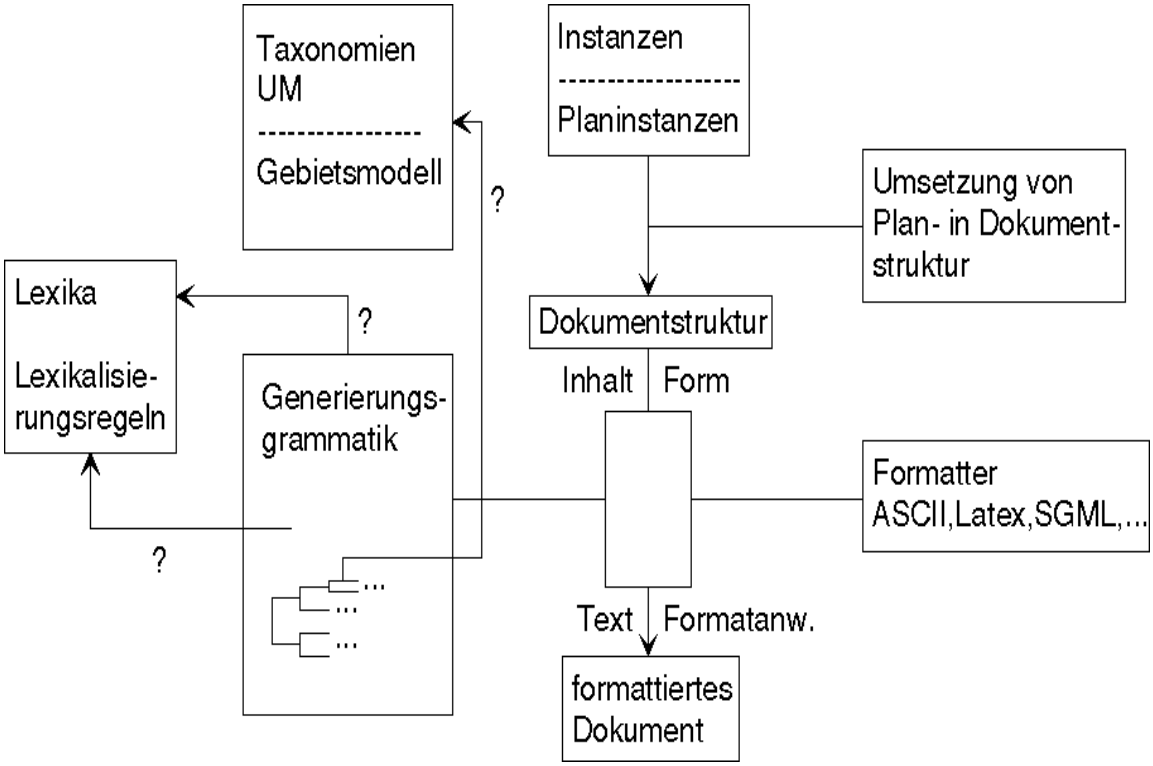

Abbildung 7.1: TECHDOC: Systemarchitektur

tungsvorgang repräsentieren. Planinstanzen werden umgesetzt in Dokumentstrukturen, die dann sowohl inhaltliche wie formale Aspekte eines Dokuments beschreiben. Die inhaltlichen Spezifikationen werden von der Generierungsgrammatik in linguistische Einheiten umgesetzt, die formalen Elemente mit Hilfe des TECHDOC–Formatter (vgl. 8.2) in Formatanweisungen bzw. Textauszeichnungen ('mark-up'), welche die Dokumentstruktur dann im gedruckten oder am Bildschirm ausgegebenen Dokument auch auf der Layout–Ebene unterstreichen. Das zu unterstützende Formatierprogramm ist wählbar. Derzeit werden unterstützt ein leicht formatiertes ASCII, SGML (vgl. 8.2.2), LaTeX und ZMACS. Weitere Formatierprogramme können mit geringem Aufwand ebenfalls versorgt werden, da eine objektorientierte Architektur (vgl. auch Abschnitt 8.2) benutzt wurde. Die generierten Textstücke und die Formatanweisungen werden wieder kombiniert in eine Quelldatei für das gewählte Formatierprogramm. Damit läßt sich dann ein formatiertes Dokument erzeugen und drucken.

Die gewählte Architektur ist modular genug, um eine Reihe von Varianten sehr leicht realisierbar zu machen. Es ist z.B. nicht zwingend, daß systemische Grammatiken verwendet werden. Grammatiken für andere Sprachen oder in anderen Formalismen sind dadurch anschließbar, daß sie die Schnittstellen zum 'upper model' und Gebietsmodell unterstützen und die dort modellierten Unterscheidungen nutzen sollten. Ferner sollten sie Eingabestrukturen in der Notation der SPL (Sentence Plan

Language [Kasper 1989], vgl. Anhang B) erlauben.

Für die Umsetzung von Planstrukturen in Dokumentstrukturen ist ein eigenes Modul zuständig, das ohne Rückwirkung auf die Grammatik ausgetauscht werden kann.

7.1.1 Die Entwicklung der Demonstrations–Versionen

Ein Aspekt des Vorgehens im Projekt TECHDOC ist der, daß hier 'Reverse Engineering' betrieben wurde: Ausgehend von vorgefundenen Texten wurde versucht, die verschiedenen Komponenten des Systems so zu gestalten, daß sie in der Lage sind, Äquivalente dieser Texte zu reproduzieren. Dabei ging es nicht darum, die Texte in ihrer vorgefundenen Form lediglich reproduzieren zu können, vielmehr sollte das System in einer solchen Weise gestaltet werden, daß die Übertragung auf ähnlich gelagerte Themen und verwandte Texte aufgrund der getroffenen Design–Entscheidungen leicht möglich ist.

Implementationen und Implementationsexperimente erfolgten im Rahmen von TECHDOC durch den Autor und durch von ihm geleitete Diplom-, Magister- und Praktikumsarbeiten (vgl. Anhang E). Vom Leistungsspektrum und vom Schwerpunkt der Aufgabenstellung her lassen sich zwei Demonstrations–Versionen deutlich unterscheiden. Bei der ersten Demonstrations–Version standen Fragen der grammatischen Abdeckung und der Umsetzung von Dokumentstrukturen in formatierte Texte im Vordergrund. Ergänzend wurden hier auch die Benutzeroberfläche eines Arbeitsplatzes zur Dokumentproduktion (z.B. für einen technischen Redakteur) entwickelt und implementiert (vgl. 8.3).

Der Schwerpunkt bei den Arbeiten zur zweiten Demonstrations–Version war, die Gebietswissensbasis zu reorganisieren und auszuweiten. Dabei wurde auch damit experimentiert, wie diese auch für andere Zwecke, wie Simulation und Deduktion, genutzt werden kann. Ebenfalls zur zweiten Demonstrations–Version gehört die automatische Umsetzung von Planstrukturen in Dokumentstrukturen und die Entwicklung des Autorenwerkzeuges, das es erlaubt, ohne Kenntnis der Syntax der Wissensrepräsentationssprache Planstrukturen aufzubauen und mit Inhalt zu füllen, die dann als Basis für die Textgenerierung herangezogen werden.

Das Vorgehen war also auch in dem Sinne 'Reverse Engineering', daß von der sprachlichen Oberfläche ausgehend bzw. von den Phänomenen, die als wichtig für die Reproduktion eingestuft wurden, zunächst die linguistische Abdeckung angestrebt wurde und dann deren systematische Abstützung in einer entsprechenden Wissensbasis folgte.

Die grammatikbezogenen Arbeiten für die erste Demonstrations–Version umfaßten Erweiterungen der englischen Grammatik, die durch das Fragment erforderlich waren, sowie die Neuentwicklung eines entsprechenden Fragments einer deutschen und einer französischen Grammatik (vgl. 7.2).

7.1.2 Schwache vs. starke Multilingualität

Die Gesamtarchitektur von TECHDOC macht keine Annahmen über die Ausgestaltung der Generierungskomponenten für die Einzelsprachen in einem System zur multilingualen Dokumentgenerierung.

Im Prinzip können hier sprachspezifische Generatoren eingesetzt werden, die auf sehr unterschiedlichen linguistischen Theorien basieren und sich in ihren internen Verarbeitungsmechanismen gravierend unterscheiden. Verbindende Anforderung ist lediglich, daß sie die Wissensbasen und die gemeinsame Notation für die auszudrückenden Inhalte nutzen können. Eine solche Konstellation könnte als *„schwache Multilingualität"* bezeichnet werden.

Im Kontrast dazu läßt sich als *„starke Multilingualität"* bezeichnen, wenn sich bei einem System die einzelsprachliche Generierung für alle beteiligten Sprachen auf dieselbe linguistische Theorie und einen identischen Abarbeitungsmechanismus für auf diese Theorie bezogene Grammatiken stützt.

In TECHDOC ist in diesem Sinne starke Multilingualität realisiert. Dies ist aus der Sicht der theoretischen Linguistik sicherlich wünschenswert; aus der Sicht einer anwendungsorientierten Computerlinguistik erscheint uns dies aber nicht zwingend: Implementierte Grammatikformalismen sind meist nicht in dem Sinne universell, daß sie sich ohne Modifikation auf eine Vielzahl von Sprachen übertragen lassen. In Implementationen fließen oft Annahmen über Regularitäten ein, die als invariant angesehen werden. Diese gelten dann für diejenige Sprache, mit der die Entwicklung stattgefunden hat, besitzen aber nicht notwendigerweise universelle Gültigkeit.[1]

7.1.3 Wiederverwendung linguistischer Ressourcen

Für jedes komplexe Softwareprojekt stellt sich die Frage, inwieweit für Teilaufgaben im angestrebten Gesamtsystem bereits vorhandene Ressourcen (z.B. Komponenten oder Datenbestände) genutzt werden können. Dies gilt in gleicher Weise für Systeme der Computerlinguistik.[2] Zu prüfen ist dabei, ob eine direkte Übernahme in das eigene System möglich ist oder ob zunächst Aufwand für die Anpassung an die Erfordernisse des eigenen Projektes zu leisten ist, wie hoch dieser ggf. sein wird und in welchem Verhältnis er zu dem Aufwand steht, der für eine Eigenentwicklung zu veranschlagen ist. In die Entscheidung fließt ebenfalls mit ein, ob es sich um einen zentralen Aspekt des eigenen Projekts oder um unterstützende Funktionalität handelt.

[1] Im Jahre 1988 hat sich eine ganze Tagung — der „Wiener Workshop Wissensbasierte Sprachverarbeitung" [Trost 1988] — mit der Thematik beschäftigt, inwieweit primär für das Englische entwickelte computerlinguistische Systeme und Ansätze auf Deutsch übertragbar sind und welche Probleme dabei auftreten.

[2] In der europäischen Gemeinschaft ist erkannt worden, daß die Verfügbarkeit linguistischer Ressourcen — z.B. maschinenlesbarer Lexika — von ökonomischer Bedeutung ist, und es wurden entsprechende Initiativen eingeleitet.

Es ist kein Ziel des TECHDOC–Projekts, Grammatiken zu entwickeln. Unser Interesse ist es, Grammatiken für die Generierung von Dokumenten zu nutzen. Wir haben uns daher bemüht, vorhandene linguistische Ressourcen optimal einzusetzen; Eigenentwicklung haben wir nur dort betrieben, wo keine solchen Ressourcen zur Verfügung standen.

In die in TECHDOC für Englisch verwendete Grammatik NIGEL (vgl. 7.2.1) ist ein beträchtlicher Entwicklungsaufwand eingeflossen. Diese englische Grammatik gilt als diejenige mit dem größten implementierten Fragment (dennoch besteht immer noch Bedarf an Erweiterungen, vergl. 7.1.3).

Vergleichbar umfangreiche grammatische Ressourcen für Deutsch und Französisch standen und stehen bisher nicht zur Verfügung.[3]

Wir haben uns daher dafür entschieden, NIGEL für die Generierung von Englisch einzusetzen und von NIGEL ausgehend selbst Fragmente einer deutschen bzw. französischen Grammatik zu implementieren, welche die in den Handbuchtexten angetroffenen linguistischen Konstruktionen behandeln können.

Daß sich auf diese Weise eine zunächst monolinguale linguistische Ressource in einen multilingualen Kern und drei einzelsprachliche Ergänzungen umwandeln läßt, mag zunächst verwundern. Plausibel wird das Vorgehen, wenn man sich aber vor Augen führt, wie systemische Grammatiken organisiert sind: Mit funktionalen (paradigmatischen) Fragen wird in den allgemeinen Systemen des Netzwerkes begonnen (z.B.: Welcher Sprechakt liegt vor? Welche semantische Rolle ist auszudrücken? Mit welcher Funktion kann dies geschehen?), Entscheidungen über die syntagmatische Organisation (z.B.: Lexikalisierung: Welches Funktionswort ist zu verwenden?, lineare Präzedenz: Wie sind die Konstituenten anzuordnen?) werden erst in den spezifischen Systemen kurz vor Beendigung des Grammatikdurchlaufs getroffen. Bei allen Unterschieden gibt es nun aber gerade in den funktionalen Aspekten ein hohes Maß an Parallelität und Gemeinsamkeit bei den hier betrachteten Sprachen (die zudem ja auch „nahe“ Verwandte sind).

Zur technischen Umsetzung

Um NIGEL in ein multilinguales System umzuwandeln, sind wir wie folgt vorgegangen:

- Ausgangspunkt war das uns vorliegende NIGEL. Wenn Systeme und andere Datenstrukturen des Systems nicht sprachspezifisch differenziert werden mußten, blieben sie unverändert und gehören nun zum Kernbereich.

- Waren sprachspezifische Differenzierungen erforderlich (z.B. weil bestimmte Phänomene in einer Sprache zusätzlich auftauchen oder nicht bzw. auf andere Weise realisiert werden), so wurden die benötigten Systeme, Auswahlfunktionen und Systemfragen in den einzelsprachbezogenen Erweiterungen definiert bzw. redefiniert.

[3] vergl. aber die Arbeiten im KOMET–Projekt bzw. in Sydney [Bateman et al. 1991].

- Eine weitere Maxime für die Arbeiten zur Erweiterung von PENMAN im Hinblick auf Mehrsprachigkeit war, das für die monolinguale Version geschriebene komplexe Programm zur Grammatiktraversierung möglichst unverändert beibehalten zu können.

Eine mögliche Alternative wäre gewesen (so wird es in den Arbeiten in Sydney für das sog. multilinguale PENMAN getan [Bateman et al. 1991]), innerhalb von Systemen, Choosers, Inquiries usw. Code zu erlauben, der im Hinblick auf unterschiedliche Zielsprachen konditionalisiert ist (z.B. durch die Verwendung regulärer Ausdrücke, um die Gültigkeit bzw. Nichtgültigkeit von Anweisungen für bestimmte Sprachen auszudrücken).

Um nicht Gefahr zu laufen, ein ggf. schwerer wartbares und nur mit größerem Aufwand erweiterbares System zu kreieren, haben wir uns für die andere Alternative entschieden, die zudem den Vorteil hatte, daß sich die Eingriffe in die Grammatik–Traversierung in engem Rahmen halten ließen.

Der Wechsel zwischen einzelsprachlichen Ausprägungen der von uns verwendeten Grammatik erfolgt durch gezieltes Aktivieren sprachspezifischer Systeme für die aktuelle Zielsprache und Deaktivieren „konkurrierender Systeme“ der nicht benötigten Sprachen. Konkurrierend sind in diesem Zusammenhang Systeme mit identischen Ausgabemerkmalen. Die Gestaltung des PENMAN–System verlangt, daß jedes Merkmal nur in höchstens einem aktiven System als Ausgabemerkmal verwendet werden darf. Das De-/Re-Aktivieren geschieht durch Umsetzen binärer Merkmale und ist vom Zeitaufwand her vertretbar.

Aufgrund anderer Erweiterungen ist es nun möglich, textuelle Phänomene zu behandeln und satzübergreifenden Kontext auszunutzen (s.a. 7.3). In der ursprünglichen Form ist NIGEL ein reiner Satzgenerator, der allerdings innerhalb komplexer Sätze Pronominalisierung durchführen kann. Mit Hilfe des in TECHDOC implementierten Verwaltungsmechanismus für Kontextinformation ist ein über Satzgrenzen hinausreichendes Pronominalisieren nun ebenfalls möglich.

In der vorliegenden Form hatte PENMAN seine Generierungsentscheidungen im wesentlichen von Subsumptionsinformation in der Klassenhierarchie der T–Box abhängig gemacht. Einige sprachliche Phänomene lassen sich aber nur behandeln, wenn auf Wissen über Instanzen zugegriffen werden kann. So spielen bei der Entscheidung über definite Referenz (vgl. 7.3.1) Teil-von-Beziehungen zwischen Instanzen eine wichtige Rolle (z.B.: „Ist das aktuell in Rede stehende Objekt Teil eines anderen bereits vorerwähnten und im Fokus stehenden Objekts, z.B. der Motor als Teil des Autos?“). Um solche Informationen aus der A–Box von LOOM verwalten und nutzen zu können, war es erforderlich, die Zugriffsmechanismen von NIGEL in die Wissensbasis zu erweitern.

7.2 Die verwendeten Grammatiken

7.2.1 Die englische Grammatik

Als englische Grammatik wurde die NIGEL–Grammatik (New Implementation of a Grammar for the English Language) des PENMAN–Systems übernommen [PENMAN]. NIGEL wurde seit ca. 1982 am Information Sciences Institute (ISI), Marina del Rey, entwickelt und gilt als die Implementation einer englischen Grammatik mit der größten Abdeckung. Dennoch zeigte sich sehr bald, daß Erweiterungen für typische Strukturen in Handbuchtexten erforderlich waren. Beispiele hierfür:

- Zu den notwendigen Erweiterungen gehörte die Möglichkeit, Adjektive, Adverbien und Präpositionalphrasen zu intensivieren [Pan, Rösner 1992] ; Beispiele:
  - 'A plug that is *too* loose can be very hot.'
  - 'It can damage the engine *very* easily.'
  - 'The level has dropped *significantly* below the lower mark.'
- Die Positionierung von Adverbien und Adverbialphrasen war nicht flexibel genug. NIGEL kannte dafür nur eine Position, nämlich die nach dem finiten Verb.
- Sogenannte 'secondary predication' war ein weiteres Erfordernis in Handbuchtexten; Beispiele hierfür:
  - 'Check the engine oil *with the ignition key off.*'
  - 'Check the engine condition *with the engine running.*'
  - 'Check the engine oil *with the car parked on level ground.*'

Techniken in NIGEL

Im folgenden werden wir beschreiben, wie einige der für die englische Grammatik notwendigen Erweiterungen vorgenommen wurden. Dazu ist es hilfreich, mehr in die Details von NIGEL einzudringen.

Die Grundidee der systemisch–funktionalen Grammatik und damit auch die Grundidee ihrer Implementation in NIGEL ist die, daß eine Grammatik ein Netzwerk sogenannter „Systeme" darstellt. Jedes System hat Einstiegsbedingungen (input conditions), die festlegen, wann es aktiviert werden kann. Aufgabe des Systems ist, ein Merkmal als Ausgabe zu bestimmen und ggf. zusätzlich Realisierungsanweisungen auszuführen, die durch die getroffene Wahl motiviert sind. Gewählte Merkmale können dann zur Aktivierung anderer Systeme führen (nämlich solcher, bei denen diese Merkmale Teil der Einstiegsbedingung sind).

In Systemen kann ein kleiner Satz von Realisierungsanweisungen verwendet werden; die meisten davon operieren auf Funktionen im Sinne der systemisch–funktionalen Grammatik. Manche Funktionen sind stärker auf die Semantik hin orientiert, andere eher syntaktisch ausgerichtet. Zu den letzteren gehört beispielsweise die Funktion 'SUBJECT', die in jedem Satz vorhanden ist. Mit der Funktion SUBJECT können aber semantisch unterschiedliche Funktionen wie 'ACTOR', 'SENSER', usw. realisiert werden. Die Identifikation unterschiedlicher Funktionen erfolgt mit der Realisierungsanweisung 'CONFLATE', mit der ausgedrückt wird, daß die beiden Argumente identifiziert werden sollen. Mit der Operation 'ORDER' wird festgelegt, daß das erste Argument in der endgültigen Ordnung der Funktionen unmittelbar vor dem zweiten aufzutauchen hat. Weniger strikt ist die Ordnungsbedingung bei der Realisierungsanweisung 'PARTITION'. Diese fordert nur, daß das erste Argument links vom zweiten anzuordnen ist, ohne unmittelbare Nachbarschaft zu verlangen.

Wird die Grammatik um ein System erweitert, so sind die folgenden Überlegungen relevant: Wie ist die zugehörige Auswahlfunktion (der sog. 'chooser') zu gestalten, d.h., welche unterschiedlichen Ergebnisse und damit Ausgabemerkmale des Systems kann der Entscheidungsprozeß haben und wie erfolgt die Entscheidung ?

Für die verschiedenen Ausgänge ist dann jeweils folgendes zu entscheiden:

- Muß eine neue grammatische Funktion eingeführt werden?
- Kann für diese grammatische Funktion gegebenenfalls schon eine Präselektion vorgenommen werden (d.h. können Merkmalswerte festgelegt werden, die andernfalls erst beim Grammatikdurchlauf auf der nächsten Verarbeitungs–Stufe ausgewählt würden)?
- Wie ist eine ggf. neu eingeführte grammatische Funktion relativ zu anderen Funktionen anzuordnen?

Wir diskutieren dies am Beispiel der Intensifikation von Adverbien (vgl. Abbildung 7.2) und dem zugehörigen adverb-intensification-chooser. Innerhalb des adverb-intensification-chooser werden zwei Systemfragen, sog. 'inquiries', verwendet. NIGEL unterscheidet bei solchen Systemfragen zwischen sogenannten „identifizierenden" (identifying operators) und „fragenden" Operatoren (asking operators). Mit identifizierenden Systemfragen werden Assoziationen zwischen grammatischen Funktionen und denjenigen Elementen der Wissensbasis hergestellt, die durch diese grammatischen Funktionen ausgedrückt werden sollen. Die fragenden inquiries greifen auf die Wissensbasis zu, um hier die notwendige Information für die Entscheidung in der Auswahlfunktion zu erhalten.

Im Beispiel geht es um die Frage, ob die Funktion 'CIRCUMSTANCE' intensiviert werden soll. Es gibt darauf zwei mögliche Antworten: 'intensified' und 'not intensified'. Wird intensiviert, so wird die grammatische Funktion 'INTENSIFIER' assoziiert mit dem Element der Wissensbasis, das durch die identifizierende

```
(system
        :name ADVERB-INTENSIFICATION
        :inputs
                Adverbial-Group
        :outputs
                ((0.5 INTENSIFIED-ADV (INSERT INTENSIFIER)
                     (PRESELECT INTENSIFIER ADVERBIAL-GROUP)
                     (PARTITION INTENSIFIER Circumstance))
                (0.5 NOTINTENSIFIED-ADV))
        :chooser ADVERB-INTENSIFICATION-CHOOSER
        :selector CHOICE-MASTER
        :old-feature INTENSIFIED
        :region ADJECTIVAL-COMPARISON
        :metafunction IDEATIONAL)
(chooser
        :name ADVERB-INTENSIFICATION-CHOOSER
        :definition
                ((ASK (INTENSIFY-ADJECTIVE-Q Circumstance)
                      (INTENSIFIED (IDENTIFY INTENSIFIER
                                             (INTENSIFIER-ID Circumstance))
                                   (CHOOSE INTENSIFIED-ADV))
                      (NOTINTENSIFIED (CHOOSE NOTINTENSIFIED-ADV))))
        :comments ...
        :editor ...
        :date ...)
```

Abbildung 7.2: System ADVERB-INTENSIFICATION und zugehöriger Chooser

inquiry 'intensifier-id' zurückgegeben wird. Ausgabemerkmal des Systems 'adverb-intensification' wird in diesem Fall 'intensified-adv', im anderen Fall 'not-intensified-adv'.

Wenn intensiviert wird, wird in die entstehende grammatische Struktur eine Funktion INTENSIFIER eingeführt (INSERT INTENSIFIER). Es wird bestimmt (präselektiert), daß diese als Adverbialgruppe zu realisieren ist (PRESELECT INTENSIFIER ADVERBIAL-GROUP) und sie wird links von 'CIRCUMSTANCE' angesiedelt (PARTITION INTENSIFIER Circumstance).

Offensichtlich analog geschieht das für die Intensifikation von Präpositionalphrasen (vgl. Abbildung 7.3); hierfür sorgt das System 'pp-intensification' mit seinem chooser 'pp-intensification-chooser' (da NIGEL verlangt, daß ein Merkmal nur von genau einem System gewählt werden kann, müssen unterschiedlich benannte Merkmale kreiert werden, also für Adverb–Identifikation 'intensified-adv', für die Intensifikation von Präpositionalphrasen 'intensified-pp' und für Adjektivgruppen 'intensified-adj').

Systeme und Auswahlfunktionen werden, wie in den Beispielen ersichtlich, deklarativ definiert. In den Inquiries wird dann allerdings direkt Lisp–Code verwendet.

```
(system
        :name PP-INTENSIFICATION
        :inputs
                Prepositional-Phrase
        :outputs
                ((0.5 INTENSIFIED-PP (INSERT INTENSIFIER)
                     (PRESELECT INTENSIFIER ADVERBIAL-GROUP)
                     (PARTITION INTENSIFIER MINORPROCESS))
                (0.5 NOTINTENSIFIED-PP))
        :chooser pp-INTENSIFICATION-CHOOSER
        :selector CHOICE-MASTER
        :old-feature INTENSIFIED
        :region ADJECTIVAL-COMPARISON
        :metafunction IDEATIONAL)
(chooser
        :name PP-INTENSIFICATION-CHOOSER
        :definition
                ((ASK (INTENSIFY-ADJECTIVE-Q Minirange)
                        (INTENSIFIED (IDENTIFY INTENSIFIER
                                                (INTENSIFIER-ID Minirange))
                                        (CHOOSE INTENSIFIED-PP))
                        (NOTINTENSIFIED (CHOOSE NOTINTENSIFIED-PP))))
        :comments ...
        :editor ...
        :date ...)
```

Abbildung 7.3: System PP-INTENSIFICATION und zugehöriger Chooser

Die Interpretation der System– und Chooser–Definitionen erfolgt durch ein umfängliches, in Lisp geschriebenes Abarbeitungsprogramm für die NIGEL–Grammatik.

Eine Ergänzung noch zur 'pp-intensification': In der systemischen Terminologie wird eine Präpositionalphrase durch zwei grammatische Funktionen konstituiert, nämlich durch 'MINORPROCESS' und 'MINIRANGE'. 'MINORPROCESS' deckt die Präposition ab und 'MINIRANGE' die Nominalphrase, die auf die Präposition folgt. Im 'pp-intensification-chooser' wird die Frage, ob intensiviert werden soll oder nicht, anhand von Informationen entschieden, die von Elementen der Wissensbasis zu erhalten sind, die mit der 'MINIRANGE' assoziiert sind. Im 'pp-intensification-system' wird dann die Ordnungsbeziehung mit dem 'MINORPROCESS' definiert (durch die Realisierungsanweisung (PARTITION INTENSIFIER MINORPROCESS)), da der INTENSIFIER ja vor der Präposition zu plazieren ist.

Adverbien und adverbiale Ergänzungen

Eine ganze Reihe von Ausweitungen und Modifikationen der englischen Grammatik hatten mit Adverbien und adverbialen Ergänzungen zu tun. So wurde damit experimentiert, unterschiedliche Positionen von Adverbien generieren zu können. In der

ursprünglichen Form konnte die Grammatik das Adverb nur unmittelbar nach dem Hauptverb plazieren, was in vielen Fällen nicht korrekt ist.

Die gefundene Lösung basiert auf einer zunächst syntaktisch orientierten Klassifikation. Dabei werden unterschieden:

- Adverbien, die nur am Anfang des Satzes stehen können (z.B. *surprisingly, unfortunately, honestly*).

- Adverbien, die sowohl eine Position am Satzanfang als auch eine Position vor dem Hauptverb einnehmen können (z.B. *personally, possibly*).

- Eine andere Gruppe ist immer in der Position vor dem Hauptverb, aber nach dem Subjekt (z.B. *still, almost, just*).

- Noch eine weitere Gruppe kann die Position nach der Verb–Phrase einnehmen (z.B. *easily, quickly, securely*). Diese können mit leichten Bedeutungsunterschieden aber auch an anderen Positionen — nämlich Satzanfang, -mitte und -ende — auftauchen.

- Zu einer letzten Gruppe gehören die negativen Adverbien wie *'barely', 'hardly', 'never', 'seldom'*.

In der Ausgangsversion von PENMAN wurden drei Typen von 'manner phrases' unterschieden: quality, means und comparison. Dabei deckt 'quality' die gängigen Adverbien ab (z.B. *'securely'*), 'means' Ergänzungen wie *'by hand'* und 'comparison' Beispiele wie *'like a fly'*. Es war nicht möglich, gleichzeitig ein Adverb und eine instrumentelle Ergänzung oder ein Adverb und eine 'means'–Ergänzung zu haben. Dazu wurde ebenfalls eine Grammatik–Erweiterung vorgenommen, so daß es nun möglich ist, zu generieren *'Tighten the spark plugs securely with a spark plug wrench.'* oder auch *'Tighten the spark plugs securely by hand to prevent crossthreading.'*

Secondary Predication

Die sog. 'secondary predication' ist im Englischen weit verbreitet. Mit ihr lassen sich z.B. zeitliche Zusammenhänge zwischen zwei Prozessen (Vorzeitigkeit, Gleichzeitigkeit) ausdrücken. In unseren Texten fanden sich — in Vor- und Durchführungsbedingungen — zahlreiche Beispiele insbesondere für die 'with'–Konstruktion (dt. Entsprechung: „bei" mit Partizipialkonstruktion). Die vorgenommenen Erweiterungen erlauben nun, Sätze wie die folgenden zu generieren:

- *'Check the car with the engine running.'*

- *'Check the engine with the car parked on level ground.'*

7.2.2 Die deutsche Grammatik

Bei der Entwicklung der deutschen Grammatik wurde die englische Grammatik als Vorlage herangezogen. Systeme, die übernommen werden konnten, blieben gemeinsam. Differenzen zwischen den beiden Sprachen wurden dann durch getrennte Systeme und durch die Einführung neuer Systeme behandelt.

Die englische Morphologie war mit einigen wenigen Prozeduren abgehandelt. Für das wesentlich formenreichere Deutsch wird die morphologische Komponente MORPHIX [Finkler, Neumann 1988] benutzt. Die morphologischen Prozesse sind daher nicht in die Grammatik integriert, sondern werden in einer Art Post–Prozeß auf den von der Grammatik produzierten funktionalen Strukturen und ihren Merkmalsmengen durchgeführt. Eine Alternative hierzu könnte sein, auch die morphologischen Prozesse in die systemische Grammatik zu integrieren; ein solcher Ansatz wird im KOMET–Projekt in Darmstadt untersucht [Teich 1992].

Die Unterschiede zwischen den relevanten Fragmenten der englischen und der deutschen Grammatik liegen hauptsächlich im syntagmatischen Bereich. Dazu gehören:

- flexiblere Stellung von Konstituenten,
- Unterschiede zwischen Haupt- und Nebensatz,
- erweiterte Möglichkeiten der Topikalisierung.

Unterschiede auf syntagmatischer Ebene konnten meist durch die Änderung von Gates bzw. der darin vorkommenden Realisierungsanweisungen (z.B. 'ORDER' u.a.) behandelt werden.[4]

Andere Unterschiede zwischen den Sprachen sind nicht rein syntagmatischer Natur, sondern beruhen darauf, daß eine Ausdrucksmöglichkeit in der einen Sprache vorliegt, die in der anderen nicht vorhanden ist. Ein Beispiel hierfür sind die unterschiedlichen möglichen Imperativ–Formen des Deutschen, die nicht in der gleichen Vielfalt im Englischen vorkommen. Beispiel:

- deutsch:
  - *'Zündkerzen fest anziehen'*
  - *'Ziehen Sie die Zündkerzen fest an'*
  - *'Ziehe die Zündkerzen fest an'*
- englisch: *'Tighten the spark plugs securely'*

Wo Sprachen solche zusätzlichen Wahlmöglichkeiten bieten, muß die Grammatik entsprechende Unterscheidungen unterstützen, d.h. es müssen zusätzliche systemische

[4] Als Gates werden in NIGEL solche (terminalen) Systeme bezeichnet, in denen keine weiteren Merkmale bestimmt werden, sondern nur Realisierungsanweisungen erfolgen.

Merkmale eingeführt werden und die Auswahl zwischen ihnen muß von Choosers und Inquiries geleistet werden können. Um die möglichen Unterscheidungen auch tatsächlich nutzen zu können, müssen die Wissensbasen der Umgebung des Systems (d.h. die Taxonomien, die textuelle Wissensbasis, usw.) die benötigen Informationen enthalten oder es müssen Defaultwerte gesetzt werden, die dann greifen, wenn die Informationen aus der Umgebung (noch) keine Entscheidung erlauben.

Beispiele für die Generierung von Präpositionen

Deutsch und Englisch unterscheiden sich erheblich bei der Entscheidung, welche Präposition für welchen Sachverhalt gewählt werden muß. In einer im Rahmen des TECHDOC–Projekts durchgeführten Magisterarbeit [Grote 1993] wurde daher ein Teilnetzwerk zur Behandlung von PPs für die deutsche Grammatik von TECHDOC entwickelt und implementiert.

Damit hat sich die Abdeckung der deutschen Grammatik für diesen Bereich erheblich erhöht (vgl. [Grote 1993]):

- Es existiert eine wesentlich detailliertere funktionale Differenzierung der unterschiedlichen *Minorprocess*–Typen im Systemnetzwerk (experientielle Region). Dieses erlaubt die Verbalisierung von Relationen zwischen dem Bezugswort und dem Sachverhalt des Satzes, die vorher nicht generierbar waren.

- Die Präpositionen können neben Nominalgruppe nun auch Präpositionalphrase und Adverb als Bezugswort nehmen. Dadurch sind Konstruktionen wie *von links* oder *bis unter die obere Pegelmarke* produzierbar.

- Die Modifikation von Präpositionen ist möglich: *fast bis unter die Pegelmarke.*

- Die Grammatik ermöglicht die Generierung von kontrahierten Präpositionen.

- Pronominalisierungen von Präpositionalphrasen werden generiert: *Ein erwärmter Motor und das Öl darin sind heiß.*

- Idiosynkrasien der Kasusrektion und Stellung sind handhabbar.

- Die Pronominalisierung der Präpositionalphrase als Ausdruck textueller Bedeutung kann generiert werden.

Im folgenden werden Beispiele für die damit generierbaren Strukturen gegeben (s.a. [Grote 1993]).

P (PP) und Artikelverschmelzung

```
(penman::example
  :logicalform
  (fi / fill
      :actor (penman::hearer / penman-kb::person)
      :actee (r / reservoir)
      :limit (d / direction
                :range (m / mark
                          :property-ascription (u / upper))
      :speechact penman::imperative))
```

Den Tank **bis zur** *oberen Pegelmarke auffüllen.*

Bis–Konstruktionen zur Angabe eines Endpunkts sind in Handbüchern häufig. Sie stehen in der Regel mit einer zweiten PP. Im englischen NIGEL war nur eine Nominalgruppe als Realisierung der *Minirange* möglich, eine SPL–Spezifikation wie die obige konnte nicht verarbeitet werden. In dem SPL–Ausdruck markiert das *keyword :limit* den Typ der Relation zwischen dem Objekt *mark* und dem Rest der Klause. Die eingebettete Relation *direction* repräsentiert das Komplement dieser Relation, und regiert selbst wiederum die *:range*. Der Kasus der Nominalgruppe *oberen Pegelmarke* wird von der Präposition *zu* regiert.

Da der nachfolgende Artikel unbetont ist, kann die Präposition mit dem Artikel verschmolzen werden.

Modifikation und Artikelverschmelzung

```
(penman::example
  :logicalform
  (rst / rst-volitional-cause
       :domain (dp / drop
               :actor (le / level)
               :limit (d / direction
                         :range (m / mark
                                   :size-property-ascription (l / lower-adj)
                                   :modified-by (c / close))
               :tense penman::present-perfect)
       :range (
           ... )))
```

Falls der Stand **fast bis zur** *unteren Pegelmarke abgesunken ist, ...* vs.
If the level has dropped **close to the** *lower mark, ...*

Dieses Beispiel zeigt wiederum die Verschmelzung von Artikel und Präposition, ein Sprachspezifikum des Deutschen (vgl. Englisch *to the*). Zum anderen illustriert es die Modifikation einer PP: Das Keyword *:modified-by* markiert die funktionale Rolle

des assoziierten Konzepts *close*, welches in der deutschen Ausgabe als *fast* lexikalisiert wird.

Pronominaladverb

```
(penman::example
  :logicalform
  (pa / property-ascription
      :domain ((en / engine
                   :property-ascription (wu / warmed-up))
               (ol / oil
                   :spatial-locating en))
       :range (ho / hot)))
```

Ein erwärmter Motor und das Öl **darin** *sind heiß.* vs.
A warmed-up engine and the oil **in it** *are hot.*

Im Deutschen steht das Pronominaladverb **darin** als Mittel zur Realisierung der textuellen Metafunktion zur Verfügung. Aus dem Beispiel wird ersichtlich, wie aufgrund der vorherigen Erwähnung des Konzepts *engine* die topologische Relation *:spatial-locating*, welche die Beziehung zwischen dem Öl und dem Motor spezifiziert, durch eine Proform realisiert wird. Proformen dieser Art existieren im Englischen nicht. Die PP tritt in Attribut–Funktion, als *Qualifier*, auf.

Die topologische Relation wird aufgrund der dimensionalen Eigenschaften der Instanz des Konzepts *engine*, wie sie in der Objekthierarchie der Wissensbasen modelliert sind, durch *in* realisiert.

Topologische Relation

```
(penman::example
  :logicalform
  (rf / refill
      :actor (penman::hearer / penman::person)
      :actee (en / engine)
                :instrumental (ol / oil
                                  :property-ascription (re / recommended))
      :limit (d / direction
                :range (m / mark
                :property-ascription (up / upper)
                            :spatial-locating (d / dipstick)))
      :speechact penman::imperative))
```

Den Motor mit dem empfohlenen Öl bis zur oberen Pegelmarke **am** *Tauchmeßstab auffüllen.* vs.
Refill the engine with the recommended oil to the upper mark **on** *the dipstick.*

An diesem Beispiel zeigt sich deutlich die unterschiedliche semantische Distribution der topologischen Präpositionen entsprechend der Konzeptualisierung der Dimension des *dipstick*. Der *dipstick* wird in beiden Sprachen als [vertical-surface] konzeptualisiert, die topologische Relation, durch das Keyword *:spatial-locating* repräsentiert, wird aber im Deutschen durch die Präposition *am*, im Englischen durch *on* verbalisiert.

Lokational/Direktional

```
(penman::example
  :logicalform
  (c / check
     :actor (penman::hearer / penman-kb:person)
     :actee (cl / coolant-level
               :spatial-locating (r / reserve-tank))
     :speechact penman::IMPERATIVE)
```

Den Kühlmittelstand **in dem** *Reservetank überprüfen.* vs.
Check the coolant level **in the** *reserve tank.*

In ist eine der Präpositionen, die bei direktionalen vs. lokationalen Lesarten unterschiedliche Kasus regieren. Die räumliche Relation in obiger Inhaltsspezifikation ist statischer Art, die Präposition allein reicht zur Markierung dieser Bedeutung nicht aus, erst durch die Wahl des Dativ wird der Typ des *Minorprocess* realisiert.

```
(penman::example
  :logicalform
  (f / fill
     :actor (penman::hearer / penman-kb:person)
     :actee (cl / coolant)
     :spatial-locating (rt / reserve-tank)
     :speechact penman::IMPERATIVE)
```

Kühlmittel **in den** *Reservetank füllen.* vs.
Fill the coolant **into** *the reserve tank.*

Der Akkusativ der Nominalgruppe *den Reservetank* markiert die Zielgerichtetheit der toplogischen Relation. Was im Deutschen durch Kasusopposition geleistet wird, wird im Englischen durch eine alternative lexikalische Einheit (*into*) realisiert.

7.2.3 Französische Grammatik

Vor der Implementierung des Grammatik–Fragments für die Handbuchtexte war eine kontrastive Analyse und ein Vergleich der französischen Texte mit den englischen und deutschen notwendig. Trotz vieler Gemeinsamkeiten zeigten sich dabei auch einige wichtige Unterschiede (vgl. [Stede, Weis 1993] in [TECHDOC 1993]).

Manche Informationen, die im Deutschen oder Englischen als Adjektive oder in Form von Partizipialkonstruktionen ausgedrückt werden, tauchen in der französischen Version in untergeordneten Nebensätzen oder Gerundiv–Strukturen auf (z.B. *'warmed-up engine'* vs. *'un moteur ayant tourné un moment'*).

In Anweisungstexten sind Infinitiv–Konstruktionen vorherrschend; dennoch wird gelegentlich auch grammatische Zeit relevant. Hier unterscheiden sich die Sprachen wiederum, z.B. französisch: *'Si le niveau s'approche de ...'* im Unterschied zum englischen *'If the level has dropped to ...'*.

Ähnlich gelagerte Unterschiede gibt es im Hinblick auf Aspekt. So gibt es im Französischen keine direkten Entsprechungen zu den englischen continuous–Formen; allerdings ist eine Form *'être en train de faire qc.'* möglich.

Auch im Hinblick auf die Granularität von Lexikalisierungsentscheidungen gab es Unterschiede zwischen den Texten. In den französischen Texten wurden im Kontrast zu den eher allgemeinen Verben im Englischen eher spezifische Verben verwendet; Beispiel: englisch *'to remove'*, französisch *'extraire'* oder *'retirer'*.

Handbuchtexte sind auf Knappheit ausgelegt, aber selbst in diesem Genre sind im Französischen Ellipsen seltener als in den beiden anderen Sprachen. Zusätzlich wird mit dem Pronomen *'ceci'* auf vorangehend Geäußertes zurückverwiesen. Beispiele:

'...remplir le moteur à l'aide d'une huile preconisée, ceci jusqu'au repère supérieur de la jauge.'

'...serrer de 1/2 tour à l'aide d'une clé a bougie, ceci afin de compresser la rondelle.'

Die wichtigsten Grammatik–Erweiterungen

Die englische Grammatik wurde für das Französische als Ausgangspunkt genommen. Für das für Handbuchtexte relevante Fragment waren die folgenden Änderungen notwendig:[5]

- Stellung der Adjektive: Adjektive werden im Französischen fast immer dem Nomen nachgestellt. Diese Position wird daher als Default verwendet. Es gibt allerdings auch Adjektive, die immer vorangestellt werden; diese sind in einer Ausnahmenliste enthalten (neben allen Zahl–Adjektiven z.B. *'grand'*, *'petit'*, *'bon'*, *'mauvais'*, ...). Adjektive, die in Prä- bzw. Postpositionen unterschiedliche Bedeutung haben, werden durch getrennte Lexikon–Einträge behandelt.

[5] Alle in dieser Arbeit abgedruckten Generierungsergebnisse – wie z.B. in Abb. 7.4 – wurden ohne Nachbearbeitung – auch mit ihren Schwächen und Fehlern – übernommen. Sie dokumentieren den Leistungsstand des TECHDOC–Systems im Dezember 1993.

Changing of oil

1. Warm up the engine.
2. Remove the engine oil filler cap and the drain bolt, and drain the oil.

   CAUTION: The warmed-up engine and the oil in the engine are hot. Be careful, in order not to burn you.
3. Reinstall the drain bolt with a new washer, and refill the engine to the upper mark on the dipstick with the recommended oil.

   ENGINE OIL CHANGE CAPACITY:
4. Start the engine, and make sure that oil is not leaking from the drain bolt.

Le remplacement de l'huile

1. Laisser chauffer le moteur.
2. Retirer le bouchon de remplissage d'huile et le boulon de vidange, puis vidanger l'huile.

   L'ATTENTION: Le moteur chaud et l'huile dans le moteur etre chaud. Etre prudent, pour ne pas bruler vous.
3. Remettre en place le boulon de vidange avec une rondelle neuve, puis remplir le moteur jusqu'a le repere superieur sur la jauge avec l'huile preconise.

   LA CONTENANCE EN HUILE MOTEUR:
4. Demarrer le moteur, puis veiller que l'huile ne NIL pas fuir de le boulon de vidange.

Abbildung 7.4: Ein generierter Beispieltext in Englisch und Französisch

- direktes Objekt: Das direkte Objekt steht im Französischen hinter dem Verb–Komplex, solange es nicht pronominalisiert wird. Ein Pronomen als direktes Objekt allerdings verlagert seine Position vor das konjugierte Verb bzw. zwischen die finite und infinite Verbform. Im PENMAN–System erfolgt die Festlegung der Konstituenten–Ordnung schon **vor** der Lexikalisierung (und auch Prozessen wie Pronominalisierung); daher muß die 'Umstellung' des Pronomens als Korrektur der zuerst erreichten Linearisierung erfolgen.

- Imperativ: Im Französischen gibt es analog zum Deutschen zwei Möglichkeiten, Imperative zu bilden: Einerseits in infiniter Verbform, die unpersönlich ist (z.B. *'Laissez chauffer le moteur'*), oder in finiter Form, wobei dann zwischen einer eher persönlichen und einer eher formalen Form (z.B. *'Assied toi'* im Unterschied zu *'Asseyez vous'*) unterschieden werden kann. In Anweisungstexten spielt die zweite Möglichkeit nur eine geringe Rolle, unpersönlich gehaltene Handlungsaufforderungen überwiegen.

- Determinatoren: Innerhalb eines Satzes sind — im Unterschied z.B. zu sehr knapp gehaltenen deutschen Sätzen — Determinatoren obligatorisch. Eine Ausnahme hiervon sind Überschriften, hier kann der Artikel weggelassen sein (z.B. *'Remplacement de l'huile'*).

- Negation: Kann im Englischen und Deutschen die Verneinung mit einem Partikel ausgedrückt werden (*'not'* bzw. *'nicht'*), so benötigt das Französische in der Regel die beiden Partikel *'ne'* und *'pas'*. Diese können nicht als lexikalische Einheit gesehen werden, da ihre Stellung im Satz variieren kann. Wird ein finites Verb negiert, so umschließen *'ne'* und *'pas'* das Verb; ansonsten stehen sie zusammen vor der infiniten Form. Um diese Stellungsunterschiede zu behandeln, werden zwei getrennte systemische Funktionen eingeführt und die ORDER–Angaben in den Realisierungsweisungen entsprechend angepaßt. Ähnlich wäre dann für weitere Verneinungsmöglichkeiten vorzugehen: *'ne ... point'*, *'ne ... plus'*, *'ne ... guère'*, *'ne ... jamais'*, ...

- Nachbearbeitung: In der englischen Grammatik–Version wird die Umwandlung des unbestimmten Artikels 'a' vor vokalisch anlautenden Substantiven bzw. Adjektiven in 'an' auf dem zur Ausgabe anstehenden String durchgeführt. In analoger Weise werden in der französischen Version die folgenden Phänomene behandelt:

  - Euphonie–Regeln
    Elision: - la alimentation => l'alimentation
    Hiathus: - sa amie => son amie
  - Verschmelzungsregeln:
    la couleur de le moteur => la couleur du moteur

Die Entwicklung des Fragments der französischen Grammatik hatte vor allem explorativen Charakter. Für tatsächliche Anwendungen müßte hier eine gründliche Erarbeitung mit entsprechendem Personalaufwand erfolgen.

7.2.4 Anforderungen des Textmaterials

Arbeitet man mit Textmaterial, wie es in der Praxis verwendet wird, so stößt man auf zusätzliche Probleme und Fragestellungen, die bei Arbeiten an solchem Textmaterial nicht entstehen, das aus rein linguistischer Sicht zusammengestellt wurde. Für

die Akzeptanz eines sprachverarbeitenden Systems ist es allerdings unabdingbar, sich auch mit den Anforderungen alltäglicher Sprachverwendung auseinanderzusetzen.

Einige Beispiele für solche Probleme:

- Bei der tatsächlichen Sprachverwendung werden nicht nur komplette Sätze und — wie in Überschriften — Nominalgruppen als Einheiten verwendet, sondern auch andere, im grammatischen Sinne unvollständige Strukturen wie das *„bei kaltem Kühlmittel bis zur Markierung im Behältereinfüllstutzen“* in der Aufzählung verschiedener Alternativen (vergl. 9.1).

- Die im Lexikon des PENMAN–Systems möglichen Einträge sind für Anwendungen im Deutschen in vielen Fällen noch zu unstrukturiert. Zwar kann im Englischen ein String *'rear compartment'* als Lexikoneintrag problemlos verarbeitet werden. Im Deutschen müßte man für das entsprechende *„hinteres Abteil“* entweder ein Vollformenlexikon verwenden oder — was vorzuziehen ist — es muß möglich sein, die Nominalgruppe als Struktur im Lexikon darzustellen und dann morphologischen Prozessen zugänglich zu machen.

- Auch für Funktionsverbgefüge wie *„sich Zugang verschaffen zu“* muß im Lexikon eine syntaktische Struktur und nicht nur ein String abgelegt werden können. Die Grammatik muß solche Strukturen dann integrieren und weiterverarbeiten können.

Für die Bezeichnung technischer Objekte werden im Deutschen wie im Englischen komplexe Nominalkomposita verwendet (z.B. Englisch: *engine oil level warning lamp*, Deutsch: *Motorölstand–Warnleuchte*). Ist das bezeichnete Objekt hinreichend genau modelliert, d.h. sind Angaben zu seinem Material, seiner Funktion, seiner Form repräsentiert, so müssen solche Komposita nicht im Lexikon stehen, sondern nur die Wörter für die jeweiligen Komponenten. Die Komposita können dann aus den Komponenten synthetisiert werden. Damit ist nicht nur morphologische Synthese gemeint, sondern die Möglichkeit, Benennungen dieser Art systematisch gewinnen zu können.

In [Schmitt 1992] werden hierfür folgende Regeln vorgeschlagen:

1. Bei der Benennung technischer Teile sind Form und/oder Funktion anzusprechen. Haben Teile die gleiche Form, aber unterschiedliche Funktionen, so ist die Nennung der Funktion vorrangig, entsprechendes gilt für die gleiche Funktion bei unterschiedlicher Form.

2. Gleiche Teile sollten auch gleich benannt werden.

Er gibt Beispiele hierfür wie das folgende: engl. *sealing ring*
Funktion: Abdichten; Form: Ring;
Mögliche dt. Bezeichnung: *Dichtring*

7.3 Behandlung textueller Phänomene

In der bisherigen Form ist PENMAN ein reiner Satzgenerator. Bei komplexen Sätzen treten aber bereits textuelle Phänomene wie Pronominalisierung auf, die vom System in Ansätzen geleistet werden. Im folgenden wird anhand der Beispieltexte aufgezeigt, welche textuellen Phänomene für das Genre „Handbuch“ relevant sind und aufgrund welcher Wissensquellen sie behandelt werden können.

Für die systemische Grammatik ist es charakteristisch, daß an den Entscheidungspunkten Fragen an die Umgebung gestellt werden. In der Ausprägung als Satzgenerator betreffen diese Fragen in erster Linie die Wissensquellen „Gebietsmodell“ und sog. ‘upper model’, d.h. sie beziehen sich auf die Klassifikation von Konzepten, Relationen usw. in bezug auf dieses Modell. Zur Behandlung textueller Phänomene ist die Nutzung einer zusätzlichen Wissensquelle erforderlich. Die folgende Diskussion soll auch erarbeiten, welche Information in dieser textbezogenen Wissensbasis zur Verfügung gestellt werden muß.

Es gibt verschiedene Antworten auf die Frage, an welcher Stelle im Generierungsprozeß die hier diskutierten Entscheidungen angesiedelt werden sollen. So gibt es mindestens zwei Stellen, an denen beispielsweise Entscheidungen über Artikel eingebaut werden können.

- Die eine Möglichkeit ist die, bereits im SPL–Ausdruck (vgl. 2.3.3) den Artikel festzulegen. Dafür spricht, daß die notwendige Information zu diesem Zeitpunkt bereits vorhanden ist.

- Eine Alternative ist, diese Entscheidung erst beim Durchlaufen der Grammatik zu fällen. Da einige der Kriterien, wie sich im folgenden zeigen wird, aber auch genrespezifische Informationen berücksichtigen, müssen diese dann in die grammatischen Entscheidungen direkt einfließen können.

Auf die Frage der möglichst sinnvollen Plazierung der Entscheidungsstellen wird daher im folgenden ebenfalls Wert gelegt.

7.3.1 Artikel in Nominalphrasen

Bei Durchsicht der Wartungsanleitungen ergibt sich das folgende Modell für die Bestimmung des Artikel:

Ein Abschnitt hat ein Thema, zum Beispiel das Durchführen einer der Grundoperationen (Prüfen, Ergänzen, Austauschen) bezogen auf ein Objekt oder eine Substanz. Diese Aktion ist an einem physischen Ort durchzuführen, der damit einen räumlichen Kontext konstituiert. Die für die Aktion zentralen Aggregate und ihre Teile sind damit als präsent zu betrachten. Im Text kann daher mit definiten Artikeln auf diese Objekte verwiesen werden. Zu den als physisch vorhanden betrachteten Objekten gehören hierbei auch “abstrakte Objekte“, wie Flüssigkeitsstände oder Ergebnisse von Meßvorgängen.

Für die Wissensrepräsentation lassen sich daraus die folgenden Forderungen ableiten: Es muß möglich sein, für abstrakte wie konkrete Objekte zu entscheiden, ob sie durch das globale Thema eines Abschnitts in der geschilderten Weise eingeführt sind und damit exophorische Referenz möglich ist.

Die geschilderte Möglichkeit zur Referenz kann bereits durch die Überschrift aktiviert sein. Beispiel: Nach der Überschrift „Checking coolant" kann definit referenziert werden auf *„the coolant level"* und *„the reserve tank"*. Die Referenz mit *„the engine"* ist nicht begründbar aus einer Teil-von-Beziehung in bezug auf das Hauptaggregat (also: Kühlmitteltank), sondern ergibt sich durch das implizite Globalthema, daß ein Automobil gewartet werden soll. Der Motor als eines der zentralen Teile des Automobils ist damit ebenfalls definit referenzierbar.

Es muß also zusätzlich modelliert werden, welche „globalen" Aggregate typischerweise für jede Art von Kfz–Wartung sofort zur Referenz verfügbar sind.

Zusätzlich zu den allgemein verfügbaren Objekten können dann durch Aktionen bzw. ihre Teilschritte weitere Objekte referenzierbar werden. Kontrollieren des Kühlmittelstands kreiert beispielsweise einen lokalen Kontext, in dem die Marken auf der Ableseskala (diese sind auch im begleitenden Bild hervorgehoben) zur Verfügung stehen, weshalb ebenfalls definit verwiesen werden kann (*„below the Max mark"*, *„unterhalb der 'Max'–Marke"*, *„bis zur 'Max'–Linie"* usw.)

7.3.2 Referenz durch Abstraktion

Bestimmte Nominalphrasen sind erst dann zur Referenz nutzbar, wenn das zur Diskussion stehende Objekt tatsächlich textuell schon explizit erwähnt worden ist. Ein Beispiel hierfür: Mit *„the level"* bzw. *„der Stand"* kann auf den Kühlmittelstand verwiesen werden im zweiten Satz des Textes „Kühlmittel" (vgl. Abb. 5.10), da *der Kühlmittelstand* im ersten bereits eingeführt ist. Diese Referenz durch Abstraktion sollte sinnvollerweise in der Grammatik bei der Lexikalisierung behandelt werden. Hierfür ist die derzeitige Auswahl von lexikalischen Termen sinngemäß wie folgt zu erweitern:

Es ist zu prüfen, ob das zur Debatte stehende konkrete oder abstrakte Objekt im Text bereits erwähnt worden ist. Weiter ist zu prüfen, ob andere Formen, wie Pronomina u.ä. Vorrang haben. Wenn dies nicht der Fall ist (beispielsweise weil Pronomina Ambiguitäten einführen würden), kann ein abstrakterer Terminus herangezogen werden (also *„Stand"* statt *„Kühlmittelstand"*).

7.3.3 Substanzen

Nominalphrasen, die auf unspezifische Mengen von Substanzen verweisen, haben typischerweise keinen Artikel. Wir finden also: *„eine Lösung von Antifrost und Wasser"* oder *„Kühlmittel in den Kühler nachfüllen"*. Allerdings können verallgemeinerte Quantoren Verwendung finden, wie *„kein Kühlmittel"* oder *„enough coolant"*. „Objektiviert" und damit definit referenzierbar werden Substanzen erst in bestimmten

Konstellationen. Eine davon kann sein, daß sie sich in einem Behälter befinden. Auf eine solche Substanz kann dann definit verwiesen werden. Im Beispieltext: *„Den Kühlerdeckel nicht bei heißem Motor entfernen; das Kühlmittel steht unter Druck und könnte Sie ernsthaft verbrennen.“* Hier ist ein „Objektcharakter“ gegeben: Das Kühlmittel *im Kühlsystem* und zusätzlich die Situation *„bei heißem Motor“*.

7.3.4 Sprachspezifische Variationen

Ein großer Teil der bisher skizzierten Heuristiken trifft sowohl für Deutsch als auch für Englisch zu. Es gibt allerdings auch Unterschiede. So heißt es in der deutschen Textversion: *„Kühlmittel beschädigt den Lack“* und in der englischen: *„Radiator coolant will damage paint“*. Im Deutschen ist hier die definite Referenz zwingend, möglich wäre lediglich ein Übergang zu *„Kühlmittel beschädigt Lacke“* oder besser gar *„ . . . Lackfarben“*. *„Radiator coolant will damage the paint“* scheint aber für das Englische auch akzeptabel.

Ein anderer Unterschied ist bei der Verwendung von *„any“*, die im Deutschen zwar nachempfunden werden kann, aber dann etwas ungewöhnlich klingt, also: *‘Quickly rinse any spilled coolant from painted surfaces’, „Jegliches verschüttete Kühlmittel schnell von lackierten Flächen abwischen.“* oder *‘After any remaining pressure has been relieved . . . ’, „Nachdem jeglicher verbliebene Druck abgelassen ist . . . “*

7.3.5 Ellipsen als textuelles Phänomen

Der Beispieltext ‘Checking coolant’ (vgl. Abb. 5.10) enthält verschiedene Stellen, an denen insbesondere in koordinierten Strukturen Ellipsen auftreten. Dies führt dazu, daß der Text insgesamt knapper wird, was einer generellen Vorgabe für Instruktionstexte entgegenkommt.

Elisionen erfolgen in den deutschen und englischen Textversionen oft an den gleichen Stellen, wenngleich sie sich in Einzelheiten dann unterscheiden.

Was sind häufig vorkommende Formen von Elision?

- Auslassen lediglich des Subjekts:

  „Falls der Stand unterhalb der ‘Max’–Marke liegt, aber noch sichtbar ist.“

- Auslassen von Subjekt und dem finiten Element der Verbalgruppe:

  „If the level is below the Max mark, but still visible“ (lang: *„ . . . but the level is still visible“*).
  „Falls . . . , sollte das Kühlsystem auf Leckstellen überprüft und erforderlichenfalls repariert werden“ (lang: *„ . . . und das Kühlsystem sollte erforderlichenfalls repariert werden“*).

Daß ein Potential für Elision vorhanden ist, läßt sich zwar bereits auf der SPL–Ebene feststellen, ob und in welcher Form sie allerdings tatsächlich realisiert wird,

interagiert auch mit anderen Entscheidungen, so zum Beispiel mit der, ob Pronomina verwendet werden.

Beispiel: Wird für den auf *„the cooling system should be checked for leaks“* folgenden Halbsatz *„... and repaired, if necessary“* ein Subjekt realisiert, zum Beispiel als Pronomen *„it“*, dann ist die Elision des *„should“* nicht mehr möglich. Daher spricht einiges dafür, Ellipsenbildung in die Grammatik zu verlagern und erst zu einem relativ späten Zeitpunkt vorzunehmen, wenn die bis dahin realisierte syntaktische Struktur vorliegt. Dies auch deshalb, weil es andere nicht–lokale Phänomene gibt, die mit der Lexikalisierung interagieren. Ein Beispiel hierfür ist: *„Falls der Stand unterhalb der 'Max'–Marke liegt, aber noch sichtbar ist ...“* Wird hier statt *„liegt“* im ersten Halbsatz ebenfalls *„ist“* lexikalisiert, so kann dieses nachträglich elidiert werden: *„Falls der Stand unterhalb der 'Max'–Marke, aber noch sichtbar ist ...“*

Zur kontrastiven Analyse

Die Unterschiede bei der linguistischen Realisierung korrespondierender Inhalte zwischen den deutschen und englischen Versionen in unserem Korpus können in lexikalische und syntaktische klassifiziert werden. Zur ersten Kategorie gehören zunächst einfache Fälle, in denen eine 1:1–Abbildung zwischen sprachspezifischen lexikalischen Einheiten (die möglicherweise phrasaler Art sind) möglich ist.

Im Deutschen sind zum Beispiel Nominalkomposita bevorzugt, denen im Englischen oft dann eine Nominalgruppe entspricht (z.B. *„Kühlervorderseite“ = 'front of the radiator'*). Es gibt aber auch Fälle von inhaltlicher Entsprechung und differierender linguistischer Struktur, bei der dann die englische Form knapper ist (Beispiel: *„Schalter mit 3 Stellungen“ = 'three position switch'*).

Dann gibt es den Fall idiomatischer Ausdrücke, die zwar dieselbe syntaktische Funktion haben, aber leicht unterschiedliche Bedeutungen tragen; so entspricht etwa *„für längere Zeit“* der englischen Phrase *'for more than a short time'*. Versucht man diese wörtlich in die jeweils andere Sprache zu übersetzen, so klingen beide Varianten inakzeptabel.

Ein anderer Typ lexikalischer Unterschiede ergibt sich entlang der Dimension Spezifizität. Oft steht einem relativ allgemeinen englischen Verb eine ganze Reihe spezifischer deutscher Entsprechungen gegenüber, so etwa dem *'remove'* die spezifischeren Ausprägungen *„herausnehmen“*, *„herausziehen“*, *„abnehmen“*, *„entfernen“*. Die spezifischeren deutschen Verben tragen oft zusätzliche Bedeutung, so etwa über die räumliche Charakteristik der Aktion oder die Art der beteiligten Objekte.

Ein weiteres Phänomen ist Inkorporation. Verben können z.B. Bedeutung inkorporiert haben, die sonst durch ein Adverb ausgedrückt werden muß. Im Englischen ist dies etwa durch die Vorsilbe *'re'* möglich: *'reinstall the dipstick'*; im Deutschen muß hier das Adverb *„wieder“* (*„den Tauchmeßstab wieder einführen“*) zusätzlich eingefügt werden, um den restitutiven Charakter der Aktion auszudrücken (vgl. 7.5.1).

Eigenschaften von Objekten können sowohl durch Adjektive wie durch defi-

nierende Relativsätze ausgedrückt werden; das Englische hat hier allerdings Einschränkungen, wenn das Adjektiv zusätzlich intensiviert ist. Hier gibt es für *„eine zu lose Zündkerze“* nicht die wörtliche Entsprechung *‘ * a too loose spark plug’*, sondern es muß auf einen Relativsatz übergegangen werden: *‘a plug that is too loose’*.

Ähnlich liegt der Fall bei Partizipien. Hier kann im Englischen lediglich das reine Partizip vor das Nomen plaziert werden; komplexe Partizip–Konstruktionen wie *„der auf ebenem Boden abgestellte Wagen “* sind im Englischen nicht möglich. Hier muß dann ein Relativsatz verwendet werden.

Ein SPL–Ausdruck für mehrere Sprachen Die meisten der soeben diskutierten Unterschiede zwischen Englisch und Deutsch machen es noch nicht nötig, unterschiedliche Eingabespezifikationen für die Grammatiken anzusetzen. Die SPL–Terme (vgl. 2.3.3) sind ja auch entworfen als Möglichkeit, die ‘message’ zu kennzeichnen, die übermittelt werden soll, und schreiben die syntaktische Realisierung nicht bereits vollständig fest.

Erst beim Traversieren der Grammatik — im systemischen Sinne gesehen als Potential zum Ausdrücken von Bedeutung — wird diese Botschaft dann in die entsprechende Form für die jeweilige Sprache umgewandelt.

Einige der Elemente in SPL–Termen, insbesondere die Schlüsselwörter, die als Abkürzungen oder sogenannte *„SPL–Makros“* für komplexe Merkmalskombinationen stehen, könnten fehlgedeutet werden als lediglich syntaktische Vorgaben (z.B. das Schlüsselwort ‘:TENSE’ und seine verschiedenen Wertbelegungen). Der Realisierungsprozeß beim Durchlaufen der Grammatik ist aber wesentlich komplexer, da die grammatischen Systeme, die durch die SPL–Makros gegebenen Merkmale und die an die Wissensbasis gestellten Inquiries miteinander interagieren. Das SPL–Makro ‘:TENSE’ beispielsweise expandiert in eine Konfiguration aus Ereigniszeit, Sprechzeit und Referenzzeit und deren relativen Positionen. Es hat zweierlei Funktionen: Zum einen drückt es aus, daß temporale Beziehungen ausgedrückt werden sollen, und es gibt an, von welcher Art diese sind.

Ein Beispiel: Wenn für eine Eigenschaftszuweisung kein ‘:TENSE’ spezifiziert ist, so wird dadurch implizit der zeitliche Aspekt in der ‘message’ als nicht relevant gekennzeichnet. Dies führt dann beim Grammatik–Durchlauf dazu, daß die Eigenschaft als Adjektiv ausgedrückt wird. Ist in der Struktur für die Eigenschaftszuweisung zusätzlich das ‘:TENSE’–Keyword angegeben, dann wird eine Ausdrucksmöglichkeit gewählt, die es erlaubt, die zeitlichen Beziehungen linguistisch zu realisieren. Im Englischen führt dies zur Wahl eines Relativsatzes; im Deutschen gibt es dazu noch die zusätzliche Option einer pränominalen Partizip–Konstruktion.

Die Notwendigkeit für sprachspezifische SPL–Ausdrücke In den untersuchten Texten fanden sich auch Beispiele für Unterschiede in den einzelsprachlichen Ausprägungen, die nicht mehr durch die Grammatik allein aus identischen SPL–Termen realisiert werden können. Als Beispiel das folgende Paar:

'They will adversely affect engine performance and durability.'.
„Sie beeinträchtigen Motorleistung und Haltbarkeit."

Hier liegt der Unterschied zwischen den einzelsprachlichen Versionen darin, daß die englische Verbalisierung kompositional ist. Das allgemeine Verb *'affect'* wird kombiniert mit dem Polaritätswert (*'adversely'*). Im Unterschied dazu ist dieselbe Bedeutung im Deutschen in ein Lexem inkorporiert (*„beeinträchtigen"*). Eine kompositionale Realisierung ist im Deutschen ebenfalls möglich, z.B. als *„negativ beeinflussen"*. Spezifischere Lexeme scheinen aber Vorrang zu haben.

Der allgemeine Aspekt ist hier, daß nicht einzelnen Konzepten, sondern Konstellationen aus Konzepten und Rollen Lexikalisierungen zugeordnet werden, und daß diese Zuordnungen nicht rein kompositional sein müssen, sondern die Füller von Rollen inkorporiert haben können.

Ein anderer Fall wird durch folgendes Beispiel illustriert:

'Reinstall the cap and twist clockwise until it stops.'
„Den Deckel wieder anbringen und im Uhrzeigersinn bis zum Anschlag drehen."

Auch hier fassen wir die beiden korrespondierenden Sätze als einander inhaltlich entsprechend auf. Die zugrundeliegende semantische Struktur drückt die Beziehung zwischen einer Aktion aus und dem Endzustand, der erreicht werden muß. Englisch scheint hierfür eine eher prozeßorientierte Verbalisierung zu bevorzugen (*'until it stops'*); im Deutschen ist eine ergebnisorientierte Version häufiger, bei welcher der Endzustand der Drehbewegung mit einer Präpositionalphrase (hier: *„bis zum Anschlag"*) beschrieben wird.

7.4 Fragen der Lexikalisierung

7.4.1 Lexikalisierung in PENMAN

In PENMAN erfolgt die Lexikalisierung einer Konstituente erst am Ende des entsprechenden Grammatikdurchlaufs. Beim Durchlaufen der Grammatik werden (durch die Realisierungsanweisungen CLASSIFY und OUTCLASSIFY) Festlegungen dahingehend getroffen, welche Merkmale für eine Lexikalisierung zu gelten haben bzw. nicht vorhanden sein dürfen. Mit dieser Menge von erforderlichen bzw. nicht erlaubten Merkmalen können dann aus dem Lexikon mögliche Lexikalisierungen bestimmt werden (im Idealfall sind diese eindeutig; angewendet wird diese Art der Lexikalisierung in erster Linie für Funktionswörter, aber auch für Modalverben).

Eine andere Lexikalisierungsmöglichkeit ergibt sich dadurch, daß mit Konzepten der Wissensbasis mögliche Wörter assoziiert sind. Die Lexikoneinträge enthalten die für ein Wort gültigen grammatischen oder lexikogrammatischen Merkmale. Die beim Grammatikdurchlauf angesammelte Menge von erforderlichen bzw. verbotenen Merkmalen wird nicht zur direkten Bestimmung des Worts, sondern nur zur Entscheidung

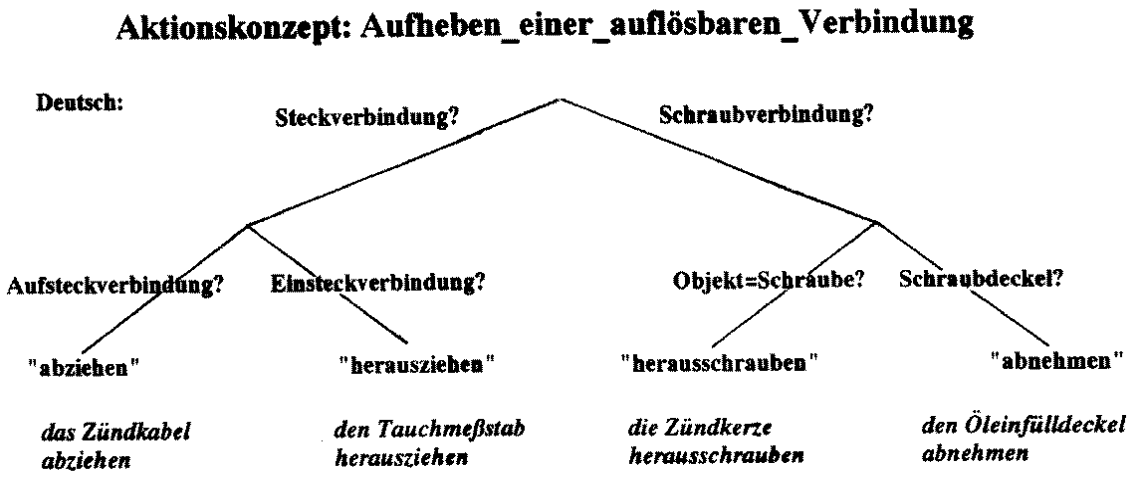

Abbildung 7.5: Lexikalisierungen eines abstrakten Aktionskonzepts für Deutsch ...

über seine Zulässigkeit verwendet, d.h. sind erforderliche Merkmale nicht vorhanden oder verbotene Merkmale vorhanden, so wird das Wort nicht akzeptiert und es wird rekursiv bei abstrakteren Konzepten nach einer möglichen Lexikalisierung gesucht. Mit einem Konzept muß nicht notwendigerweise nur ein Wort assoziiert sein. Zwar sind typischerweise Prozesse mit Verben, Objekte mit Nomen, Eigenschaften mit Adjektiven assoziiert. Mit Prozesskonzepten können aber auch Nomen im Sinne einer Nominalisierung assoziiert sein. Diese treten dann typischerweise auf, wenn ein Prozesskonzept in der systemischen Funktion THING realisiert wird, d.h. zu einem Nomen in einer Nominalgruppe führen soll.

7.4.2 Abstrakte Aktionskonzepte

Die bearbeiteten Texte beinhalten vor allem Handlungsaufforderungen und Aussagen. Um eine sprachunabhängige Repräsentation zu erreichen, müssen sowohl für Handlungsaufforderungen wie für allgemeine Aussagen geeignete abstrakte Darstellungen gefunden werden, die dann als Basis für sprachspezifische Lexikalisierungen dienen können.

Es ist nicht direkt offensichtlich, bezüglich welcher Aspekte die Abstraktionen vorgenommen werden sollen. Bei den in Handlungsaufforderungen angesprochenen Aktionen bietet sich an, die Vor- und Nachbedingungen heranzuziehen und Aktionen, die hier Gemeinsamkeiten haben, in abstrakte Klassen zusammenzufassen.

Verb–Bedeutungen sind aber oft nicht eindimensional, sondern nehmen Bezug auf unterschiedliche Aspekte. Dem kann Rechnung getragen werden durch Mehrfachvererbung.

Ein Beispiel: *'remove the engine oil filler cap and drain bolt'* bzw. *„Motoröl–Einfülldeckel und Ablaßschraube entfernen“*. Hier liegt zum einen im Hinblick auf den

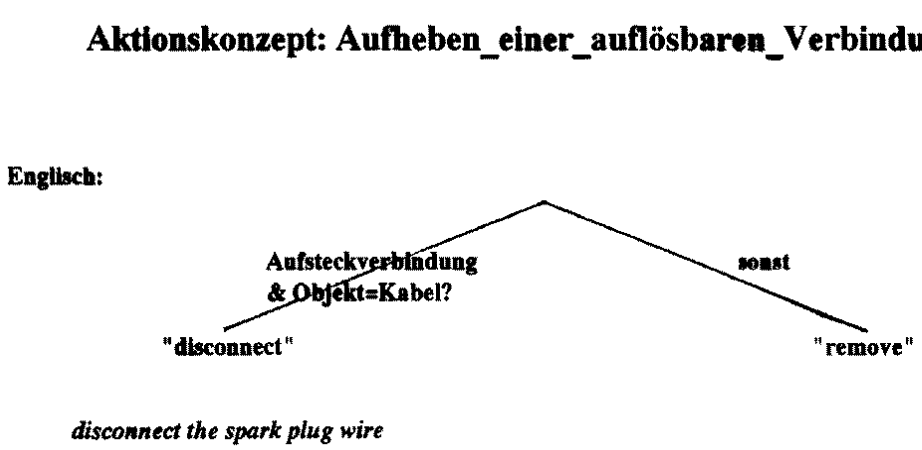

Abbildung 7.6: ... und für Englisch

Aspekt des Verbindungszustands eine Aktion vom Typ 'Aufheben einer auflösbaren Verbindung' vor; in der Verb–Bedeutung ist aber auch noch der Aspekt einer Transportaktion enthalten: Ein Teil wird vom aktuellen Ort an einen bestimmten oder unbestimmten anderen Ort bewegt.

Für einige Aktivitäten liegt Dualität vor in dem Sinne, daß einem abstrakten Aktionskonzept auch ein „gegenläufiges" zugeordnet werden kann, das die Effekte des ersten wieder aufzuheben in der Lage ist.

Ein Beispiel hierfür: *' ... and drain the oil'* bzw. *„ ... und Öl ablassen"*. Abstrahiert liegt hier der Aktionstyp 'Substanz aus Behälter entfernen' vor, dem der komplementäre Aktionstyp 'Substanz in Behälter einbringen' zugeordnet werden kann (im Text: *' ... refill the engine with the recommended oil ... '* bzw. *' ... den Motor mit dem empfohlenen Öl auffüllen ... "*).

Weitere Beispiele für konkrete Formulierungen und ihre zugeordneten Abstraktionen:

'warm up the engine' bzw. *„Motor warmlaufen lassen"*
einen Zustand eines Objekts herbeiführen,

'reinstall the drain bolt' bzw. *„die Ablaßschraube ... anbringen"*
Verbindung herstellen (hier zusätzlich mit restitutivem Aspekt),

'start the engine' bzw. *„Motor anlassen"*
einen Betriebszustand eines Aggregats herbeiführen,

'make sure that oil is not leaking from the drain bolt' bzw. *„sichergehen, daß kein Öl an der Ablaßschraube ausläuft"*
prüfen, daß unerwünschter Zustand nicht vorliegt,

'replace plugs' bzw. *„Zündkerzen auswechseln"*
ein Teil bzw. ein Kollektiv austauschen,

' ... so you dont get the wires mixed up' bzw. *„ ... damit die Zündkabel nicht durcheinandergebracht werden"*
(nicht) einen irregulären Zustand (unabsichtlich) herbeiführen,

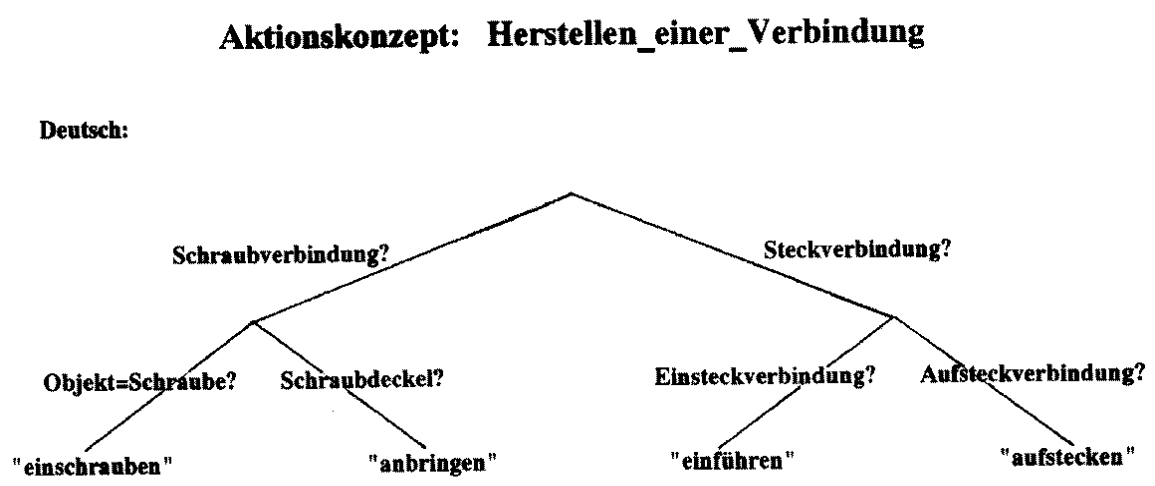

Abbildung 7.7: Lexikalisierungen von Herstellen-einer-Verbindung

'clean any dirt from the spark plug base ' bzw. *„die Zündkerzenbasis von jeglichem Schmutz befreien“*
einen erwünschten Zustand absichtlich herbeiführen und eine Substanz von einem Ort entfernen,

'discard the old plug' bzw. *„die alte Zündkerze wegwerfen“*
ein Teil entsorgen,

'check the engine oil' bzw. *„den Motorölstand kontrollieren“*
einen Meßvorgang durchführen,

'wipe off the dipstick' bzw. *„den Tauchmeßstab abwischen“*
eine Substanz von einem Ort entfernen,

'add oil' bzw. *„Öl nachfüllen“*
eine Substanz hinzufügen (zu vorhandener Substanz).

7.4.3 Lexikalisierung durch Klassifikation

Eine wichtige Maxime für den Entwurf der Wissensbasis war, die Sprachunabhängigkeit der Modellierung sicherzustellen. Die Arbeit mit mehrsprachigem Material gab dabei wichtige Anhaltspunkte. Wird nur mit einsprachigem Material gearbeitet, so ist oft wesentlich weniger deutlich, wie sehr die Entscheidung über einzuführende Konzepte vom Vorhandensein bzw. Nichtvorhandensein von Wörtern als zugehörigen Lexikalisierungen beeinflußt ist.

Bei der Modellierung von Objekten spielen diese Fragen nicht die Rolle wie bei der Modellierung von Aktionen, sind aber prinzipiell auch hier zu berücksichtigen. Im folgenden diskutieren wir in erster Linie den Fall der Aktionen.

Die Einführung abstrakter Aktionskonzepte dient nicht nur der Herstellung der Sprachunabhängigkeit, sondern führt auch zu einer Ökonomie der Repräsentation.

Bei der Abstraktion steht ja im Vordergrund, zu welchem Zweck die Repräsentation eingesetzt werden soll. Gemeinsamkeiten werden im Hinblick auf diesen Aspekt bewertet und formuliert. Da die für alle Unterklassen gemeinsamen Aspekte nur bei einer abstrakten Oberklasse formuliert werden müssen, ergibt sich eine ökonomische Form der Darstellung.

Wie sich bei den kontrastiven Untersuchungen zeigte, unterscheiden sich die einzelsprachlichen Ausprägungen auf der anderen Seite sehr deutlich in der Wahl ihres lexikalischen Materials. Eine der Dimensionen der Variabilität ist hier die entlang einer Achse von sehr spezifischen zu eher allgemeinen Verben. Während im englischen Text eher allgemeine Verben bevorzugt werden, zeichnet sich der deutsche — und in vergleichbarem Maße auch der französische — durch die Wahl spezifischer Verben aus.

Das Wissen über die Möglichkeiten zur Realisierung wird nun in jeweils sprachspezifischen Wissensbasen mit Lexikalisierungsregeln abgelegt. Diese Lexikalisierungsregeln haben Schnittstellenfunktion zwischen Gebietswissen und Generierung. Die Bedingungen für die Lexikalisierung nehmen Bezug auf Konstellationen der an einer Aktion beteiligten Aktanten. Ein häufig anzutreffender Fall ist der, daß spezifische Lexikalisierungen möglich sind, wenn bei einer Aktion bestimmte Typen bzw. Kombinationen von Typen der Füller der Rollen vorliegen. Die allgemeineren Verbalisierungen sind dadurch in vielen Fällen nicht blockiert, sondern ebenfalls möglich, klingen aber nicht in dem Maße adäquat oder sprachgewandt wie die spezifischeren.

Gelegentlich geht eine spezifische Verbalisierung mit Inkorporation einher, d.h. eine vorhandene Rolle wird an der sprachlichen Oberfläche nicht mehr realisiert, weil sie bereits in der Verb–Bedeutung vollständig aufgeht. Ein Beispiel hierfür ist die Verbalisierung 'fliegen' für ein allgemeines Aktionskonzept der Ortsveränderung, wenn die Instrumentenrolle mit einer Instanz von 'Flugzeug' gefüllt ist (also: *„Der Manager flog nach Paris."*, und nicht: *„Der Manager flog mit dem Flugzeug nach Paris."*). Allgemeinere Formulierungen sind ebenso möglich (also: *„Der Manager reiste mit dem Flugzeug nach Paris."*). Ist die Information über den Füller der Rolle nicht vollständig in die Verb–Bedeutung inkorporiert, was insbesondere dann der Fall ist, wenn eine zusätzliche Information über die bloße Typ–Information hinaus vermittelt werden soll, so wird der Rollenfüller in seiner detaillierteren Version dennoch realisiert (also z.B. *„Der Manager flog mit seinem eigenen Flugzeug nach Paris."* oder, wenn dies relevant ist, *'"Der Manager flog mit einer Lufthansa–Maschine nach Paris."*).

Eine mögliche Darstellungsform für solche Lexikalisierungsentscheidungen sind die bereits von Goldman [Goldman 1975] in seinem BABEL–System verwendeten sogenannten 'Diskriminationsnetze'. Steht ein Klassifikationssystem zur Verfügung, so läßt sich der Aufwand für die Definition und Nutzung von Lexikalisierungsregeln im Vergleich zu den traditionellen Ansätzen (wie Diskriminationsnetzen) erheblich reduzieren. Der Klassifikator kann hier nämlich sowohl den Aufbau der Subsumptionshierarchie zwischen verschiedenen Lexikalisierungen übernehmen wie auch die automatische Einordnung einer speziellen Aktionsinstanz in diese Subsumptionshier-

archie. Auf genau diese Weise wird die sprachspezifische Lexikalisierung in TECH-DOC realisiert.

Die Lexikalisierungswissensbasis verweist auf die Gebietswissensbasis, ist aber von ihr separierbar, da sie unter einem eigenen Wurzelknoten angesiedelt ist ('lexicalized-concept' genannt). Lexicalized–concepts werden subklassifiziert bezüglich der Sprache, in der sie verwendet werden können. Sprachspezifisch lexikalisierte Konzepte werden definiert als Unterkonzepte von 'lexicalized-concept' und zumindest einem, gegebenenfalls auch mehreren, gebietsspezifischen Konzepten, sowie durch die Angabe von hinreichenden Bedingungen für die Klassifikation einer Instanz unter das lexikalisierte Konzept.

Diese Bedingungen nehmen dann Bezug auf Information über die Rollenfüller. In der Abbildung 7.5 werden etwa unterschiedliche Lexikalisierungen für das abstrakte Aktionskonzept „Aufheben einer auflösbaren Verbindung" definiert, bei denen die Auswahl jeweils erfolgt nach dem Typ des Objekts, mit dem die Aktion durchgeführt wird. Der Klassifikator liefert automatisch das spezifischste lexikalisierte Konzept zu einer Aktionsinstanz (wenn dieses existiert) oder eine Auswahlliste, falls es mehrere spezifische, aber nicht miteinander in Subsumptionsbeziehungen stehende Lexikalisierungen gibt.

In einem anderen Modus kann auch eine strukturierte Liste aller möglichen Lexikalisierungen ausgegeben werden. Im Falle konkurrierender, miteinander nicht vergleichbarer Lexikalisierungen muß dann eine Feinentscheidung getroffen werden.

Einige der Möglichkeiten zeigen die im folgenden dokumentierten Beispiele anhand von Systemläufen.[6]

In der Gebietswissenbasis werden u.a. verschiedene Arten von Verbindungen unterschieden:[7]

```
> (overview 'connection-type)

                THING
             USER:UM-THING
          PENMAN-KB:OBJECT
       PENMAN-KB:NON-CONSCIOUS-THING
    PENMAN-KB:ABSTRACTION
CONNECTION-TYPE
CONNECTION-TYPE
    DISSOLVABLE-CONNECTION
       PLUG-IN-CONNECTION
       PLUG-ON-CONNECTION
```

[6] In mit `>` gekennzeichneten Zeilen beginnen Eingaben, nach diesen folgen – ebenfalls in `Schreibmaschinenschrift` – ggf. die Ausgaben des Systems.

[7] OVERVIEW erlaubt, auf ein Konzept (hier: CONNECTION-TYPE) fokussierte Ausschnitte aus dem jeweils aktuellen Subsumptionsgraphen einer LOOM–Wissensbasis zu zeigen. Zunächst werden (ausgehend vom Wurzelknoten, hier: THING) die Vorfahren gezeigt, danach die Nachkommen und die Instanzen (hier zur besseren Übersicht nur an den Blättern). Die Einrückungstiefe verdeutlicht die Vererbungstiefe.

```
        instances :
        SPARK-PLUG-WIRE-INSTANCE
        ********************
      SCREW-CONNECTION
         SCREW-IN-CONNECTION
           instances :
           SPARK-PLUG-INSTANCE
           ********************
         SCREW-ON-CONNECTION
           instances :
           OIL-FILLER-CAP-INSTANCE
           ********************
      STICK-IN-CONNECTION
        instances :
        DIPSTICK-INSTANCE
        ********************
NIL
```

Ebenfalls zur Gebietswissensbasis gehören abstrakte Aktionskonzepte wie „Aufheben einer auflösbaren Verbindung“:

```
> (defconcept disconnect-dissolvable-connection
  :is (:and penman-kb::directed-action :primitive))
```

Sprachspezifische Lexikalisierungswissensbasen hängen unter dem Wurzelkonzept LEXICALISED-CONCEPT. Durch Angabe der jeweiligen Sprache wird zwischen den sprachspezifischen Unterbäumen „hin- und hergeschaltet“.

```
> (defrelation for-language
     :domain lexicalised-concept  :range languages-supported
      :ATTRIBUTES (:CLOSED-WORLD))

> (defset languages-supported
  :is (:one-of penman::*english* penman::*french*  penman::*german*))

> (defconcept lexicalised-concept-for-german
 :is (:and lexicalised-concept (:filled-by for-language penman::*german*)))

> (defconcept lexicalised-concept-for-english
 :is (:and lexicalised-concept (:filled-by for-language penman::*english*)))
```

Wie läßt sich damit erreichen, daß im Englischen ‘remove’ gewählt wird und im Deutschen spezifischere Lexikalisierungen verwendet werden?

Zunächst die Definition für das Englische:[8]

[8]Lies: „remove-lex ist lexicalised-concept für das Konzept disconnect-dissolvable-connection und die Sprache Englisch“.
Der Appendix -LEX dient hier lediglich zurVermeidung von Namenskonflikten, da auch im Gebietsmodell aus dem Englischen inspirierte Bezeichner verwendet werden.

```
> (defconcept remove-lex
 :is (:and disconnect-dissolvable-connection lexicalised-concept
   (:filled-by for-language penman::*english*)))
```

Für die spezifischeren Lexikalisierungen im Deutschen müssen Selektionsbedingungen definiert werden. Hier wird dabei auf den Typ des Füllers der ACTEE–Rolle Bezug genommen:[9]

```
(defconcept abziehen
 :is (:and disconnect-dissolvable-connection lexicalised-concept-for-german
   (:the penman-kb::actee PLUG-ON-CONNECTION)))

(defconcept herausziehen
 :is (:and disconnect-dissolvable-connection lexicalised-concept-for-german
   (:the penman-kb::actee STICK-IN-CONNECTION)))

(defconcept abnehmen
 :is (:and disconnect-dissolvable-connection lexicalised-concept-for-german
   (:the penman-kb::actee SCREW-ON-CONNECTION)))

(defconcept herausschrauben
 :is (:and disconnect-dissolvable-connection lexicalised-concept-for-german
   (:the penman-kb::actee SCREW-IN-CONNECTION)))
```

Auch die unspezifische Lexikalisierung 'entfernen' für disconnect-dissolvable-connection (analog zu remove-lex) soll zugänglich sein. Dabei sind keine Selektionsbedingungen mit Bezug auf das Gebietsmodell definiert:

```
(defconcept entfernen
 :is (:and disconnect-dissolvable-connection lexicalised-concept
   (:filled-by for-language penman::*german*)))
```

LOOMs Klassifikator subsumiert die spezifischeren Lexikalisierungen korrekt:[10]

```
> (overview 'entfernen)

                    THING
                 USER:UM-THING
              PROCESS
           PENMAN-KB:MATERIAL-PROCESS
        PENMAN-KB:CONCEPT_30
     PENMAN-KB:DIRECTED-ACTION
   DISCONNECT-DISSOLVABLE-CONNECTION
```

[9] Lies z.B.: „abziehen ist lexicalised-concept für das Konzept disconnect-dissolvable-connection und die Sprache Deutsch genau dann, wenn der Füller der Rolle ACTEE vom Typ PLUG-ON-CONNECTION ist."

[10] Wegen der Mehrfachvererbung ergeben sich die beiden getrennten Wege zu THING.

```
         THING
      LEXICALISED-CONCEPT
   LEXICALISED-CONCEPT-FOR-GERMAN
ENTFERNEN
ENTFERNEN
   ABNEHMEN
   ABZIEHEN
   HERAUSSCHRAUBEN
   HERAUSZIEHEN
NIL
```

Generiert wird aus Instanzen in der Gebietswissensbasis, also z.B. aus[11]

```
> (po (fi disc-2))

(TELL
  (:ABOUT DISC-2
          DISCONNECT-DISSOLVABLE-CONNECTION
          PENMAN-KB:CONCEPT_28
          (PENMAN-KB:ACTEE DIPSTICK-INSTANCE)))
```

Diese Aktionsinstanz kann nun unterschiedlich lexikalisiert werden.

```
> (lexicalize 'disc-2 'PENMAN::*GERMAN*)
HERAUSZIEHEN
> (lexicalize 'disc-2 'PENMAN::*ENGLISH*)
REMOVE-LEX
```

In den Einzelsprachen wird die jeweils spezifischste Lexikalisierung gewählt. Alle definierten Lexikalisierungen (in abnehmender Spezifizität) für eine Instanz und eine Sprache liefert

```
> (all-lexicalisations 'disc-2 'penman::*german*)
(HERAUSZIEHEN ENTFERNEN)
> (all-lexicalisations 'disc-2 'penman::*english*)
(REMOVE-LEX)
```

Die Lexikalisierungsregeln sind also eine sprachabhängige und in Teilen auch gebietsabhängige Wissensbasis. Der Grad der Gebietsabhängigkeit ist allerdings unterschiedlich. In dem Maße, wie für die Auswahl der Lexikalisierungen lediglich abstrakte Konzepte herangezogen werden, ist eine höhere Gebietsunabhängigkeit gewährleistet.

[11] lies: „DISC-2 ist eine Instanz von DISCONNECT-DISSOLVABLE-CONNECTION mit der Instanz DIPSTICK-INSTANCE als Füller der Rolle PENMAN-KB:ACTEE.“. Mit fi (für find-instance) kann zu einem Bezeichner die zugehörige Instanz aus der Wissensbasis geholt werden, mit po (für print-object) wird die mit dieser Instanz assoziierte Information formatiert ausgegeben.

7.4.4 Lexikalisierung und syntaktische Struktur

Eine abstrakte Aktion mit technischen Geräten ist die, sie gemäß ihrem bestimmungsgemäßen Gebrauch handzuhaben. Für unterschiedliche Geräte wird diese abstrakte Operation zumindest im Deutschen unterschiedlich lexikalisiert. Das Englische neigt auch hier wieder zu allgemeinen Termen, wie etwa in dem folgenden Beispiel: *'apply both brakes'*, hier stand im deutschen Text demgegenüber dann *„Handbremse anziehen und Fußbremse betätigen"*. Mit anderen Worten: da für die Handbremse eine sehr spezifische Lexikalisierung der Bedienungsaktion existiert (*„anziehen"*) und andererseits das allgemeine *„betätigen"* für diese Bremse zumindest unelegant klingt, wird hier im Deutschen jede Bremse einzeln angesprochen. Im Englischen wird lediglich gesagt, daß beide Bremsen zu betätigen sind.

7.4.5 Perspektive bei der Lexikalisierung

In die Lexikalisierung fließt u.a. die Perspektive ein, also die Betonung eines speziellen Aspekts eines Objekts oder einer Aktion.

Ein Beispiel: Ein Benutzer wird ein und dieselbe Aktion durchführen bei den folgenden alternativen Aufforderungen:

- *„Öleinfülldeckel entfernen"*.
- *„Öleinfülldeckel öffnen"*.
- *„Öleinfülldeckel abschrauben"*.

„Entfernen" ist dabei die am wenigsten spezifische Formulierung und kann unabhängig von der Art der Verbindung verwendet werden. *„Abschrauben"* wiederum läßt sich nur auf Teile anwenden, die als Schraubverbindungen klassifiziert sind (ein Ersetzungstest mit beispielsweise *„Zündkerze"* macht dies deutlich). *„Öffnen"* schließlich kann nur für Teile verwendet werden, zu deren *Funktion* das Verschließen eines Behälters gehört (*„* Zündkerze öffnen"* ist daher inakzeptabel).

In einer Aktion des Nachfüllens oder Flüssigkeit–Ablassens ist der Aspekt des Verschlusses des Behältnisses besonders relevant. Hier ist *„Öffnen"* als Lexikalisierung der Aktion gut geeignet. Soll zusätzlich auch die Art und Weise, in der dies geschieht, unterstrichen werden, so läßt sich dies durch den Zusatz *„durch Drehen im Uhrzeigersinn"* o.ä. ausdrücken.

7.5 Qualitative Simulation von Abläufen

In der dem TECHDOC–System zugrundeliegenden Modellierung werden Aktionskonzepte dadurch charakterisiert, daß mit ihnen Vorbedingungen und Nachbedingungen assoziiert werden. Vorbedingungen müssen erfüllt sein, damit es möglich bzw. sinnvoll ist, die Aktion durchzuführen. Die Nachbedingungen charakterisieren die durch die Aktion in der Welt dann herbeigeführten Veränderungen (vgl. Abb. 7.8).

```
(defconcept warm_up
  :is (:and change-attribute
    (:filled-by attribute temperature)
    (:filled-by direction-of-change increase))
  :annotations
       ((action-preconditions
           ((:fail (at-least-warm-object (penman-kb::actee ?self)))))
        (action-postconditions
           ((temperature-value (penman-kb::actee ?self) warm)))))
```

Abbildung 7.8: Definition von Vor- und Nachbedingungen für eine Aktion

Diese Informationen können in deklarative Weise mit den Aktionskonzepten assoziiert werden und sind dann für eine Reihe von Prozessen nutzbar. Zum einen lassen sich damit Wartungsabläufe simulieren: Für jede Aktion in einem Plan wird geprüft, ob die Vorbedingungen für ihre Durchführung erfüllt sind. Ist dies der Fall, werden die Nachbedingungen herbeigeführt. Technisch wird dies dadurch realisiert, daß Vorbedingungen mit 'ASK' in der Wissensbasis abgefragt werden. Nachbedingungen werden in einer solchen Form formuliert, daß sie als Assertionen mit 'TELL' in die Wissensbasis eingetragen werden (vgl. 6.5.3). Dies führt zu Veränderungen (z.B. durch Reklassifikation von Instanzen); ein neuer Zustand der Wissensbasis stellt sich ein. Oft sind die neu eingetretenen Zustände dann realisierte Vorbedingungen für die nachfolgende Aktion.

Die in den Vor- und Nachbedingungen kodierte Information kann auch für das Planen herangezogen werden. Wenn eine Bedingung (z.B. die Vorbedingung einer Aktion) nicht erfüllt ist, kann nach Aktionen gesucht werden, durch deren Nachbedingung die gesuchte Bedingung erfüllt wird. Sind die Vorbedingungen einer solchen Aktion bereits erfüllt, so kann diese durchgeführt werden, und die gewünschte Bedingung stellt sich ein. Ist dies nicht der Fall, wird das Verfahren rekursiv weitergeführt, bis eine Kette von Aktionen gefunden wurde, die vom aktuellen Zustand der Wissensbasis aus erlaubt, die gewünschte Bedingung herbeizuführen. Läßt sich eine solche Kette nicht finden, wird der Benutzer darüber informiert und hat dann die Option, von sich aus die Gültigkeit der Bedingung zu assertieren.

Die geschilderten Möglichkeiten der aktiven Nutzung des modellierten Wissens über Aktionen stellen in gewisser Weise eine operationale Semantik der Aktionskonzepte dar und sind eine wichtige Verankerung (grounding) der aus ihnen produzierten sprachlichen Äußerungen.

Ein weiterer wichtiger Aspekt ist, daß modellierte Vor- und Nachbedingungen die Ableitung von Information erlauben, die zur Beantwortung von Fragen herangezogen werden kann (vgl. 8.3.1). Zu den Fragetypen gehören 'warum'–Fragen („Warum wird eine bestimmte Aktion durchgeführt?“ — zum Beispiel, um Vorbedingungen für eine andere Aktion oder bestimmte wünschenswerte Zustände herbeizuführen) und Fragen

nach möglichen Konsequenzen (also: „Welche Auswirkungen hat die Durchführung einer Aktion?“).

Abb. 7.9 auf Seite 166 zeigt einen kleinen Ausschnitt aus einer qualitativen Simulation des Plans für Ölwechsel an einem PKW.

7.5.1 Verbindung von Simulation und Lexikalisierung

Sowohl Objekte wie auch Aktionen können inhärente wie auch situationsbezogene Eigenschaften haben. Wissen über inhärente Eigenschaften wird bereits bei der Definition der jeweiligen Konzepte abgelegt. Bei Substanzen kann hierzu etwa gehören, gegen welche anderen Materialien sie sich bei Oberflächenkontakten aggressiv verhalten. Bei den Aktionen gehören zu den inhärenten Eigenschaften die definierenden Vor- und Nachbedingungen.

Situationsbezogene Eigenschaften können nicht von vornherein festgelegt werden, sondern müssen sich in einer konkreten Situation für die jeweiligen Instanzen bestimmen lassen. Bei Instanzen von Objekten kann es etwa der aktuelle Temperaturwert oder die Werte anderer veränderlicher Größen sein. Bei Instanzen von Aktionen ist für die Lexikalisierung wichtig zu wissen, ob es sich um eine repetitive oder restitutive Aktion handelt. Werden die Schritte der Pläne (wie in 7.5) simuliert, so läßt sich dieser Sachverhalt aus dem aufgezeichneten Verlauf der Simulation ableiten.

Eine Aktionsinstanz wird als *repetitiv* angesehen, wenn dieselbe Aktion im Verlauf der aktuellen Wartung bereits mindestens einmal mit demselben Objekt durchgeführt wurde.

Restitutiv ist eine Aktion, die einen Zustand wiederherstellt, der zu einem früheren Zeitpunkt bei der aktuellen Simulation gegolten hat und der durch eine Aktion verändert worden ist.

Beispiele für Paare von solchen Aktionen:

DRAIN <substance> from <container> vs.
FILL <substance> into <container>
DISCONNECT-DISSOLVABLE-CONNECTION <object> vs.
ESTABLISH-DISSOLVABLE-CONNECTION <object>.

Es gibt eine ganze Reihe englischer Verben, die durch die Vorsilbe *‘re-’* den restitutiven Aspekt in die Verbalisierung inkorporiert haben (z.B. *reinstall*, *refill* u.a.). Gibt es kein entsprechendes deutsches Verb mit Inkorporation, so kann der restitutive Charakter durch das Adverb *„wieder“* Ausdruck finden. Der repetitive Charakter einer Aktion kann durch das Adverb *„erneut“* unterstrichen werden.

Ein Beispiel:

```
> (defconcept establish-dissolvable-connection
  :is (:and penman-kb::directed-action :primitive))

;;;*************************************************************
```

```
;;; lexicalised-concepts

(defconcept install-lex
 :is (:and establish-dissolvable-connection lexicalised-concept
   (:filled-by for-language penman::*english*)))>

;;; Die spezifischeren dt. Lexikalisierungen:

(defconcept anbringen
 :is (:and establish-dissolvable-connection lexicalised-concept
   (:filled-by for-language penman::*german*)
   (:the penman-kb::actee
 (:or PLUG-ON-CONNECTION SCREW-CONNECTION))))

(defconcept einfuehren
 :is (:and establish-dissolvable-connection lexicalised-concept-for-german
   (:the penman-kb::actee STICK-IN-CONNECTION)))

  ...
```

In der englischen Lexikalisierungswissensbasis wird REINSTALL-lex definiert, als INSTALL-lex mit action-modifier restitutive. Zusätzlich wird angemerkt, daß bei dieser Lexikalisierung dann die Rolle action-modifier inkorporiert ist (incorporated-roles) und damit nicht mehr sprachlich ausgedrückt wird.

```
(defconcept reinstall-lex
  :is (:and install-lex
    (:filled-by action-modifier restitutive))
  :annotations ((incorporated-roles (action-modifier))))
```

Sei EDC-1 eine Instanz von ESTABLISH-DISSOLVABLE-CONNECTION mit DIPSTICK-INSTANCE als ACTEE.

```
> (po (fi edc-1))
(TELL
  (:ABOUT EDC-1
          ESTABLISH-DISSOLVABLE-CONNECTION
          PENMAN-KB:CONCEPT_28
          (PENMAN-KB:ACTEE DIPSTICK-INSTANCE)))

> (lexicalize 'edc-1 'PENMAN::*german*)
EINFUEHREN
> (lexicalize 'edc-1 'PENMAN::*english*)
INSTALL-LEX
```

Zusätzliche Information über den ACTION-MODIFIER ...

```
> (tellm (:about edc-1 (action-modifier restitutive)))
|I|EDC-1(ESTABLISH-DISSOLVABLE-CONNECTION PENMAN-KB:CONCEPT-28)
```

... und die Konsequenz bei der Lexikalisierung:

```
> (lexicalize 'edc-1 'PENMAN::*german*)
EINFUEHREN
> (lexicalize 'edc-1 'PENMAN::*english*)
REINSTALL-LEX
```

7.6 Zusammenfassung

Das System TECHDOC dient als Plattform für weitere Experimente und wird bei den geplanten Folgearbeiten in diesem Sinne genutzt werden. Zu den Fragestellungen, die dabei von Interesse sind, gehören:

- Wie kann die Umsetzung von Planstrukturen in Dokumentstrukturen möglichst flexibel gestaltet werden?
- Welche anderen Arten von strukturierten Plänen sollten unterstützt werden können?
- Wie kann Besonderheiten der einzelsprachlichen Realisierung noch besser Rechnung getragen werden?
- Welche Stilparameter sind für unterschiedliche Nutzungen und Adressaten relevant und wie wirken sie sich bei der linguistischen Realisierung aus?

Andere Fragestellungen betreffen die Wissensrepräsentation, so muß z.B. für Aspekte wie Modifikatoren für Prozesse oder Graduierungen von Eigenschaften die Repräsentation erweitert werden.

Schließlich soll in der Zusammenarbeit mit industriellen Partnern anhand realer Informationsbestände für ein technisches Produkt bearbeitet werden, mit welchen Techniken und mit welchem Aufwand Informationen aus z.B. Produktdatenmodellen oder Prozeßmodellierungen für das TECHDOC–System nutzbar gemacht werden können.

```
> (simulate (list (fi 'warm_up-engine_instance)))

 ACTION |I|WARM_UP-ENGINE_INSTANCE(WARM_UP PENMAN-KB:CONCEPT_28) to perform:

(TELL
  (:ABOUT WARM_UP-ENGINE_INSTANCE
          WARM_UP
          PENMAN-KB:CONCEPT_28
          (PENMAN-KB:ACTEE ENGINE_INSTANCE1)
          (ATTRIBUTE 'TEMPERATURE)
          (DIRECTION-OF-CHANGE 'INCREASE)))

 as generated text: "Warm up the engine."

  CHECK-PRECONDITIONS called for |I|WARM_UP-ENGINE_INSTANCE(WARM_UP ...)
   and ?action-type WARM_UP

 PRECONDITIONS: ((:FAIL (AT-LEAST-WARM-OBJECT (PENMAN-KB:ACTEE ?SELF))))

 check-preconditions: overall result is TRUE

 establish-postconditions for |I|WARM_UP-ENGINE_INSTANCE(WARM_UP ...) :

 current state of (PENMAN-KB:ACTEE ?SELF)
(TELL
  (:ABOUT ENGINE_INSTANCE1
          NORMAL-OR-COLD-OBJECT
          ENGINE
          PENMAN-KB:CONCEPT_10
          (PART-OF GENERIC-VEHICLE_INSTANCE)
          (TEMPERATURE-VALUE 'VERY-COLD)))

 will TELL that (TEMPERATURE-VALUE (PENMAN-KB:ACTEE ?SELF) WARM)
  shall hold for |I|ENGINE_INSTANCE1(NORMAL-OR-COLD-OBJECT ENGINE ...)

 changed state of (PENMAN-KB:ACTEE ?SELF):

(TELL
  (:ABOUT ENGINE_INSTANCE1
          AT-LEAST-WARM-OBJECT
          ENGINE
          PENMAN-KB:CONCEPT_10
          (PART-OF GENERIC-VEHICLE_INSTANCE)
          (TEMPERATURE-VALUE 'WARM)))

 as generated text: "The engine was warmed up."
```

Abbildung 7.9: Ausschnitt aus einer qualitativen Simulation