

Generierung natürlicher Sprache

Stephan Busemann

Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH
Stuhlsatzenhausweg 3, W-6600 Saarbrücken 11

Hans-Joachim Novak

IBM Deutschland GmbH, AE SAT
Postfach 80 08 80, W-7000 Stuttgart 80

Zusammenfassung

Dieser Aufsatz beschreibt das interdisziplinäre Forschungsgebiet Generierung natürlicher Sprache und gibt einen Überblick über den gegenwärtigen Stand der Kunst. Behandelt werden Ansätze aus der Psycholinguistik, Planungs- und Entscheidungsverfahren aus der sprachverarbeitenden KI und Verfahren auf der Grundlage moderner Grammatikformalismen. Die jeweiligen Forschungsziele und -methoden werden dargestellt.

This report describes the interdisciplinary research field of natural language generation and gives an overview of the current state of the art. The paper presents psycholinguistic approaches, AI planning and decision-making processes, and generators based on modern grammar formalisms. For each case, the research goals and methods are described.

Dieser Aufsatz erscheint in geringfügig geänderter Fassung in dem Band *Künstliche Intelligenz* (Hrsg. G. Görz), Addison-Wesley Verlag, München, 1993.

Diese Arbeit wurde teilweise finanziell unterstützt durch das Bundesministerium für Forschung und Technologie (FKZ ITW 9002).

© Deutsches Forschungszentrum für Künstliche Intelligenz 1992

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Deutsches Forschungszentrum für Künstliche Intelligenz, Kaiserslautern, Federal Republic of Germany; an acknowledgement of the authors and individual contributors to the work; all applicable portions of this Copyright notice. Copying, reproducing, or republishing for any other purpose shall require a licence with payment of fee to Deutsches Forschungszentrum für Künstliche Intelligenz.

Inhaltsverzeichnis	?
1 Einleitung	2
1.1 Entwicklung des Forschungsgebiets.	2
1.2 Die Modularisierung des Generierungsprozesses.	4
1.3 Zentrale Fragestellungen und Forschungsrichtungen.	4
1.4 Überblick	7
2 Ein Modell menschlicher Sprachproduktion	8
2.1 Inkrementelle Sprachproduktion	10
2.2 Die präverbale Bedeutungsstruktur.	11
2.3 Die Konstruktion der präverbalen Bedeutungsstruktur.	13
3 Planungs- und Entscheidungsprozesse	19
3.1 Verschiedene Architekturen	19
3.2 Inhaltsbestimmung: Das Auswahlproblem	20
3.3 Formbestimmung	29
3.4 Sprachliche Realisierung	31
3.5 Wortwahl	33
3.6 Systemische Grammatik	36
4 Generierung mit modernen Grammatikformalisen	38
4.1 Inkrementelle Generierung mit Baumadjunktionsgrammatiken (TAG)	40
4.2 Generalisierte Phrasenstruktur-Grammatiken (GPSG).	42
4.3 Lexikalisch-Funktionale Grammatik (LFG).	46
4.4 Steuerung durch den „semantischen Kopf“.	47
4.5 Bidirektionale Verarbeitung	51
Literatur	52
Index	60

1 Einleitung

Generierung von natürlicher Sprache (NL)¹ ist ein interdisziplinäres Forschungsgebiet im Schnittbereich von Informatik, KI, Linguistik und Psychologie/Kognitionswissenschaft, das durch zahlreiche verschiedene Strömungen gekennzeichnet ist. Die zentrale Fragestellung lautet:

- Wie wird aufgrund eines zugrundeliegenden Ziels eines Sprechers² ein für den Hörer kommunikativ adäquater, natürlichsprachlicher Diskurs maschinell erzeugt?

Nicht nur unterschiedliches Erkenntnisinteresse der Forscher, sondern auch unterschiedliche Randbedingungen und Zwecke für Computersysteme zur Generierung natürlicher Sprache führen zu einer zunehmenden Diversifizierung des Gebiets. Andererseits werden neuerdings Brücken geschlagen, um Ideen aus unterschiedlichen Ansätzen miteinander zu verbinden.

Ziel dieses Beitrags ist es, diese Entwicklungen strukturiert zu beschreiben und einen Überblick über den gegenwärtigen Stand der Kunst zu geben. Dabei wird besonderer Wert auf die verschiedenen Forschungsziele und -methoden sowie die unterschiedlichen Grundannahmen gelegt.

Natürlich ist diese Arbeit durch unsere Rezeption des Forschungsgebiets geprägt. Insofern wird im Unterschied zu früheren Übersichtsartikeln [McDonald, 1987; McKeown und Swartout, 1988] zum einen der interdisziplinäre Charakter des Gebiets stärker betont und zum ändern den kognitiven und den linguistischen Aspekten ein höheres Gewicht beigemessen.³

1.1 Entwicklung des Forschungsgebiets

Frühe Arbeiten zur Generierung blieben abseits der Hauptforschungsströmungen. Im Vergleich zum Aufwand im Bereich Analyse, die seit den frühen Sechziger Jahren systematisch erforscht wird, wurde Generierung bis Anfang der Achtziger Jahre vernachlässigt. Verschiedene Gründe lassen sich hierfür anführen:

- Vielfach waren frühe NL Systeme so ausgelegt, daß sie ohne oder mit geringen Sprachgenerierungsfähigkeiten auskamen. Die Dialog-Initiative liegt dann ausschließlich beim Benutzer und nicht auch beim System.
- Die noch recht niedrigen Anforderungen konnten mithilfe einfacher Techniken (Musterabgleich) erfüllt werden.

¹Wir bevorzugen die international gebräuchliche Abkürzung NL für *natural language* und bezeichnen damit sowohl *natürliche Sprache* als auch *natürlichsprachlich* (in den verschiedenen Flexionsformen).

²Im folgenden werden männliche und weibliche Personen sowie Maschinen als *Sprecher* bezeichnet, wenn sie gesprochene oder geschriebene sprachliche Information absenden und als *Hörer*, wenn sie sie empfangen.

³Wir danken Katharina Morik, Hans Uszkoreit und Michael Zock für wertvolle Anregungen und Hinweise.

- Von Menschen werden auch qualitativ schlechte Systemäußerungen verstanden und akzeptiert, weshalb die Sprachgenerierungskomponenten nur soweit entwickelt waren, wie es für die Performanz des Gesamtsystems unbedingt nötig war.

Diese Situation hat sich seit etwa einem Jahrzehnt drastisch geändert. Mit Beginn der Achtziger Jahre wurde Generierung zu einem eigenständigen Teilgebiet der sprachorientierten KI. Damit ist es jung im Vergleich zum Bereich der Analyse von NL, und noch immer arbeiten weitaus weniger Wissenschaftler an Generierung als an Analyse. Doch tragen regelmäßige Generierungsworkshops zum gegenseitigen Kennenlernen der Wissenschaftler und ihrer Ideen bei, so daß das Gebiet sich noch immer schnell entwickelt. Aus den Beiträgen zu diesen Workshops⁴ gehen regelmäßig Sammelbände hervor, die einen Überblick über die aktuelle Forschung bieten [Kempen, 1987b; Paris *et al.*, 1991; Zock und Sabah, 1988; Dale *et al.*, 1990; Horacek und Zock, 1992; Dale *et al.*, 1991]; vgl. außerdem [McDonald und Bolc, 1988].

Seit 1990 gibt es eine *special interest group on NL generation* (SIGGEN) in der ACL, womit nach der faktischen nun auch die formelle Etablierung des Gebiets feststeht.

Was wurde bisher erreicht? Fast vor einem Jahrzehnt hatten führende Forscher in der sprachorientierten KI die Anforderungen an das Wissen eines leistungsfähigen und flexiblen Generierungssystems formuliert [Mann *et al.*, 1982]. Benötigt wird demzufolge

- eine umfassende, linguistisch fundierte Grammatik,
- ein Wissensrepräsentationsformalismus, der unterschiedliche Arten von Wissen darstellen kann,
- ein Modell des intendierten Lesers eines Texts und
- ein Modell für Diskursstruktur und Kontrolle.

In den vergangenen Jahren wurde auf jedem dieser Teilgebiete Beachtliches geleistet und in Generierungssystemen nutzbar gemacht. Jedoch ist uns kein System bekannt, das alle genannten Bereiche in gleichermaßen tief ausgeprägter Weise berücksichtigt. Nach wie vor besteht der Bedarf, Einzelanstrengungen in ein umfassenderes System zu integrieren. Dies ist außerordentlich schwierig, da die theoretischen Grundlagen nur Teilbereiche abdecken und mühsam aufeinander abgestimmt werden müssen.

Ein wichtiger Fortschritt bestünde in der Möglichkeit, Generierungssysteme evaluierend zu vergleichen. Dies stößt auf die Schwierigkeit, daß den Systemen unterschiedliche Ausgangsstrukturen zugrundeliegen. Ein weiteres Problem besteht in der Tatsache, daß eine theoretisch saubere Modellierung nicht unmittelbar zu einer guten Performanz des Systems führt. Umgekehrt können *ad hoc*-Methoden ein erstaunlich gutes Systemverhalten erzeugen.

⁴1985 auf Burg Stettenfels bei Stuttgart, 1986 in Nijmegen, 1987 auf Catalina Island (California), 1988 in der Abbeie de Royaumont bei Paris, 1989 in Edinburgh, 1990 Dawson (Pennsylvania), 1991 in Judenstein bei Innsbruck, 1992 in Trento, 1993 in Pisa.

Für eine Evaluation müssen Maße definiert werden, die dies alles berücksichtigen.⁵ Diese Maße müssen schließlich gewichtet werden, um zu einer Gesamtbewertung zu gelangen. Überlegungen hierzu werden gegenwärtig auf Workshops in den USA angestellt. Als Ergebnis erhofft man sich u.a. Hinweise, wie eine allgemeine Technologie entwickelt werden müßte, die sich Anwender zunutze machen könnten.

1.2 Die Modularisierung des Generierungsprozesses

Die sprachliche Äußerung eines Gedankens kann als komplexer *Entscheidungsprozeß* aufgefaßt werden [McDonald, 1983]. Der Sprecher muß einerseits entscheiden, *was* gesagt werden soll und andererseits, *wie* etwas zu sagen ist. Zum ersten Bereich gehört die Auswahl des Redethemas und der Aufbau einer argumentativen Struktur (etwa bei Erklärungen). Zum zweiten Bereich zählen u.a. Wahl des Verbgenus, Verwendung von Proformen, Bestimmung von Topikalisierung oder Extraposition.

Die Dichotomie geht auf [Thompson, 1977] zurück (dort: strategische vs. taktische Generierung) und findet sich in den meisten heutigen Ansätzen, allerdings in unterschiedlichen Ausprägungen, was auf die andauernde theoretische Unklarheit bezüglich des Generierungsprozesses schließen läßt, und in unterschiedlicher Terminologie. So unterscheidet Kempen zwischen Konzeptualisierer und Formulierer [Kempen und Hoenkamp, 1987] oder McDonald zwischen Sprecherkomponente und Sprachkomponente [McDonald, 1983].

Die intuitiv naheliegende Aufteilung in inhaltsbestimmende und formbestimmende Generierung hat in vielen Generierungssystemen zu einer zweistufigen Architektur geführt (vgl. etwa TEXT [McKeown, 1985]). Die damit verbundene Modularisierung in zwei voneinander getrennte Prozesse (Sprachgebrauch: *What-to-Say-* bzw. *How-to-Say-It-Komponente*) erwies sich als nützlich; sie ermöglicht grundsätzlich die Arbeit an der einen Seite, ohne sich mit der vollen Komplexität der Fragen auf der anderen Seite auseinandersetzen zu müssen. Andererseits erwies sich die Modularisierung aus theoretischen Gründen als überaus problematisch, wie im folgenden ersichtlich wird.

Die Menge der getroffenen Entscheidungen führt zu einer oberflächennahen Darstellung der Bedeutung einer Äußerung, aus der aber erst noch eine Endkette erzeugt werden muß. Diesen Prozeß, der selbst keine Entscheidungen mehr erfordert, nennen wir *Realisierung*.⁶ (siehe Abb. 1). Zur Realisierung zählen Etablierung von Kongruenzrelationen, Wortstellung und Wortflexion.

1.3 Zentrale Fragestellungen und Forschungsrichtungen

Wir betrachten nun drei zentrale Fragestellungen bzw. Zielsetzungen im Bereich Generierung, um die verschiedenen Motivationen herauszuarbeiten. Allen gemeinsam ist die

⁵Dies ist keine triviale Aufgabe; nach welchen Kriterien mißt man z.B. die Leistung einer konkreten Grammatik, so daß verschiedene Ansätze vergleichbar werden?

⁶Viele englischsprachige Autoren verstehen unter *linguistic realization* die formbestimmende Generierung einschließlich Realisierung. Dies liegt für gering flektierende Sprachen wie das Englische durchaus nahe.

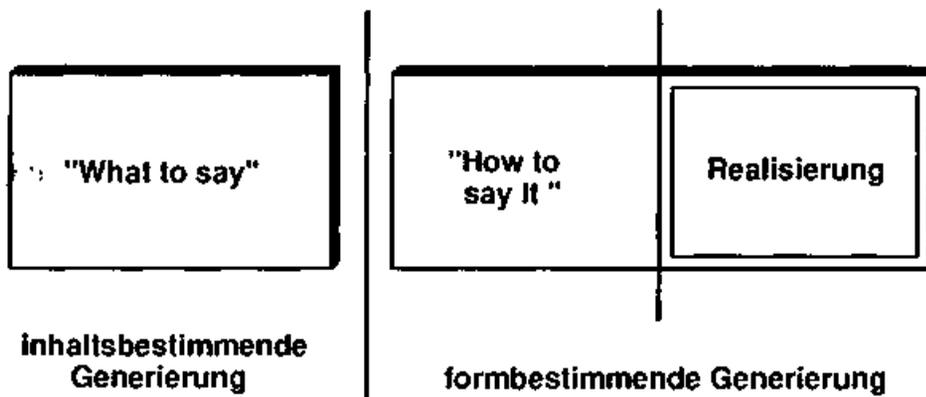


Abbildung 1: Die Unterteilung des Generierungsprozesses.

Verwendung *informatischer Mittel* insofern, als die theoretischen Ergebnisse stets implementiert werden, wenn auch wiederum mit unterschiedlichen Zielen. Die erste Frage

- Wie funktioniert der menschliche Sprachproduktionsprozeß?

ist psycholinguistisch und kognitionswissenschaftlich beeinflusst. Gegenstand dieser Frage ist der Mensch. Forschungsziel ist die kognitiv adäquate Modellierung der menschlichen Sprachproduktion, die Phänomene wie Pausen, Selbstkorrekturen und Ergänzungen erklären kann. Der Zusammenhang zwischen Inhaltsbestimmung („Denken“) und Formbestimmung („Sprechen“) steht im Zentrum vieler psycholinguistischer Untersuchungen an Probanden. Offensichtlich liegt eine Verzahnung beider Aktivitäten vor. Kennzeichnend für diese Ansätze ist ein *inkrementelles* Vorgehen, bei dem bereits formbestimmende Entscheidungen getroffen und Teile der Äußerung produziert werden, nachdem erst ein Teil der inhaltsbestimmenden Entscheidungen gefallen ist. Das inkrementelle Vorgehen legt eine parallele, kaskadierte Architektur nahe. Die Stufen der Kaskade bestehen z.B. aus der inhaltsbestimmenden, der formbestimmenden und der artikulatorischen Ebene [De Smedt, 1990]. Eine zentrale Frage bei diesem Ansatz betrifft die Größe der Segmente, die auf die jeweils nächste Kaskadenstufe fließen. Sind sie zu klein, lassen sich keine fundierten Entscheidungen hinsichtlich ihrer sprachlichen Form treffen; sind sie zu groß, geht der Effekt der flüssigen Rede verloren, da unerwünschte Pausen zwischen den Teiläußerungen auftreten. Interessant ist u.a., welche linguistischen Kategorien eine psychologische Relevanz haben. Ferner wird der Einfluß beschränkter Ressourcen (z.B. beschränkte Zeit bei Reportagen) untersucht.

Ein weiterer, kognitionswissenschaftlicher Aspekt betrifft die Verwendung *analoger* Repräsentationen, etwa bei der Modellierung von Wegbeschreibungen. Der Sprecher vollzieht eine imaginäre Wanderung vor dem Sprechen, benutzt also neben propositionalen auch depiktionale (bildhafte) Repräsentationen.

Meistens wird die Produktion *gesprochener* Sprache modelliert, auch wenn in Computersystemen Tastatur und Bildschirm als Ein-/Ausgabe-Medien verwendet werden. Implementierungen erfolgen vorwiegend zum Zweck der Theorievalidierung und beanspruchen

psychologische Adäquatheit.

Die zweite Fragestellung

- Durch welche Planungs- und Entscheidungsprozesse kann Generierung modelliert werden?

dominiert bei Sprachgenerierungsansätzen in der KI. Ähnlich wie bei der psycholinguistisch motivierten Fragestellung steht der komplexe Informationsverarbeitungsprozeß des Menschen Pate für Computermodelle. Er wird jedoch in der KI nicht durch aufgrund von empirischen Untersuchungen postulierte Gehirnaktivitäten modelliert, sondern durch Verarbeitung von explizit repräsentiertem Wissen gepaart mit Inferenzmechanismen. Die Modellierung erfolgt auf verschiedenen Granularitätsebenen und schließt oft die Ebene der Referenzsemantik ein, d.h. den expliziten Bezug sprachlicher Ausdrücke zu einem Diskursbereich. Linguistische Arbeiten vor allem in den Bereichen der Textstrukturierung, der Dialogführung und der Sprechakttheorie betonen den interdisziplinären Charakter des Gebiets.

Das theoretisch orientierte Interesse findet in der KI seinen Niederschlag in einer anwendungsorientierten Richtung, die die Implementation von Systemen mit guter Performanz anstrebt. Hier zeigt sich ein Spektrum hochgradig verschiedener Anwendungssituationen für Generierungssysteme, die unterschiedliche Leistungen erfordern:

NL Dialogsysteme. Generierung berücksichtigt die Ziele des Dialogpartners und erzeugt kommunikativ adäquate Antworten. Vielfach kommen kurze, oft elliptische Sätze vor. In einem Dialogschritt generiert ein Sprecher Bezüge auf frühere Äußerungen von Dialogpartnern und von ihm selbst. Entscheidungen werden oft aufgrund einer semantischen Repräsentation, eines Benutzermodells sowie von Wissen über Dialogstruktur und Dialogkontext getroffen.

Maschinelle Übersetzung. Texte werden unter Berücksichtigung einer anderssprachlichen Formulierung generiert. Ausgangsstrukturen sind entweder einzelsprachunabhängige Resultate des ausgangssprachlichen Analyseprozesses oder die Ergebnisse eines Transferprozesses, der bereits zielsprachliche Aspekte berücksichtigt. In beiden Fällen besteht die Aufgabe in der Formbestimmung, denn die Inhaltsbestimmung erfolgt durch vorangehende Teilprozesse. Deren Ergebnis liegt in Form einer linguistisch motivierten Repräsentation vor.

Textgenerierung. Ausgehend von Repräsentationen des propositionalen Gehalts des Textes und der Argumentationsabsicht des Sprechers werden Texte in Absatzlänge generiert, die oft mithilfe von semantischen und rhetorischen Relationen strukturiert sind. Bezüge auf vorangegangene Textteile und Berücksichtigung von außersprachlichem Wissen sichern die Kohärenz des Gesamttextes. Solche Texte werden in sehr verschiedenen Szenarios generiert, z.B. Verbalisierung von Schlußfolgerungen in Expertensystemen, Reportagen bei beschränkten zeitlichen Ressourcen, Zusammenfassungen von Arbeitsmarktberichten oder Wetterberichte.

Die dritte Fragestellung entstand relativ spät (Mitte der Achtziger Jahre) zusammen mit der Entwicklung von unifikations- bzw. constraintbasierten Grammatikformalismen:

- Gegeben sei eine semantische Repräsentation und eine Grammatik. Welche sind die gemäß der Grammatik zulässigen Endketten? .

Gegenstand der Forschung ist die Verarbeitung von NL unter der Maßgabe, daß die Kodierung von sprachlichem Wissen linguistisch-theoretischen Ansprüchen genügen muß. Im Unterschied zu den oben skizzierten Fragen steht weniger der Mensch und seine Umgebung im Mittelpunkt als vielmehr die Sprache selbst sowie die linguistischen Methoden, grammatisch wohlgeformte Sätze (selten: Texte) zu beschreiben.⁷

Die Entwicklung moderner Grammatikformalismen (z.B. Lexikalisch-Funktionale Grammatik (LFG), generalisierte Phrasenstruktur-Grammatik (GPSG), Head-gesteuerte Phrasenstruktur-Grammatik (HPSG)) hat die Beschreibung der Syntax und der logisch-semantischen Beziehungen von NL Sätzen zum Schwerpunkt. Die Verwendung solcher Formalismen bei der Generierung erlaubt eine linguistisch angemessenere Behandlung syntaktischer und syntaktisch-semantischer Phänomene, als dies in der sprachverarbeitenden KI bisher angestrebt wurde.⁸ Linguistische Methoden bewirken eine breite Abdeckung von sprachlichen Konstrukten, da die gesamte Bandbreite sprachlicher Phänomene betrachtet wird, ohne diskursbereichsspezifische Einschränkungen zu machen.

Die constraintbasierte Herangehensweise ist eine Voraussetzung dafür, daß die Grammatiken non-direktional sind: sie verhalten sich neutral bezüglich Parsing oder Generierung (häufig wird auch von *reversiblen* oder *bidirektional verwendbaren* Grammatiken gesprochen). Infolge der deklarativen Repräsentationsweise besteht Sprachverarbeitung in dem Problem, eine Relation zwischen wohlgeformten Endketten und den ihnen durch die Grammatik zugeordneten logischen Formen (LF) zu definieren. Generierung und Parsing werden hier formal als zueinander inverse Abbildungen definiert. Die Suche nach effizienten Kontrollstrategien stellt eine wichtige Aufgabe dar.

Im Rahmen der dritten Fragestellung dienen Implementationen vorrangig der Theorievalidierung und dem Test von Grammatiken. Der Computer fungiert als „Testbett“ und als Entwicklungswerkzeug. Unter wachsendem Anwendungsdruck werden constraintbasierte Generatoren oft als Realisierungskomponente eines Generierungssystems implementiert. Erst in jüngster Zeit wird der Gegenstandsbereich auf semantisch-pragmatische Phänomene (anaphorische Bezüge) und Diskursphänomene ausgedehnt.

1.4 Überblick

Der Rest des Aufsatzes gliedert sich in drei Hauptabschnitte, die sich den im vorigen Abschnitt genannten drei Fragen widmen. Abschnitt 2 beschreibt die Teilaufgaben beim Sprechen und diskutiert psycholinguistisch motivierte Sprachproduktionsmodelle. Er vermittelt außerdem einen Einblick in die Zusammenhänge zwischen dem Ziel des Sprechers,

⁷Auf Verschmelzungen mit der sprachorientierten KI wird weiter unten eingegangen.

⁸Polemische Stimmen in der KI-Szene halten syntaktische Probleme für „weitgehend gelöst“ und die Bearbeitung der „wesentlichen Fragen“ (d.h. der pragmatischen Probleme) für vordringlich.

dem vermuteten Wissensstand und den Zielen des Hörers und den sprachlichen Mitteln, die dem Sprecher zur Verfügung stehen.

In Abschnitt 3 gehen wir zuerst auf verschiedene Architekturen von Generierungssystemen ein. Danach betrachten wir Fragen der Inhaltsbestimmung, der Formbestimmung, der Realisierung und der Wortwahl. Am Ende dieses Abschnitts gehen wir kurz auf systemische Grammatiken ein, die in der Sprachgenerierung eine besondere Rolle spielen.

In Abschnitt 4 stehen Kontrollstrategien im Zentrum, die eine effiziente Realisierung mithilfe moderner linguistischer Grammatikformalismen erlauben. Dabei werden einige subtile Abhängigkeiten zwischen dem jeweiligen Formalismus, der Kontrollstrategie und der Art der Wissenskodierung aufgezeigt.

2 Ein Modell menschlicher Sprachproduktion

In der folgenden Darstellung beschränken wir uns auf den Bereich der *Produktion gesprochener Sprache*.

In [Levelt, 1989]⁹ wird ein Prozeßmodell eines Sprechers vorgestellt, das auf dem Informationsverarbeitungsparadigma beruht. Die komplexe Aufgabe, eine Äußerung zu produzieren, wird in Teilprozesse zerlegt, von denen jeder bestimmte Eingabestrukturen verarbeitet und Ausgabestrukturen erzeugt, die als Eingabe für den nachfolgenden Prozeß dienen. Levelt schlägt drei große Prozeßblöcke vor:

1. den Konzeptualisierer;
2. den Formulator;
3. den Artikulator.

Die Aufgabe des *Konzeptualisierers* besteht darin, unter Zugriff auf ein Diskursmodell, Wissen über die Sprechsituation, enzyklopädisches Wissen, etc. den Inhalt einer Äußerung festzulegen. Diese Inhaltsbestimmung ist ein komplexer Vorgang, der eine Intention des Sprechers erfordert, die Auswahl der relevanten Information, die Linearisierung der ausgewählten Information sowie permanente Buchführung darüber, was bereits gesagt wurde. Weiterhin beobachtet der Sprecher permanent Inhalt und Form seiner eigenen Äußerungen, ein Prozeß, den Levelt *monitoring* nennt. Das Ergebnis aller dieser Prozesse des Konzeptualisierers ist eine *vorsprachliche Bedeutungsstruktur* (engl. *preverbal message*). Bei der Erstellung dieser Ausgabestruktur werden zwei Planungsphasen unterschieden: Makroplanung und Mikroplanung. Makroplanung betrifft die Ebene der kommunikativen Intention des Sprechers, z.B. Aufspalten des Kommunikationsziels in mehrere Teilziele und Bestimmung der zum Erreichen der Teilziele relevanten Information. Mikroplanung hingegen legt die Form der zu äußernden Propositionen fest. Es ist also Aufgabe der Makroplanung festzulegen, daß die Assertion eines Sachverhalts (z.B. *Peter gab Maria das Buch*)

⁹Die Darstellung des Prozeßmodells folgt im wesentlichen dem Buch von Levelt, das dem interessierten Leser empfohlen wird.

der kommunikativen Intention, diesen Sachverhalt mitzuteilen, entspricht. Die Festlegung der Form (*Peter gab Maria das Buch* versus *Das Buch gab Peter Maria*), d.h. auch der Perspektive, ist Aufgabe der Mikroplanung.

Die Aufgabe des *Formulators* besteht darin, die konzeptuelle Struktur des *Konzeptualisierers* in eine linguistische Struktur zu übersetzen. Diese Übersetzung erfolgt in zwei Schritten:

1. grammatische Kodierung der konzeptuellen Struktur deren Ergebnis eine Oberflächenstruktur (eine lineare Abfolge korrekt flektierter Wörter) ist;
2. phonologische Kodierung der Oberflächenstruktur deren Ergebnis ein *Artikulationsplan*¹⁰ ist, der die Eingabe zur letzten Komponente, zum *Artikulator* darstellt.

Die grammatische Kodierung geschieht durch Zugriff auf das mentale Lexikon und durch Prozesse zur Erzeugung syntaktischer Strukturen. Es wird davon ausgegangen, daß die Information, die im mentalen Lexikon mit einer lexikalischen Einheit assoziiert ist, aus einem phonologischen, morphologischen, syntaktischen und einem semantischen Teil besteht. Die semantische Information zu einer lexikalischen Einheit assoziiert das Wort mit dem Konzept¹¹, das seine Bedeutung ausdrückt. Die syntaktische Information enthält Kategorien (Artikel, Nomen, Verb, etc.) und Subkategorisierungen (das Verb *geben* nimmt einen Agens, ein Objekt und einen Rezipienten). Ein lexikalischer Eintrag wird aktiviert, wenn seine Bedeutung (d.h. das mit dem Eintrag assoziierte Konzept) mit einem Teil der konzeptuellen Bedeutungsstruktur der Eingabe übereinstimmt. Dadurch kann auch sofort auf die syntaktische Information des Eintrags zugegriffen werden, die ihrerseits Prozesse zur Erzeugung syntaktischer Strukturen anstößt. Syntaktische Prozesse betreffen den Aufbau von Nominalphrasen, Präpositionalphrasen, Verbalphrasen, etc. Das Ergebnis der grammatischen Kodierung ist eine Oberflächenstruktur.

Aufgabe der phonologischen Kodierung ist es nun, für jedes Wort der Oberflächenstruktur und für die Struktur als Ganzes einen *Artikulationsplan* zu erstellen. Dies geschieht zum einen durch Zugriff auf die morphologische und phonologische Information, die ebenfalls im mentalen Lexikon mit den lexikalischen Einheiten assoziiert ist und zum anderen durch phonologische Prozesse, die Informationen aus dem Lexikon modifizieren und näher spezifizieren (z.B. Bestimmung der Betonung in einem Satz). Das Ergebnis dieser Teilkomponente ist die Eingabe für die nächste Komponente: der *Artikulationsplan*.

Die tatsächliche Artikulation einer Äußerung ist die Durchführung des *Artikulationsplans* durch die Muskulatur unserer Atmung und der Sprechwerkzeuge¹².

Eine wesentliche Eigenschaft eines Sprechers (weiter oben bereits kurz erwähnt) ist seine Fähigkeit zur Selbstbeobachtung (engl. *self-monitoring*). Ein Sprecher ist auch immer sein eigener Zuhörer und zwar auf zwei Ebenen: Er kann sich selbst zuhören, so wie er auch anderen Sprechern zuhört. Er kann aber auch seiner *internen Sprache* zuhören, d.h. bevor

¹⁰Der *Artikulationsplan* wird auch als *interne Sprache* (engl. *internal speech*) bezeichnet.

¹¹Wie üblich verwenden wir das Wort *Konzept* im Sinne von *Begriff* (engl. *concept*).

¹²In den weiteren Ausführungen gehen wir auf die phonologische Kodierung und den Artikulator nicht weiter ein, da sich die meisten heutigen Systeme auf die Ausgabe geschriebener Sprache beschränken.

sie tatsächlich geäußert wird. In beiden Fällen nimmt Levelt an, daß die Sprachverstehenskomponenten dafür eingesetzt werden, die eigenen Äußerungen wieder zu verstehen. Somit kann ein Sprecher Probleme bei Elementen seiner internen Sprache entdecken, bevor die Elemente tatsächlich geäußert werden. Durch Selbstbeobachtung ist der Sprecher also in der Lage, das bereits Gesagte und das intern für eine Äußerung Vorbereitete mit dem Intendierten zu vergleichen und zu korrigieren. Aus weiteren Experimenten und Daten zur Selbstkorrektur kann man schließen, daß Selbstbeobachtung außer der Überprüfung der intendierten Bedeutung auch linguistische Wohlgeformtheit einschließt, z.B. (1) (aus [Fay, 1980]):

(1) How long does that has to - have to simmer?

Weiterhin ist der Sprecher ebenfalls in der Lage, den Aufbau der vorsprachlichen Bedeutungsstruktur (die Eingabe an den *Formulator*) selbst zu beobachten, bevor sie an den Formulator gegeben wird. Damit liegt die Hauptaufgabe beim *Konzeptualisierer*, der sowohl bereits erzeugte Bedeutungsstrukturen als auch interne und geäußerte Sprache beim Aufbau neuer Bedeutungsstrukturen berücksichtigt.

2.1 Inkrementelle Sprachproduktion

In diesem Teilabschnitt gehen wir der Frage nach, ob mit dem bisher vorgestellten Sprechermodell zwangsläufig auch eine sequentielle Abfolge der einzelnen Prozesse gegeben ist.

Gesprochene Sprache wird mit einer Rate von zwei bis drei Wörtern pro Sekunde produziert. Jedes Wort wird aus mehreren 10 000 Wörtern des mentalen Lexikons mit dieser Geschwindigkeit ausgewählt. Artikulation findet mit einer Geschwindigkeit von ca. fünfzehn Phonemen pro Sekunde statt. Wenn wir davon ausgehen, daß alle Schritte sequentiell durchgeführt werden, bedeutet dies, daß die langsamste Komponente die Geschwindigkeit der Verarbeitung bestimmt, in unserem Fall der *Konzeptualisierer*. Erst wenn eine Bedeutungsstruktur aufgebaut ist, fängt der Formulator an zu arbeiten und danach der Artikulator. Das würde zu bedeutenden Pausen zwischen einzelnen Äußerungen führen, die jedoch nicht beobachtet werden. Weiterhin können Pausen *während* der Verarbeitung eines Satzes psychologisch nicht plausibel erklärt werden, da ja angenommen werden muß, daß die gesamte Information im Artikulationsplan kodiert ist.

Durch das Sprechermodell ist jedoch nicht *per se* eine serielle Verarbeitung vorgegeben. Obwohl der Formulator ohne *Konzeptualisierer* keine Eingabe zur Verarbeitung hätte und der Artikulator ohne Artikulationsplan nicht arbeiten kann, können die Komponenten doch teilweise parallel arbeiten. [Garret, 1976] schlägt vor, daß einzelne Komponenten wie eine Kaskade arbeiten können. Für den vorliegenden Fall bedeutet das: der Formulator beginnt mit der Verarbeitung der präverbalen Bedeutungsstruktur bevor diese vollständig erstellt ist und gibt sein Teilergebnis an den Artikulator weiter. Auf diese Weise können alle drei Komponenten parallel arbeiten, jede jedoch an verschiedenen Teilen der Äußerung. [Kempen, 1978] nennt diese Art der Verarbeitung *inkrementelle Generierung* (siehe auch Abb. 2).

Trotz der parallelen Verarbeitung durchläuft jedes Element einer Äußerung die unter-

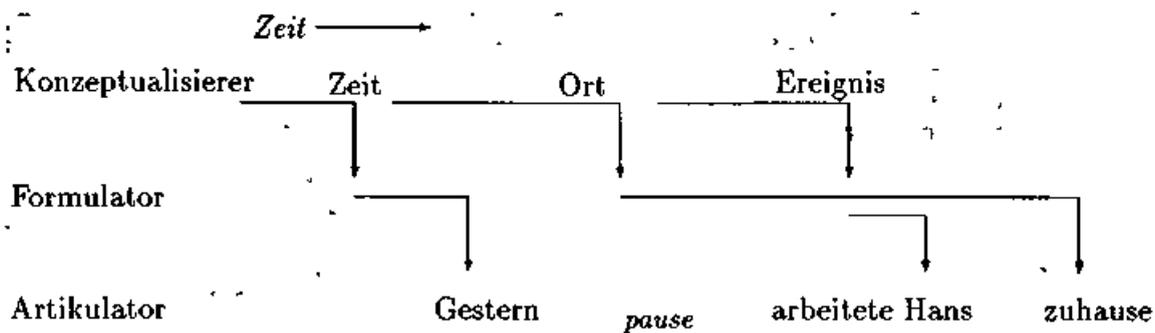


Abbildung 2: Inkrementelle Generierung.

schiedlichen Komponenten sequentiell. Abb. 2 verdeutlicht auch, daß der Formulator in gewissen Fällen auf Grund syntaktischer Beschränkungen die Reihenfolge der vom Konzeptualisierer bereitgestellten Einheiten vertauschen muß. Das Schema der inkrementellen Generierung erlaubt in einfacher Weise die Erklärung von Hesitationsphänomenen sowohl innerhalb eines Satzes als auch zwischen Sätzen. Ebenso können Abbrüche und Korrekturen erklärt werden, die oft dadurch entstehen, daß ein Sprecher wegen syntaktischer Beschränkungen kein neues Material in einen bereits teilweise geäußerten Satz einfügen kann. In [De Smedt, 1990] wird eine vollständig implementierte Computersimulation inkrementeller Satzgenerierung vorgestellt.

In Abb. 3 sind die Teilkomponenten des skizzierten Sprachproduktionsmodells mit ihren Aufgaben nochmals im Überblick dargestellt.

2.2 Die präverbale Bedeutungsstruktur

Aus Abb. 3 geht hervor, daß die präverbale Bedeutungsstruktur die einzige Eingabe an den Formulator ist (abgesehen vom mentalen Lexikon mit dessen Informationen die syntaktische Struktur aufgebaut wird). Daher muß diese Struktur alle Informationen enthalten, die eine Übersetzung in eine sprachliche Oberflächenstruktur erlauben.

Die präverbale Bedeutungsstruktur wird als eine propositional dargestellte semantische Repräsentation angesehen, die bestimmte Anforderungen erfüllen muß, damit sie in natürlicher Sprache ausgedrückt werden kann. Zwei Gesichtspunkte spielen bei der Beurteilung einer semantischen Repräsentation eine wesentliche Rolle: Vollständigkeit und Verarbeitung. Vollständigkeit bezieht sich darauf, ob in der Repräsentation alle semantischen Unterschiede ausgedrückt werden können, die ein Sprecher in seiner Muttersprache trifft. Verarbeitung betrifft die algorithmische Verwendung der Repräsentation. Bisher wurde kaum eine Repräsentationssprache mit einem psychologisch realistischen Algorithmus zur Generierung von Äußerungen benutzt ([Levelt, 1989, S. 78]), so daß Aussagen über Vor- und Nachteile anhand eines funktionsfähigen Computermodells nicht getroffen werden können. Ohne uns nun auf die eine oder andere Repräsentation festzulegen, wollen wir hier allgemeine Anforderungen an die präverbale Bedeutungsstruktur aus der Sicht der

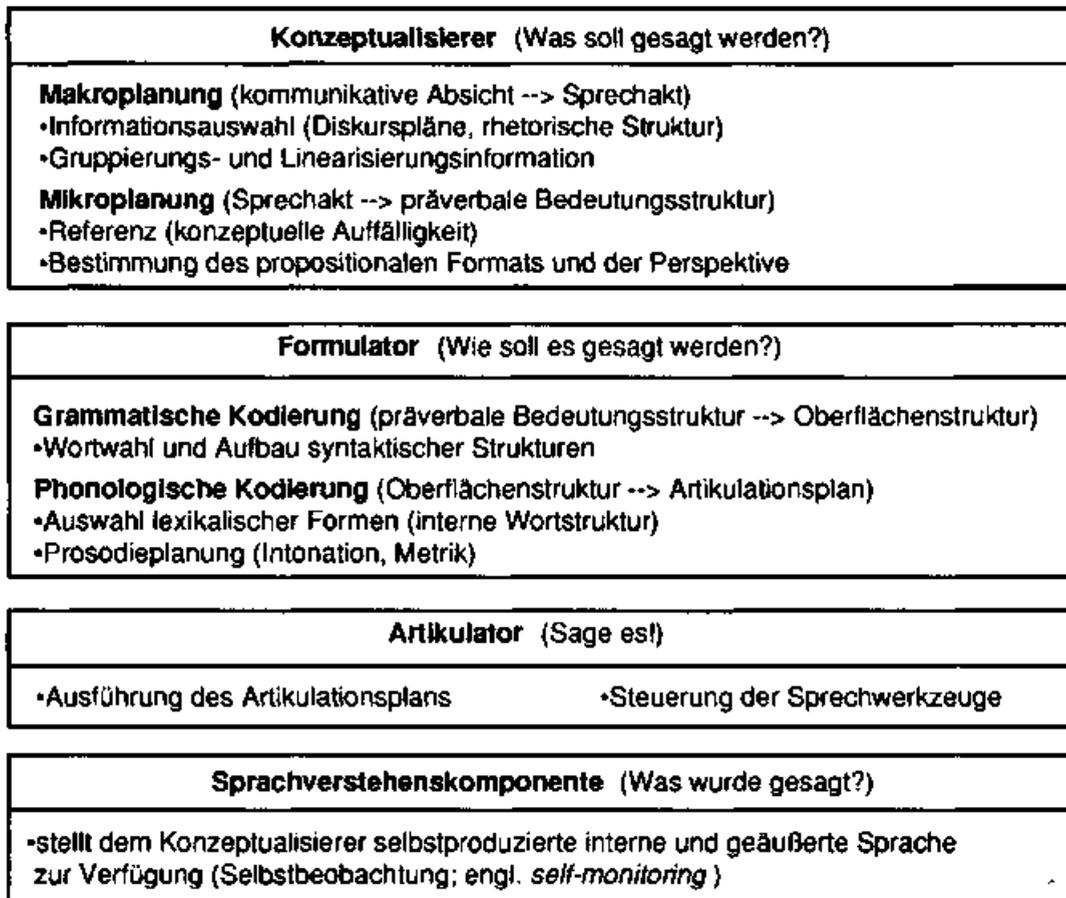


Abbildung 3: Die Prozesse im Überblick

Psycholinguistik vorstellen.

Nach Levelt kann eine präverbale Bedeutungsstruktur eine einfache Kategorie (PERSON, EREIGNIS) sein, eine Funktor/Argument-Struktur und eine Head/Modifikator-Struktur. Damit ist jedoch eine Äußerung für den Formulator noch nicht vollständig spezifiziert. Dazu müssen zusätzlich noch die thematische Funktion der Argumente festgelegt werden (*thematische Rollen*), und die Perspektive des Sprechers (d.h. Topic und Fokus seiner Äußerung) muß markiert werden.

Diejenigen Komponenten des Formulators, die für die Wortstellung und Flektionsmorphologie sorgen, benötigen entsprechende Markierungen in der präverbalen Bedeutungsstruktur für Modus (deklarativ, interrogativ, imperativ), Tempus und Deixis. Deixis betrifft Ausdrücke wie *hier, dort, heute*, etc. die unmittelbar vom Sprecher und seiner Sprechsituation abhängen. Um diese Ausdrücke korrekt verwenden zu können, muß die Sprechsituation ebenfalls repräsentiert sein.

Allgemein kann man sagen, daß die Kodierung einer präverbalen Bedeutungsstruktur für Sprecher verschiedener Sprachen unterschiedlich ist, da Sprachen sich in ihren grammatisch relevanten semantischen Merkmalen unterscheiden. Wir haben insbesondere gesehen, daß nicht nur kommunikationsrelevante Einheiten in der präverbalen Bedeutungsstruktur

kodiert werden, sondern alle Merkmale, die zur Erzeugung einer Äußerung notwendig sind.

2.3 Die Konstruktion der präverbalen Bedeutungsstruktur

Aus Abb. 3 geht hervor, daß die präverbale Bedeutungsstruktur in zwei Schritten erzeugt wird. Im ersten Schritt, der Makroplanung, wird die kommunikative Absicht des Sprechers in einen Sprechakt überführt (z.B. *informiere, frage, befehle*). Eine kommunikative Absicht eines Vaters V seiner Tochter T gegenüber kann z.B. (2) sein.

(2) (KNOW (T, INTEND (V, BELIEVE (T, HAUPTSTADT-VON(FRANKREICH, PARIS))))))

(3) (INFORMIERE (T, HAUPTSTADT-VON(FRANKREICH, PARIS)))

(4) (DEKLARATIV (HAUPTSTADT-VON (FRANKREICH, PARIS)))

Dies bedeutet: die Absicht besteht darin, daß die Tochter nach der Äußerung weiß, daß der Vater beabsichtigt, daß sie glaubt, daß Paris die Hauptstadt von Frankreich ist. Der Sprechakt, der diesen Zustand bei der Tochter hervorruft, kann z.B. (3) sein. Dieser Sprechakt muß nun wiederum in eine präverbale Bedeutungsstruktur übersetzt werden, die wie (4) aussehen kann.

Im ersten Schritt müssen also mithilfe von Regeln kommunikative Absichten in Sprechakte übersetzt werden. Dabei kann oft mehr als ein Schritt notwendig sein. Bei der Erzeugung einer Wegbeschreibung z.B. weiß der Sprecher oft den Anfangs- und den Zielort. Um den Weg zu beschreiben, wird er das Gesamtziel in mehrere Unterziele aufspalten: Beschreibe zuerst wie sie zur Autobahn kommt, dann wo sie runterfahren muß, danach wie sie zur Stadtmitte kommt, etc. Für jedes dieser Unterziele müssen dann entsprechende Sprechakte und darauf folgend präverbale Bedeutungsstrukturen generiert werden. Das Ergebnis der Makroplanung ist eine Sequenz von Sprechakten, bei denen der zu äuernde Inhalt und der Modus (deklarativ, interrogativ, imperativ) festgelegt sind.

Unabhängig davon, ob ein Dialog geführt oder ein längerer Text generiert wird, ob Sprecher und Hörer kooperieren oder entgegengesetzte Ziele verfolgen, will ein Sprecher stets einen *Effekt* durch seine Äußerung beim Hörer erreichen. Der Effekt kann z.B. darin bestehen, daß der Hörer sein Wissen erweitert (siehe obiges Beispiel), daß er von einem Argument überzeugt wird oder daß er eine Handlung ausführt. In jedem Fall muß der Sprecher den Hörer mit seinem Wissensstand, seinen Überzeugungen und seinen Zielen einbeziehen, kurz, er muß seine Äußerung *planen*. Weiterhin muß der Sprecher sorgfältig abwägen, was explizit gesagt werden soll und was implizit vermittelt wird (also vom Hörer erschlossen werden kann).

Die Auswahl der sprachlichen Mittel zum Erreichen des Ziels¹³ wird in der KI oft als Planungsproblem behandelt. Wissensstand, Überzeugungen und Ziele des Hörers sowie der sprachliche und außersprachliche *Kontext* beeinflussen die Planung. Der sprachliche Kontext erlaubt (oder verbietet) z.B. anaphorische Referenzierungen. Der Einfluß des

¹³Ein kommunikatives Ziel kann nicht nur durch sprachliche, sondern auch durch außersprachliche Mittel (Gestik, Mimik, Handlungen) erreicht werden. Wir beschränken uns hier auf sprachliche Mittel, was allerdings eine wesentliche Einschränkung bei der Modellierung realer Kommunikation bedeutet.

außersprachlichen Kontexts soll an folgendem Beispiel erläutert werden. Die Frage (5) kann auf viele verschiedenen Weisen interpretiert werden. Sie kann eine indirekte Aufforderung sein, die Zeit zu sagen (6). Die direkte Antwort *ja* wäre dann nicht kommunikativ adäquat. Die Frage kann einen Vorwurf beinhalten, auf den der Sprecher z.B. mit (7) reagiert. Hier würde die Antwort *ja* die Konfrontation wahrscheinlich verstärken. Schließlich kann die Frage auch direkt interpretiert werden, wenn z.B. A sichergehen will, daß B die Voraussetzung hat, eine Verabredung zu einem späteren Zeitpunkt einzuhalten. Dann ist eine Antwort wie (8) angemessen.

- (5) A: Wissen Sie, wie spät es ist?
- (6) B: Es ist viertel vor drei.
- (7) B: Tut mir leid, ich bin aufgehalten worden.
- (8) B: Ja, ich habe eine Uhr dabei.

Die Planerkennung und -generierung beruht auf gemeinsamen Überzeugungen (engl. *shared beliefs*). Wenn z.B. B die Frage (5) von A in der ersten Version interpretiert hat, so glaubt B, daß A keine Uhr dabei hat (sonst könnte A sich die Frage nach der Zeit leicht selbst beantworten). Außerdem glaubt B, daß A weiß, daß A keine Uhr dabei hat (sonst hätte A nicht gefragt). Schließlich glaubt B auch, daß A glaubt, daß B glaubt, daß A keine Uhr dabei hat. Daher würde die Antwort (9) gegen beider Überzeugungen verstoßen.

- (9) B: Wenn Sie die Zeit wissen wollen, so schauen Sie doch auf Ihre Uhr.

Nachdem der Sprecher die einzelnen Sprechakte bestimmt hat, muß er nun für jeden einzelnen Sprechakt die genaue Informationsstruktur festlegen: Perspektive, Topic und Fokusstruktur sowie alle weiteren Angaben, die in der präverbalen Bedeutungsstruktur notwendig sind.

Beide Planungsebenen sind extrem kontextabhängig. Dem eigenen Chef gegenüber kann z.B. eine Handlungsaufforderung als informiere-Sprechakt überbracht werden, wohingegen das bei einem Untergebenen als *befehle* möglich ist. Welche Perspektive einer Äußerung gegeben wird, hängt auch von dem Glaubens- und Wissenszustand des Hörers ab. Wenn man weiß, daß der Hörer gerade das Gegenteil dessen glaubt, was man ihm sagen möchte, wird nicht nur die Tatsache als solches genannt, sondern auch versucht, weitere Begründungen dafür anzuführen.¹⁴

Die Notwendigkeit der sprecherinternen Buchführung über den Verlauf des bisherigen Diskurses, die *Diskurshistorie*, und des sprecherinternen mentalen Modells des Diskurses von dem der Sprecher glaubt, daß es von der Hörerin geteilt wird, des *Diskursmodells*, ist in der Psycholinguistik unbestritten. Wir gehen hier nicht weiter auf die Problematik von Diskurshistorie und Diskursmodell ein, sondern stellen einige weitere Einflußgrößen bei der Planung von Äußerungen vor, denen gemeinsam ist, daß sich der Sprecher ihrer bewußt sein muß.

¹⁴Einen guten Überblick über Forschungen auf dem Gebiet der Formalisierung von Intentionen und Sprechakten sowie der Modellierung von Glaubens- und Wissenszuständen gibt [Cohen *et al.*, 1990].

Diskurstyp und Diskursthema

Als erstes ist der *Diskurstyp* zu nennen, besonders *informelle alltägliche Konversation*. Diese zeichnet sich dadurch aus, daß sich der Sprecher des informellen Charakters bewußt ist, daß die Gesprächspartner dieselben Rechte haben und die Freiheit, das Thema jederzeit zu ändern. Weiterhin herrscht weitgehend Einigkeit darüber, nach welchen Regeln die Dialogführung (engl. *turn-taking*) wechseln kann.

Verschiedene Diskurstypen sind bisher untersucht worden: Erzählungen, Vorlesungen, Wegbeschreibungen.¹⁵ Den unterschiedlichen Diskurstypen gemeinsam ist, daß der Sprecher sich des Typs bewußt sein muß, um die entsprechende Rolle einnehmen zu können. Kategorienfehler auf der Sprecherseite können folgeschwer sein, wenn man z.B. an Funkgespräche der Flugsicherung denkt. Um welchen Diskurstyp es sich aktuell handelt, wird entweder explizit verhandelt, ist durch die Situation vorgegeben oder wird implizit durch die Art und Weise eines Sprechers, ein Gespräch zu beginnen, vorgegeben [Schegloff, 1987].

Ein weiteres Merkmal eines Diskurses, dessen sich ein Sprecher stets bewußt sein muß, worüber er also buchführen muß, ist das *Thema* des Diskurses, das, *worüber* gesprochen wird. Weicht er vom Thema ab, so trägt er nicht zur Kohärenz des Diskurses bei. Es gibt jedoch sprachliche Möglichkeiten, eine gewünschte Themaänderung (engl. *topic shift*) explizit zu signalisieren. Eine Unterbrechung kann z.H. durch *Das nächste Thema, daß ich diskutieren möchte, ist ...* direkt deklariert werden. Weitere Möglichkeiten sind *Dabei fällt mir ein, . . .*, um eine Abweichung vom Thema anzukündigen, nach der meistens das ursprüngliche Thema wiederaufgenommen wird, oder *Halt, ich habe vergessen zu sagen, daß ...*, um ein früheres Thema nochmals aufzunehmen. Viele Autoren gehen von einer Hierarchie von Themen aus, die sie mit Diskurszielen identifizieren. Wird ein Unterziel angesprochen, so wird in der Hierarchie eine Ebene nach unten gegangen. Ist das Unterziel erfüllt, kann auf eine höhere Ebene zurückgegangen werden.

Informationsauswahl: empirische Befunde

Was der Sprecher sagt, hängt völlig von seinen Intentionen und dem gegenwärtigen Status seines Diskursmodells ab. Für komplexe Beschreibungen ist die Zeit zur Auswahl geeigneter Information höher als für einfache, bekannte Aufgaben. Dies wurde in einem Experiment von [Good und Butterworth, 1980] zur Wegbeschreibung gezeigt. Die Probanden sollten einen ihnen bekannten Weg (von Zuhause zur Arbeitsstelle) und einen ihnen unbekannt Weg beschreiben. Die Beschreibung des bekannten Weges enthielt weniger Pausen (Zeiten, in denen nichts geäußert wurde) als die Beschreibung des unbekannt Weges, 33 Prozent bzw. 37 Prozent. Bei der wiederholten Beschreibung des bekannten Weges fiel der Prozentsatz auf 27 Prozent. Die Annahme ist, daß die Auswahl der Information im Wiederholungsfall einfacher ist, da die Information im Langzeitgedächtnis präsent und leicht zugreifbar ist. Es wird vermutet, daß Sprecher flüssige Äußerungen anstreben und versuchen, Pausen in Sätzen zu minimieren. Im Experiment von [Good und Butterworth, 1980] wurde gezeigt, daß Pausen in Sätzen (von mehr als 250 msec Länge) im Wiederholungsfall der Beschreibung wesentlich weniger häufig auftraten als in den Originalbeschreibungen,

¹⁵Einen guten Überblick über verschiedene Diskurstypen gibt [Van Dijk, 1985].

28 Prozent bzw. 41 Prozent der Gesamtanzahl der Pausen. Daraus kann man schließen, daß bei einer Aufgabe, die keinen hohen Aufwand für die Informationsauswahl erfordert, mehr Aufmerksamkeit für die flüssige Äußerung eines Satzes aufgewendet wird, d.h. für die Mikroplanung.

Als nächstes gehen wir auf die Frage ein: *Nach welchen Kriterien wird Information ausgewählt, um sprachlich auf Objekte zu referieren?* Von [Olson, 1970] stammt die Beobachtung, daß Referenz auf Objekte von den wahrgenommenen und inferierbaren Alternativen abhängt. Liegt von vielen Notizzetteln auf einem Schreibtisch nur einer auf einem Buch und es liegt nur ein Buch auf dem Tisch, kann man *der Zettel auf dem Buch* sagen. Ist ein weißes und ein rotes Buch auf dem Tisch und der Zettel liegt auf dem weißen, sagt man: *der Zettel auf dem weißen Buch*. Das Prinzip der Erzeugung referierender Ausdrücke liegt in der eindeutigen Kennzeichnung der Objekte, so daß sie vom Hörer ebenfalls eindeutig identifiziert werden können.

Generell kann man sagen, daß referierende Ausdrücke immer eindeutig sind, aber oft redundant. [Pechmann, 1984] fand in einem Experiment, daß mehr als 60 Prozent der referierenden Ausdrücke Erwachsener redundant waren. Es gibt wenigstens zwei gute Gründe für die Überspezifizierung. [Deutsch, 1976] und [Mangold, 1986] zeigen z.B., daß es für einen Hörer einfacher ist, einen überspezifizierteren Referenten zu finden als einen minimal spezifizierten. Es ist leichter ein *weißes Buch* zu identifizieren als *etwas Weißes*. Die Vermutung ist, daß durch die Verbalisierung der Art der Objekte dem Hörer die Bildung einer *Gestalt* möglich ist, nach der er dann sucht. Der zweite Grund liegt in den inferierbaren Alternativen. Ein Sprecher kann ein neues Objekt nicht nur gegenüber visuell wahrgenommene Alternativen abgrenzen, sondern auch zu Alternativen in seinem Diskursmodell kontrastieren.

In dem Experiment von [Pechmann, 1984] wurden dem Sprecher z.B. zwei aufeinanderfolgende Dias gezeigt. Auf dem ersten war ein blauer und ein roter Vogel sowie eine blaue und eine rote Tasse zu sehen. An einem der vier Objekte war ein Sternchen (dem blauen Vogel). Dieses sollte dem Hörer, der dasselbe Dia nur ohne das Sternchen sah, beschrieben werden. Beim ersten Dia sagte der Sprecher, *der blaue Vogel*, eine eindeutige, minimale Kennzeichnung. Auf dem folgenden Dia waren eine blaue und eine rote Tasse, sowie ein roter Vogel zu sehen. Der rote Vogel sollte beschrieben werden und wurde als *der rote Vogel* verbalisiert, was redundant ist, zumindest bezogen auf die visuell präsenten Alternativen. Bezogen auf das Diskursmodell ist vor dieser Äußerung sicherlich *der blaue Vogel* im Fokus, und die Äußerung *der rote Vogel* ist im Kontrast dazu, obwohl kein anderer Vogel visuell präsent ist. Perzeptuell ist also *rot* die redundante Information, bezogen auf das Diskursmodell jedoch *Vogel*. Pechmann nennt das *exophorische Redundanz* (perzeptuell gegeben) und *endophorische Redundanz* (vom Diskursmodell stammend).

Die Vermutung ist, daß endophorische Referenz besonders kooperativ ist, da sie dem Hörer gestattet, mit minimalem Aufwand eine neue Gestalt zu bilden, wenn der Unterschied zwischen den beiden Objekten nur gering ist. Ist der Unterschied zu groß, wird exophorisch referiert.

Weitere Experimente zur Auswahl und Formulierung der zu übermittelnden Information wurden u.a. von [Herrmann, 1983] und [Brown und Dell, 1987] durchgeführt. Herr-

mann fand, daß bei der Äußerung einer Anfrage oder Forderung (engl. *request*) die vom Sprecher vermutete Legitimität seiner Anfrage eine wesentliche Rolle spielt. In einem Experiment forderte der Sprecher von seinem Partner eine Pistole, die dem Partner gehörte und formulierte die Forderung folgendermaßen: *Könntest Du mir die Pistole geben, Würdest Du mir die Pistole geben, Ich würde gerne die Pistole benutzen*. War jedoch der Sprecher Besitzer der Pistole, wurde die Forderung anders verbalisiert: *Gib mir die Pistole, Du mußt mir die Pistole geben, Ich brauche die Pistole*. Generell nahm die Höflichkeit der Anfrage bei zunehmenden Rechten des Sprechers ab.

Brown und Dell baten Probanden Geschichten nachzuerzählen [Brown und Dell, 1987]. In jeder Geschichte kam eine Handlung vor, die mit einem Instrument ausgeführt wurde (z.B. erstechen, Instrument: Messer oder Eispickel). Wenn das Instrument das typische für die Handlung war (Messer), so wurde es von den meisten Probanden nicht explizit erwähnt. War das Instrument jedoch der Eispickel, war die Häufigkeit der Erwähnung größer. Insbesondere wurde das untypische Instrument in demselben Satz kodiert wie die Handlung (*Der Dieb erstach den Mann mit einem Eispickel*) wohingegen das typische Instrument in einem separaten Satz kodiert wurde (*Der Dieb ergriff ein Messer und erstach den Mann*).

Linearisierung

Eine weitere Aufgabe, die Linearisierung der ausgewählten Information, ist ebenfalls von der Makroplanung zu leisten. Wir gehen hier nicht auf alle Experimente ein, sondern stellen einige wichtige Prinzipien vor, die sich aus den Experimenten ergeben und die teilweise auch auf die menschliche Gedächtnisstruktur zurückzuführen sind. Zwei Faktoren beeinflussen die Linearisierung: inhaltliche Kriterien und prozeßbezogene Kriterien.

Inhaltliche Kriterien betreffen die natürliche Ordnung der zu übermittelnden Information. Beim Geschichtenerzählen kann der Sprecher dem Verlauf bzw. der Entwicklung der Geschichte folgen. Bei der Verbalisierung von Bildfolgen [Novak, 1987] ist die zeitliche Abfolge der einzelnen Ereignisse eine natürliche Ordnung. Bei der Aufgabe, eine Wegbeschreibung zu geben, folgen Sprecher in einem mentalen Spaziergang der Route vom Start zum Zielpunkt.

Prozeßbezogene Kriterien werden dann am auffälligsten, wenn keine natürliche Ordnung vorhanden ist. [Levelt, 1982] führte verschiedene Linearisierungsexperimente durch und entwarf ein Modell der Linearisierung. Wir gehen hier nur auf die wesentlichen Prinzipien ein. Levelt zeigte den Probanden ein Netzwerk, bestehend aus runden, farbigen Knoten, durch vertikale oder horizontale Linien mit anderen Knoten verbunden. Nun gab er einen Startpunkt vor, von dem aus die Beschreibung begonnen werden sollte. Die Prinzipien, nach denen diese Linearisierungsaufgaben durchgeführt werden, lauten:

Prinzip der Konnektivität: Wenn möglich, wähle als nächsten Knoten denjenigen aus, der eine direkte Verbindung zum gegenwärtigen Knoten hat.

Stapelprinzip: Nachdem Du in eine Verzweigung gegangen bist, kehre zum letzten Entscheidungspunkt zurück.

Prinzip des minimalen Aufwandes: Ordne alternative Fortsetzungen so an, daß der Aufwand für die Rückkehr zum letzten Entscheidungspunkt minimal ist.

Nach der Linearisierung der zu übermittelnden Information muß ihr noch eine entsprechende Informationsstruktur gegeben werden. Dies geschieht in der schon erwähnten Mikroplanung. Im folgenden gehen wir auf das mentale Lexikon und den Zugriff darauf ein.

Lexikalische Einträge

Bei der Vorstellung des Formulators (siehe oben) wurde bereits darauf eingegangen, daß die Übersetzung der präverbalen Bedeutungsstruktur in die Oberflächenstruktur durch Zugriff auf die Lexikoneinträge erfolgt. Jeder Lexikoneintrag enthält phonologische, morphologische, syntaktische und semantische Informationen, die man sich als Mengen von Merkmal-Wert Paaren vorstellen kann. Der Aufbau syntaktischer Strukturen wird durch diejenigen Elemente der präverbalen Bedeutungsstruktur ausgelöst, die mit der semantischen Information eines Lexikoneintrags übereinstimmen. Es wird also angenommen, daß das Lexikon zwischen konzeptueller und grammatischer und phonologischer Kodierung vermittelt. Diese Annahme wird auch als *lexikalistische Hypothese* bezeichnet. Sie bedeutet insbesondere, daß nichts in der präverbalen Bedeutungsstruktur *per se* syntaktische Strukturen erzeugt (wie z.B. aktiv oder passiv, Dativobjekt, etc.), sondern daß diese Entscheidungen erst durch einen Lexikonzugriff getroffen werden können.

Wie kann man sich die interne Struktur eines Lexikoneintrags vorstellen? Aufgrund verschiedener Sprachproduktionsphänomene (Sprechfehler, „es liegt mir auf der Zunge“-Phänomene; für einen Überblick siehe [Fromkin, 1980]) geht die Psycholinguistik von zwei Organisationsformen des Lexikons aus: einer nach der Bedeutung und einer nach der Form (vgl. Abb. 4, nach [Levelt, 1989, S. 188]). Die semantische und syntaktische Information,

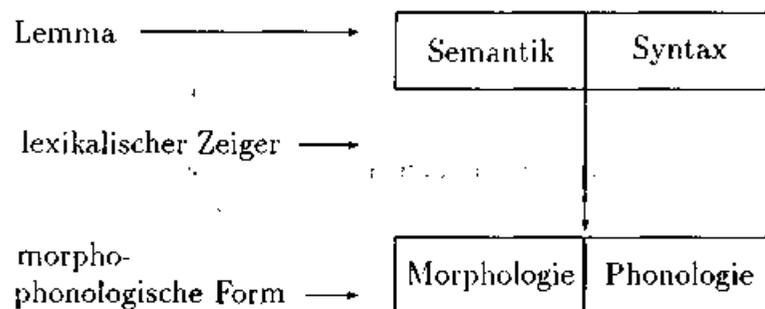


Abbildung 4: Ein lexikalischer Eintrag

die mit einem lexikalischen Eintrag assoziiert ist, wird als *Lemma* bezeichnet, die phonologische und morphologische Information als *morphophonologische Form*. Die Frage, ob für den Zugriff auf die Forminformation zuerst die Lemmainformation gefunden werden muß, ist gegenwärtig ungeklärt.

Unterschiedliche Flexionen eines Wortes werden nicht als verschiedene Einträge angesehen, sondern als ein Eintrag. Die korrekte Flexion wird durch Merkmalswerte gesteuert, die in der präverbalen Bedeutungsstruktur angegeben sind. Zu den vier bereits erwähnten Merkmalsmengen zur Beschreibung einer lexikalischen Einheit kommen sicherlich noch weitere hinzu, wie z.B. pragmatische, stilistische und konnotative Merkmale, die die Verwendung von Wörtern mit ähnlicher Bedeutung in unterschiedlichen Kontexten steuern. Weiterhin wird angenommen, daß auch ganze Phrasen, z.B. idiomatische Ausdrücke, im Lexikon abgelegt sind.

Der Zugriff auf das Lexikon und damit verbundene Probleme werden im Abschnitt über Wortwahl diskutiert.

3 Planungs- und Entscheidungsprozesse

3.1 Verschiedene Architekturen

Drei verschiedene Architekturen für Generierungssysteme können unterschieden werden: *sequentiell*, *interagierend*, *integriert*.

Sequentiell bedeutet hier, daß der Datenfluß nur in einer Richtung besteht, vom Inhaltsbestimmenden zum formbestimmenden Teil. Viele Systeme folgen diesem sequentiellen Modell, z.B. MUMBLE [McDonald, 1983], TEXT [McKcown, 1985], NAOS [Novak, 1987] und SEMTEX/GEOTEX [Rösner, 1987]. Zuerst wird der Inhalt vollständig festgelegt und danach in eine sprachliche Oberfläche überführt. Inhalt bedeutet hier soviel wie die im letzten Abschnitt eingeführte präverbale Bedeutungsstruktur, also Festlegung der konzeptuellen Struktur und der Informationsperspektive (Thema, Fokus, Topikalisierung). Im Unterschied zur präverbalen Bedeutungsstruktur wird jedoch davon ausgegangen, daß Wortwahl bereits stattgefunden hat.

Bei *interagierenden* Architekturen findet an definierten Schnittstellen eine bidirektionale Kommunikation statt. PAULINE [Hovy, 1988b] und POPEL [Reithinger, 1992] sind Beispiele für Systeme mit Rückwirkungen. Bei PAULINE erfolgen Interaktionen an fünf Entscheidungspunkten: Themawahl, Satzinhalt, Satzorganisation, Satzteilorganisation und Wortwahl. Der inhaltsbestimmende Teil plant ein Teil der Äußerung und gibt ihn an den formbestimmenden Teil weiter. Trifft er auf einen der Entscheidungspunkte, geht die Kontrolle zurück an die Inhaltsplanung.

Bei *integrierten* Architekturen wird die Trennung in Inhalts- und Formbestimmung fallengelassen. Entscheidungen, was und wie etwas zu äußern ist, werden als Teile eines einzigen, hierarchischen Planungsprozesses angesehen, wobei dieser Prozeß gleichermaßen Zugriff auf kontextuelles und sprachliches Wissen hat. KAMP [Appelt, 1985] ist ein Beispiel eines integrierten Systems.

Ein Nachteil des integrierten Ansatzes ist, daß eine explizite Trennung von linguistischem und kontextuellem Wissen nicht mehr möglich ist. Es gab z.B. bei KAMP (in der ersten Version) keine explizite Grammatik. Weiterhin muß ein einziger Planungsprozeß mit sehr verschiedenartigem Wissen umgehen.

Ein Nachteil des sequentiellen Ansatzes ist, daß stillschweigend davon ausgegangen wird, daß jede Bedeutungsstruktur auch in der angestrebten Form und Sprache verbalisiert werden kann. Die Bestimmung des Inhalts erfolgt ohne Berücksichtigung der Sprachstruktur. Falls aus syntaktischen Gründen eine Umsetzung des Inhalts in einen Text eine Umstellung des Inhalts nötig macht, kann das nicht modelliert werden. Z.B. restringieren in einem bereits teilweise geäußerten Satz die noch möglichen syntaktischen Leerstellen die weitere Reihenfolge der Argumente.

Die interagierende Architektur scheint keine Nachteile aufzuweisen. Das Problem besteht hier in der Festlegung der Schnittstellen, an denen die Interaktion stattfinden soll.¹⁶

An dieser Stelle ist ein erster Vergleich der obigen Vorstellungen mit dem bereits vorgestellten Sprachproduktionsmodell angebracht. Dort wurde argumentiert, daß keine direkte Rückkopplung zwischen Formulator und Konzeptualisierer vorliegt. Es wurde jedoch die Möglichkeit der Rückkopplung über die Selbstbeobachtung eingeräumt. Keine der drei vorgestellten Architekturen widerspricht der Idee der inkrementellen Generierung (mit POPEL liegt sogar ein inkrementeller Generator vor).

Eine Frage, die bisher nicht diskutiert wurde und auf die es nur tentative Antworten gibt, ist: *Unterhegt die Produktion gesprochener Sprache denselben Mechanismen wie die Produktion von Texten?* Intuitiv ist die Antwort: *Nein!* Bei der Produktion von Texten spielen syntaktisch zu realisierende Kohärenzrelationen eine größere Rolle als bei gesprochener Sprache. Textproduktion findet meistens ohne direkten Kontakt mit dem Rezipienten des Textes statt, sondern nur bezogen auf den gedachten Hörer, wohingegen bei der Produktion gesprochener Sprache immer das stillschweigende Übereinkommen vorhanden ist, daß der Hörer zurückfragen kann, wenn ihm etwas unverständlich ist. Von einigen Autoren wird für die Textproduktion auch eine Revisionskomponente angenommen [Mann und Moore, 1982; Meteor, 1988]. Damit soll modelliert werden, daß zuerst eine Bedeutungsstruktur spontan verbalisiert wird und danach überprüft wird, ob und wie sie zu ändern ist, um sie an den Vorgängertext und den gedachten Nachfolgetext optimal anzupassen.

3.2 Inhaltsbestimmung: Das Auswahlproblem

Die Frage *Wann sagt man was?* beschreibt das Problem kurz und bündig. Was sind die Anforderungen an die Äußerungen eines Systems? Eine sprachliche Äußerung soll *kommunikativ adäquat* und *kohärent* sein. Eine Äußerung ist kommunikativ adäquat, wenn sie in ihrer Funktion den Erwartungen des Hörers entspricht. Kohärent ist eine Äußerung dann, wenn sie in die „Umgebung“ des bisherigen Textes/Dialogs hineinpaßt, d.h. wenn es eine Äußerung zum Thema ist oder zumindest Thema- bzw. Fokuswechsel sprachlich markiert wird, alle Referenzen vom Hörer aufgelöst werden können, etc.

Anhand der Graphik zur Architektur eines Generierungssystems in Abb. 5 wollen wir einige Probleme zuerst allgemein und dann vertieft diskutieren. Die Ovale stehen für Datenstrukturen und die Rechtecke für Prozesse. Der Pfeil in der Mitte der Graphik stellt den Hauptdatenfluß dar, vom kommunikativen Ziel zur Äußerung. Die Pfeile an den Seiten

¹⁶[Danlos, 1987] zeigt, wie schwer es ist, konzeptuelle von linguistischen Entscheidungen zu trennen.

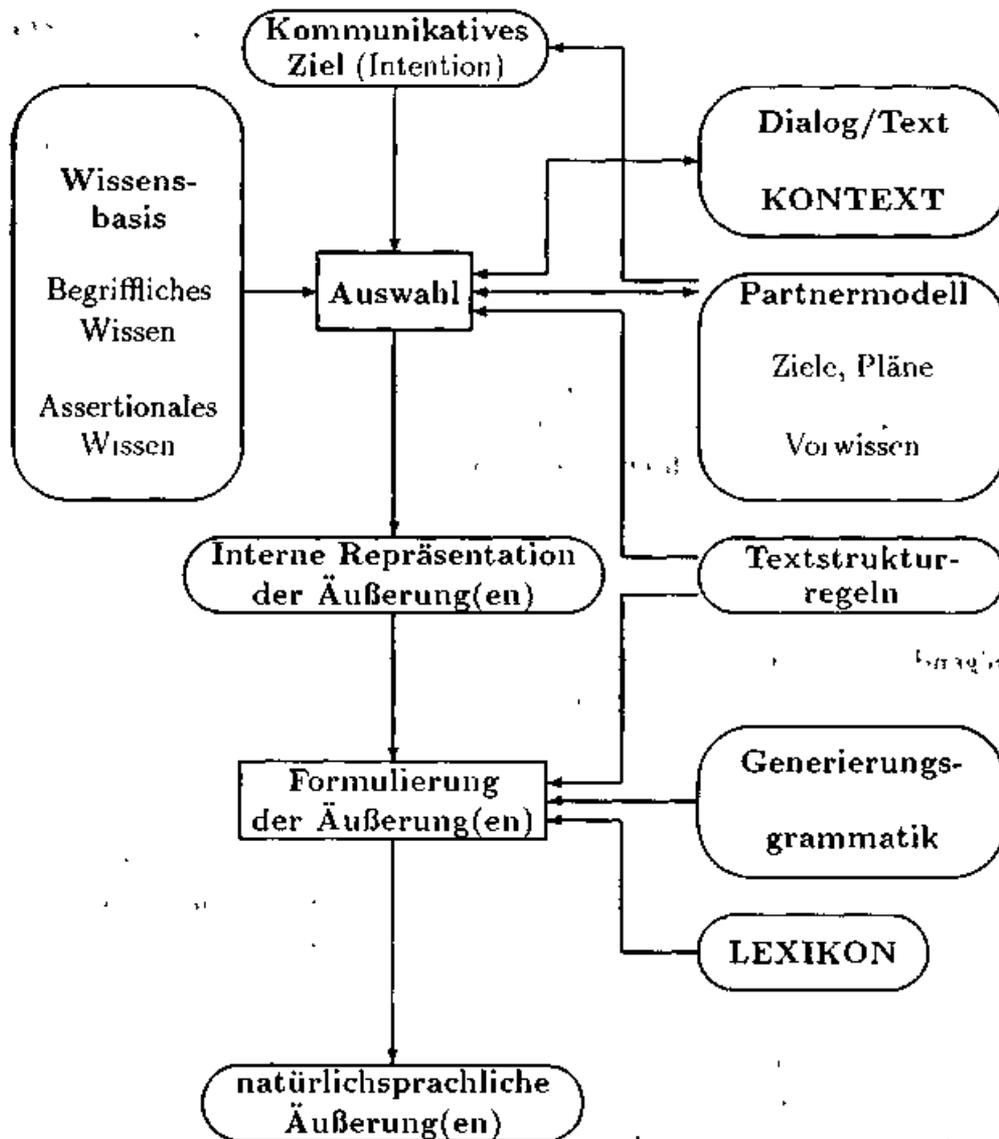


Abbildung 5: Architektur eines Generierungssystems

zeigen den jeweiligen Einfluß der Daten auf die Prozesse.

Der Auswahlprozeß hängt vor allen Dingen vom zur Verfügung stehenden Wissen des Systems ab, der Wissensbasis. In der Graphik ist diese als aus zwei Teilen bestehend dargestellt: dem begrifflichen Wissen und dem assertionalen Wissen. Für assertionales Wissen sagt man auch referentielles Wissen. Das begriffliche Wissen enthält die Definitionen und Zusammenhänge der Begriffe, die dem System bekannt sind. In heutigen Systemen wird dieses Wissen mithilfe formal wohldefinierter Repräsentationssprachen dargestellt, für die man eine eindeutige Semantik angeben kann. Das assertionale Wissen repräsentiert die tatsächlichen Objekte im Diskursbereich. Dies geschieht dadurch, daß Instanzen

von Objekten im begrifflichen Wissen gebildet werden. Der Auswahlprozeß selbst wird im wesentlichen vom Dialog/Text Kontext beeinflußt, man denke hier an die bereits diskutierten Fragen der thematischen Progression und der Fokusstrukturen. Die Abhängigkeit vom Partnermodell ist auf Grund des bisher Gesagten selbstverständlich. Der Einfluß von Textstruktureregeln hingegen ist noch nicht diskutiert worden und wird im Abschnitt *Formbestimmung* vertieft. An dieser Stelle nur soviel, wir haben bereits gezeigt, daß es verschiedene Texttypen (auch genannt Textsorten, Diskurstypen) gibt, die sich durch unterschiedlichen Textaufbau, Detaillierungsgrad, etc. auszeichnen. Bei der Erzeugung solcher Texte ist es selbstverständlich, daß Regeln zur Textstruktur den Auswahlprozeß beeinflussen.

Das Ergebnis des Auswahlprozesses ist eine Repräsentation der zu erzeugenden Äußerungen, das, was wir aus psycholinguistischer Sicht präverbale Bedeutungsstruktur nannten. Diese Struktur wird durch einen Prozeß, in der Graphik Formulierung genannt, in eine natürlichsprachliche Äußerung, bzw. mehrere Äußerungen überführt. Bei mehreren Äußerungen üben wiederum Textstruktureregeln einen maßgebenden Einfluß aus. Die Formulierung selbst erfolgt mithilfe einer expliziten Grammatik und unter Zugriff auf das Lexikon.

Im folgenden gehen wir zuerst auf Probleme und Lösungen bei Dialogäußerungen ein und betrachten dann Fragen der Textgenerierung.

Dialogäußerungen

Sowohl bei Dialogsystemen als auch bei Textgenerierungssystemen werden die Grice'schen Konversationsmaximen als Heuristik zur Behandlung des Auswahlproblems herangezogen. Sie beruhen auf einer allgemeinen Analyse von Konversationen und daraus abgeleiteten Maximen, an die sich ein Sprecher halten sollte, der eine Konversation kohärent gestalten möchte. Grice [Grice, 1975] geht vom Prinzip kooperierender Partner aus, das er folgendermaßen definiert: „*Leiste deinen Beitrag so, wie er vom gegenwärtigen Stadium des Dialoges und vom Dialogziel gefordert wird*“.

Akzeptiert man dieses Prinzip, so ergeben sich nach [Grice, 1975] speziellere Maximen, die unter die folgenden vier Kategorien fallen:

- | | |
|----------------|--|
| Quantität: | mache deinen Beitrag so informativ wie gefordert;
mache ihn nicht informativer als gefordert; |
| Qualität: | achte darauf, daß dein Beitrag wahr ist;
sage nichts, wovon du glaubst, es sei falsch;
sage nichts, wofür du keine Beweise hast; |
| Beziehung: | sei relevant; |
| Art und Weise: | fasse dich kurz;
vermeide umständliche Äußerungen;
vermeide Mehrdeutigkeiten. |

Ohne Beschränkung der Allgemeinheit betrachten wir am Beispiel eines Zugangssystems zu einem Datenbanksystem einige Auswahlprozesse, die sich aus der Anforderung,

kooperierende Antworten zu erzeugen, ergeben. Dazu unterscheiden wir zwischen *direkten* und *indirekten* Antworten auf eine Frage. Direkte Antworten enthalten genau die erfragte Information, wobei diese nicht unbedingt natürlichsprachlich ausgegeben werden muß. Fragen an ein Datenbanksystem, wie z.B. *Welche Züge fahren von Stuttgart nach Hamburg?* sollten auch nicht natürlichsprachlich beantwortet werden, wenn die Anzahl der Antworten groß ist, da in diesem Fall eine tabellenartige Ausgabe für den Benutzer viel verständlicher und übersichtlicher ist. Bei *direkten* Antworten wird in der Regel das Fragewort durch die entsprechende Information ersetzt und in angemessener Form, z.B. als Ellipse, ausgegeben. Beispiel (10) stammt aus dem Dialogsystem HAM-ANS [Hoepfner *et al.*, 1984].

- (10) Ben: Welcher PKW ist abgebogen? ' •
Sys: Der grüne.

Ein herausforderndes Problem stellen indirekte Antworten dar, die nicht die vom Benutzer erfragte Information enthalten, aber Hinweise darauf, warum die Frage nicht beantwortet werden konnte. (11) gibt ein Beispiel eines Dialogs mit einem System, das keine indirekten Antworten erzeugen kann (im Englischen wird dieser Effekt *stonewalling effect* genannt). Eine indirekte Antwort auf die erste Frage des Benutzers wäre (12) gewesen. Man beachte, daß diese Antwort des Systems impliziert, daß Restaurants sehr wohl bekannt sein können, nur keine italienischen.

- (11) Ben: Welche italienischen Restaurants bieten Menüs unter 20 DM an? •
Sys: Keine Antwort gefunden.
Ben: Bei welchen italienischen Restaurants kostet das Menü mehr als 20 DM?
Sys: Keine Antwort gefunden.
Ben: Welche italienischen Restaurants bieten Menüs an?
Sys: Keine Antwort gefunden.
Ben: Kennst Du italienische Restaurants?
Sys: Nein.
- (12) Sys: Mir sind keine italienischen Restaurants bekannt.

Linguistisch gesehen, haben wir es hier mit dem Problem der *Präsuppositionsverletzung* zu tun. Unter Präsupposition versteht man, vereinfacht gesagt, diejenigen Teile einer Äußerung, die implizit für wahr gehalten werden. In unserem Beispiel (11) präsupponiert der Benutzer, daß es italienische Restaurants gibt, die Menüs anbieten.

Kaplan hat in seinem System COOP die Technik realisiert, die Gesamtanfrage in einzelne Anfragen aufzuspalten, um kooperative Antworten zu erzeugen [Kaplan, 1983]. In unserem Beispiel wären das:

1. Gibt es Restaurants?
2. Sind sie italienische Restaurants?
3. Bieten sie Menüs an?
4. Kosten die Menüs weniger als 20 DM?

Je nachdem, welche Teilfrage bzw. Kombination von Teilfragen nicht beantwortet werden kann, werden unterschiedliche Antworten erzeugt. Weiterführende indirekte Antworten

können erzeugt werden, indem obige vier Teilfragen in jedem Fall intern beantwortet werden und dann versucht wird, Alternativen zu verbalisieren (13).

- (13) Sys: Ich kenne keine italienischen Restaurants, bei denen das Menü weniger als 20 DM kostet, aber es gibt griechische und türkische Restaurants die Menüs unter 20 DM anbieten.

Um weiterführende Antworten generieren zu können, muß das System über Fokusinformation verfügen oder versuchen, den neuen Fokus aus vorangegangenen Äußerungen zu bestimmen (14).

- (14) Ben: Welche EUREKA Projekte leitet Peter?
Sys: Keine.
Sys: Keine, aber er leitet ein DFG Projekt.

Für eine andere Art kooperativer Antworten ist es nötig, auf eine explizite Zuordnung von Attributen und Attributwerten zugreifen zu können und zu markieren, daß diese Zuordnung verbalisiert werden muß (15).

- (15) Ben: Welche Telefonnummern haben die HKC Mitarbeiter?
Sys: 447, 547, 488
Sys: Maier, H.-J.: 447, Korb, U.: 547, Willers, B.: 488

Wesentlich komplexere Anforderungen an das Verfahren zur Auswahl und Erzeugung kooperativer Antworten werden gestellt, wenn das System auf Fehlannahmen des Benutzers reagieren soll [McCoy, 1986]. Die typische Situation, die zu diesem Problem führt, ist eine NL Eingabe des Benutzers, die auf Objekte weist, die das System nicht oder anders kennt. Das Beispiel (16) stammt von McCoy.

- (16) Ben: Give me the HULL.NO of all Dcstroyer whose MAST.HEIGHT is above 190.
Sys: All Destroyers that I know about have a MAST_Height between 85 and 90. Were you thinking of the Aircraft-Carriers?

Um korrigierende Äußerungen erzeugen zu können, untersuchte McCoy, ob falsche Auffassungen über Objekte unabhängig vom Diskursbereich klassifiziert werden können, ob sich Antwortstrategien aus einer Klassifikation ableiten lassen und wie die Auswahl zwischen mehreren Antwortstrategien gesteuert werden kann. McCoy unterscheidet im wesentlichen zwei Arten von Fehlannahmen, wobei es für beide Arten je drei Subtypen mit entsprechenden Reaktionsstrategien gibt (siehe [McCoy, 1986]):

- falsche Objektklassifikation (engl. *superordinate misconception*)
- falsche Eigenschaftszuordnung (engl. *attribute misconception*)

Ein weiterer Ansatz zur Erzeugung kooperativer Antworten beruht auf der Erkennung von Benutzerzielen und Benutzerplänen. Im System TRACK [Carberry, 1983] werden auf Grund von Benutzerzielen bestimmte Klassen von Irrtümern des Benutzers (engl. *misconceptions*) erkannt und korrigierende Antworten erzeugt. Das System ARGOT [Allen, 1983] hat drei Hauptziele: Überbeantwortung (engl. *overanswering*), Beantwortung fragmentarischer Eingaben und Erkennung direkter und indirekter Sprechakte. Die Erkennung der

Sprechakte ist dabei die Voraussetzung für die anderen Ziele. Die Interaktionen in (17) und (18) sollen im Rahmen von Zugauskünften möglich sein.

(17) Ben: When does the Montreal train leave?

Sys: 4 o'clock at gate 7.

(18) Ben: The 3:15 to Windsor?

Sys: Gate 4.

Um solche Antworten erzeugen zu können, verfügt ARGOT über vorgefertigte Pläne, die ein Benutzer sinnvollerweise verfolgen kann, wenn er sich auf einem Bahnhof befindet. Eine Skizze des Algorithmus, der die erste der obigen Antworten ermöglicht, ist die folgende:

1. Interpretiere durch syntaktisch/semantische Analyse die Eingabe wörtlich
2. Versuche durch Inferenzen zwischen der Interpretation und einem erwartbaren Ziel, einen Plan des Benutzers zu erkennen
3. Versuche zusätzliche Ziele des Benutzers zu erkennen
4. Plane für die in 3. ermittelten Ziele des Benutzers eine Antwort

Im ersten Schritt wird auf Grund der obigen Eingaben vom System ermittelt, daß der Benutzer einen REQUEST an das System stellt, damit das System ein INFORM über die Abfahrtszeit erzeugt. Nun kann im zweiten Schritt abgeleitet werden, daß die Durchführung des INFORM dazu führt, daß der Hörer die Abfahrtszeit kennt. Dies ist eine der Voraussetzungen dafür, daß der Benutzer den Zug besteigen kann. Als Plan wird also das Erreichen des Zuges inferiert. Im dritten Schritt wird erkannt, daß eine weitere Voraussetzung zur Erfüllung des Plans die Kenntnis des Abfahrtsortes ist. Dieser wird daraufhin im vierten Schritt mitgenannt (Überbeantwortung).

Bei (18), es handelt sich um ein transkribiertes Beispiel, liegt die Hauptschwierigkeit darin, *the 3:15* als eindeutige Referenz aufzufassen und mithilfe der weiteren Eingabe *to Windsor* den Benutzerplan zu erkennen.

Die Beispiele zeigen, daß für viele kooperative Antworten das Erkennen bzw. Inferieren von Zielen und Plänen des Dialogpartners notwendig ist. Im System HAM-ANS wurde kein planbasierter Ansatz für Überbeantwortung verwendet, sondern durch Verbalisierung zusätzlicher Tiefenkasusrollen des vom Benutzer verwendeten Verbs wahrscheinliche Folgefragen antizipiert [Wahlster *et al.*, 1983]. In (19) verbalisiert das System zusätzlich die Rollen SOURCE und GOAL die bei der Identifizierung des infrage kommenden Abbiegen-Ereignisses quasi als Seiteneffekt gefüllt wurden.

(19) Ben: Ist ein PKW abgebogen?

Sys: Ja, einer von der Hartungstraße in die Schlüterstraße.

In jedem Fall ist die Überbeantwortung sinnvoll zu beschränken; weitere Angaben in (17) über Wagenklassen, Speisewagen, Zugtelefon und mögliche Anschlüsse würden über das

Ziel hinausschießen (engl. *pragmatic overshoot*). Die Beschränkung erfolgt in ARGOT implizit durch die Definition des Plans und in HAM-ANS implizit durch den vorgegebenen Kasusrahmen.

Textgenerierung

Textgenerierungssysteme zielen auf die Erzeugung von mehreren zusammenhängenden Äußerungen hin, meist in der Länge eines Paragraphen. Dazu müssen zwei Entscheidungen getroffen werden: was und **wie** es gesagt werden soll. Man kann die bisherigen Ansätze grob in zwei Klassen einteilen: die schemabasierten und die planbasierten. Bei den schemabasierten Ansätzen werden auf Grund empirischer Untersuchungen von Textsorten (Diskurstypen) Schemata aufgebaut, die die typische Struktur eines Textes der jeweiligen Sorte darstellen. Bei den planbasierten Ansätzen wird versucht, auf Grund elementarer Relationen zwischen Sätzen Textstrukturen aufzubauen, die einen bestimmten Zweck erfüllen.

Im folgenden stellen wir zuerst einen schemabasierten Ansatz (TEXT, [McKeown, 1985]) vor. Danach gehen wir auf planungsbasierte Ansätze zur Textgenerierung ein.

Das System TEXT wurde zur Beantwortung von Anfragen über die Struktur einer Datenbank entwickelt. Als Datenbank wurden Teile der Datenbank des ONR (Office of Naval Research) über militärische Fahrzeuge und Waffen benutzt. McKeown beschränkte sich dabei auf drei unterschiedliche Fragetypen:

1. Definitionsfragen: *What ist a frigate?*
2. Informationsfragen: *What do you know about submarines?*
3. Fragen nach dem Unterschied zwischen zwei Objekten: *What is the difference between an ocean escort and a cruiser?*

Für diese Fragetypen hat sie untersucht, wie Menschen diese Fragen beantworteten. Als Ergebnis dieser Untersuchung stellt sie eine Menge rhetorischer Prädikate vor. Ein *rhetorisches Prädikat* klassifiziert eine Äußerung auf zwei Ebenen: es gibt den Typ einer Prädikation an und ihre Funktion in Bezug auf benachbarte Äußerungen. Zwei Beispiele für rhetorische Prädikate verdeutlichen das:

specification	Mary is quite heavy. <i>She weighs 200 pounds.</i>
evidence	The audience recognized the difference. <i>They started laughing right from the very first frame of the film.</i>

Mithilfe dieser rhetorischen Prädikate hat sie Schemata aufgebaut, die typischerweise zur Beantwortung von bestimmten Fragen herangezogen werden und die Struktur einer Antwort (in Paragraphenlänge) beschreiben. In TEXT werden vier Schemata verwendet: *Identification*, *Constituency*, *Attributive* und *Contrastive (Compare & Contrast)*. (20) zeigt das Identification Schema.

- (20) Identification {class & attribute / function}
 {Analogy / Consistency / Attribute / Renaming }*
 Particular-illustration / Evidence +
 • ; {Amplification / Analogy / Attributive}
 {Particular-illustration / Evidence}

Dabei sind Elemente in geschweiften Klammern optional; solche mit einem Stern können fehlen oder n mal auftreten; solche mit einem Pluszeichen treten 1 bis n mal auf; ferner steht & für eine Konjunktion und / für eine Disjunktion von Elementen.

Die Schemata werden den drei verschiedenen Fragetypen zugeordnet: Definitionsfragen werden mit Identification und Constituency behandelt, Informationstragen mit Attributive und Constituency und Fragen nach dem Unterschied von Objekten werden mit Compare & Contrast beantwortet. Wann wird welches Schema ausgewählt? Das Constituency-Schema wird gewählt, wenn in der Datenbasis viele Informationen über die Objekte vorhanden sind, ansonsten wird Identification bzw. Attributive gewählt.

Die Instantiierung eines Textschemas geschieht in fünf Schritten:

1. Instantiiere das erste rhetorische Prädikat und sammle die Propositionen ein; wähle die relevanten Propositionen auf Grund von Fokusinformation aus;
2. führe dies auch für alle alternativen Prädikate durch und füge die neu entstehenden Propositionen zu den bereits existierenden hinzu; wähle die relevanten Propositionen auf Grund von Fokusinformation aus;
3. führe Schritt 2 auch für alle optionalen Propositionen aus;
4. markiere diejenigen Propositionen, die Teil des Antworttextes sind (sie werden somit nicht wiederholt geäußert);
5. übergebe die Sequenz von Propositionen an die taktische Komponente.

Fokusinformation wird in TEXT in zweierlei Hinsicht berücksichtigt: zum einen bei der Auswahl der Propositionen, die geäußert werden sollen, und zum anderen bei der Formbestimmung. Bei der Auswahl der Propositionen wird so vorgegangen, daß für jedes instantiierte Prädikat den einzelnen Propositionen der Fokus mitgegeben wird. Der Fokus ist dabei immer das Objekt, über das eine Aussage gemacht wird. Um in dem obigen Algorithmus die Menge der Propositionen zu ordnen, benutzt McKeown Fokusänderungsregeln, die auf Untersuchungen von Sidner basieren [Sidner, 1983].

TEXT wurde als ein Beispiel eines schemabasierten Ansatzes vorgestellt. Andere Systeme, bei denen auch Schemata in der einen oder anderen Form eingesetzt werden, deren eigentlicher Schwerpunkt jedoch anderen Fragestellungen gilt, sind z.B. SEMTEX/GEO-TEX [Rösner, 1987] und NAOS [Novak, 1987].

Das System KAMP (Knowledge And Modalities Planner) von Appelt [Appelt, 1985] verdeutlicht den planbasierten Ansatz am extremsten. KAMP ist ein integriertes System, bei dem die Trennung von was-sage-ich und wie-sage-ich-es aufgegeben wird, zugunsten eines einzigen hierarchischen Planungsprozesses. Appelt geht davon aus, daß Agenten sowohl

physikalische als auch linguistische Handlungen planen müssen, um ihre Ziele zu erreichen und, daß eine Interaktion zwischen diesen beiden Handlungstypen stattfindet. Ein Beispiel für mehrere Ziele, die ein Agent mit einer einzigen Äußerung erfüllen kann, ist in der Situation gegeben, in der zwei Agenten eine Aufgabe lösen sollen und der Sprecher auf ein Werkzeug deutet und sagt (21).

(21) Use the wheelpuller to remove the flywheel.

Damit *fordert* er den Hörer *auf*, eine bestimmte Handlung auszuführen und *informiert* ihn gleichzeitig darüber, mit welchem Werkzeug er vorgehen soll, unter der Annahme, daß er es nicht weiß. In KAMP werden die Überzeugungen des Hörers explizit in den Planungsprozeß einbezogen.

Appelt geht von einer Hierarchie linguistischer Handlungen aus, die aus vier Ebenen besteht: Auf oberster Ebene sind die illokutionären Akte (die kommunikativen Intentionen) angesiedelt, wie *informieren* oder *auffordern*. Danach kommen die Oberflächensprechakte, wie *Aussage*, *Befehl* oder *Frage*, dann Handlungen zur Konzeptaktivierung, wie *Beschreiben* oder *auf einen Gegenstand zeigen* und schließlich die Äußerungsakte, die konkreten grammatischen Formen. Spezifisches linguistisches Wissen wird erst auf der Ebene der Oberflächensprechakte bedeutsam, da sie abstrakte Repräsentationen von Sätzen mit speziellen syntaktischen Strukturen enthält. Unter der dritten Ebene, *Handlungen zur Konzeptaktivierung*, versteht Appelt folgendes: Vollzieht ein Sprecher einen Oberflächensprechakt, beabsichtigt er, daß der Hörer den propositionalen Gehalt der Äußerung erkennt. Zu diesem Zweck muß der Hörer eine Proposition konstruieren, deren Terme Objekte bezeichnen, an die der Sprecher gerade denkt und auf die er mit seiner Äußerung referiert. Erkennt der Hörer die Terme, so sind die entsprechenden Konzepte aktiviert. KAMP beschränkt sich auf die Planung von Handlungen, deren Konzepte zum gemeinsamen Wissen von Sprecher und Hörer gehören. Auf der letzten Ebene werden die Handlungen zur Konzeptaktivierung in Äußerungsakte überführt, indem die Deskriptoren aus der Konzeptaktivierung in Form von Wörtern und syntaktischen Strukturen realisiert werden.

Da Appelt das Wissen und die Überzeugungen des Hörers explizit in die Planung miteinbezieht, wählte er als Repräsentationssprache Moores Modallogik, die Inferenzen über Überzeugungen (engl. *beliefs*) und Handlungen erlaubt und formalisiert [Moore, 1980]. Die Semantik der intensionalen Operatoren wird in Form möglicher Welten axiomatisiert. Die mögliche Welten Semantik der Modallogik selbst wird mithilfe von Axiomen der Logik erster Stufe dargestellt. Dadurch ergibt sich die Möglichkeit Inferenzen mit einem Theorembeweiser für die Prädikatenlogik auszuführen.

KAMP plant nur einzelne Äußerungen und keine Texte. Es stellt einen wichtigen Beitrag zur Formalisierung von Sprechakten dar, jedoch werden die logikbasierten Planungen von Texten aus Effizienzgründen gegenwärtig wenig verfolgt.

Aus der Sicht der Textgenerierung stellen TEXT und KAMP zwei Extrempunkte dar. TEXT beschreibt den Aufbau eines ganzen Textes mithilfe von Schemata und KAMP plant einzelne Äußerungen beginnend mit der Intention des Sprechers.

3.3 Formbestimmung

Die Form einer Äußerung wird durch zwei Aspekte bestimmt: Festlegung der syntaktischen Struktur eines Satzes und Festlegung der Struktur eines Textes, der aus mehreren Sätzen besteht. Im folgenden wenden wir uns zuerst der Textstruktur zu.

Ein Ansatz, der seine Wurzeln in der Textlinguistik hat und eine Verbindung von Schema- und planbasierten Ansätzen herstellt, ist die Rhetorical Structure Theory (RST) von Mann und Thompson [Mann und Thompson, 1988].¹⁷

Die Grundidee bei der Entwicklung von RST, die als empirisch fundierte deskriptive Theorie entwickelt wurde, besteht darin, Relationen zwischen den Sätzen eines Textes zu beschreiben. In der RST gibt es vier verschiedene Arten von Objekten: Relationen, Schemata, Schemaanwendungen und Strukturen.

Relationen sind zwischen zwei nicht überlappenden Textteilen definiert, die Nukleus (N) und Satellit (S) genannt werden. Die Definition einer Relation besteht aus vier Teilen. Sie werden im folgenden anhand der Relation PURPOSE verdeutlicht:

Definition

1. Constraints für N
2. Constraints für S
3. Constraints für die Kombination von N und S
4. Intendierter Effekt

Beispiel PURPOSE

N stellt eine Handlung dar
S stellt nicht verwirklichte Situation dar
S stellt eine Situation dar, die durch die Handlung in N verwirklicht wird
Hörer erkennt, daß die Handlung in N initiiert wurde, um S zu verwirklichen

Schemata definieren den strukturellen Aufbau eines Textes. Sie bestehen aus Textteilen (mehreren Sätzen), der Relation zwischen den Teilen und einer Spezifikation der Relation der Teile zum gesamten Text. *Schemaanwendungen* sind nichts anderes als Instanzen von Schemata.

Strukturen entstehen durch wiederholte Schemaanwendungen. Ziel ist es, ein Schema zu finden, das den gesamten Text überspannt. Wenn dies gelingt, wird der Text als kohärent angesehen. RST wurde als Methode zur Analyse von Texten entwickelt. Die Definitionen der einzelnen Relationen verlangen vom Analysierenden ein Urteil über die Anwendung der Relation und sind nicht für maschinelle Verarbeitung angelegt. Die Hauptkritik an diesem Ansatz betrifft daher auch die Bedeutung und die Formalisierbarkeit der Relationen.

Inzwischen sind jedoch Formalisierungen dieses Ansatzes entstanden und erfolgreich in hierarchischen Planungssystemen eingesetzt worden. Für die Generierung von Texten werden die RST-Relationen als *Pläne* aufgefaßt und die Bedingungen als Beschränkungen, die erfüllt sein müssen, um den Plan zu instantiieren. Aus elementaren Plänen (RST-Relationen) können größere Pläne aufgebaut werden. Der Planungsalgorithmus führt größere Pläne wieder auf ihre elementaren Bestandteile zurück und versucht, diese zu instantiieren, wenn die Bedingungen erfüllt sind. Als Beispiel zeigen wir hier die Formalisierung der oben angegebenen PURPOSE-Relation (nach [Hovy, 1988a]):

¹⁷Einen direkten Vergleich von TEXT-Schemata und RST findet man in [Mann, 1988].

Constraints für Nukleus:

- 1.(BMB S H (ACTION ?act-1))
- 2.(BMB S H (ACTOR 'act-1 ?agt-1))

Constraints für Satellit:

- 1.CBMB s H (STATE ?state-i))
- 2.(BMB S H (GOAL ?agt-1 ?state-1))
- 3.(BMB S H (RESULT ?act-1 ?act-2))
- 4.(BMB S H (OBJ ?act-2 ?state-1))

Intendierter Effekt:

- 1.(BMB S H (BEL ?agt-1 (RESULT ?act-1 ?state-1)))
- 2.(BMB S H (PURPOSE ?act-1 ?state-1))

Hovy benutzt Operatoren zur Formalisierung von Sprechakten. Der Operator (BMB S H P) bedeutet: die Proposition P folgt aus den Überzeugungen des Sprechers über die gemeinsamen Überzeugungen von Sprecher und Hörer. (BEL S P) bedeutet: P folgt aus den Überzeugungen des Sprechers. Mithilfe dieser Formalisierung können Sprechakte wie *inform* direkt als Pläne auftreten (siehe [Moore und Paris, 1989]), da auch die Effekte auf den Hörer modelliert werden können.

Wir haben schon erwähnt, daß komplexe Pläne aus einer Menge von RST-Relationen aufgebaut werden können. Dadurch ergibt sich die Möglichkeit, mit dem allgemeinen Planungsmechanismus auch Schemata, wie sie von McKeown benutzt werden, aufzubauen und gleichzeitig über die Flexibilität des Planungsansatzes zu verfügen.

RST-Planung soll nun nicht als die Lösung der Probleme im Bereich Textgenerierung angesehen werden.¹⁸ RST-Relationen betreffen die Beziehungen zwischen Sätzen. Bei der Auswahl müssen aber die Einheiten, die von einem RST-Planer geeignet angeordnet werden, erst bestimmt werden. Ein generelles Problem bei der Textplanung besteht darin, sicherzustellen, daß die Ausgabe des Planers auch syntaktisch realisierbar ist (die Kluft zwischen Planergebnis und syntaktischer Form wird auch als „generation gap“ bezeichnet [Meeteer, 1990]). Wie sowohl für die Auswahl als auch für die anschließende Formbestimmung RST-basierte Planungsverfahren eingesetzt werden können, beschreiben [Dobes und Novak, 1991]. Diese Kluft ist bei den meisten Systemen vorprogrammiert, was daran liegt, daß die Planung sich auf den konkreten Zustand der Welt bezieht, wie er im assertionalen Teil der konzeptuellen Repräsentation ausgedrückt ist, der nicht primär für die Verbalisierung entworfen wurde. Bei einer zu engen Kopplung der konzeptuellen Repräsentation an die sprachliche Realisierung enthebt man sich grammatischer Auswahlmöglichkeiten. Ein Ausweg aus diesem Problem ist eine explizite, linguistisch motivierte semantische Repräsentation, die aus der konzeptuellen aufgebaut wird und besser für eine sprachliche Realisierung geeignet ist (für zwei Herangehensweisen an dieses Problem vgl. [Bateman, 1990;

¹⁸Für eine Analyse der Probleme bei der Planung von Paragraphen siehe [Hovy, 1990b], für Probleme die Anzahl der verschiedenen Relationen betreffend [Hovy, 1990a]).

Iordanskaja *et al.*, 1991]).

Die Bestimmung der Form einer Äußerung (d.h. welche syntaktische Struktur gewählt wird, wann koordinierte Phrasen benutzt werden, welche Wortstellung gewählt wird) ist ein wichtiges Forschungsthema. In den meisten Systemen werden zur Bestimmung der syntaktischen Struktur Heuristiken eingesetzt oder das Problem wird dadurch umgangen, daß die zugrundeliegende Wissensrepräsentation, aus der generiert wird, ereigniszentriert aufgebaut ist. Das bedeutet, das Generierungssystem bekommt als elementare Einheiten Ereignisse, mit denen Verben und ihre zugehörigen Rollen assoziiert sind. Daraus kann in einfacher Weise ein Aussagesatz formuliert werden. Komplexe Äußerungen entstehen durch entsprechende Verfahren zur NP-Generierung, deren Ziel die eindeutige Kennzeichnung von Objekten ist. Dieses Ziel kann zur Erzeugung restriktiver Relativsätze führen.

Im TEXT-System z.B. wird der Fokusmechanismus, der die Auswahl steuert auch zur Festlegung der syntaktischen Struktur benutzt. Ist z.B. die Objektrolle eines Verbs im Fokus, wird ein Passivsatz erzeugt, damit die Objektrolle in Subjektposition erzeugt werden kann. Bei konversen Verben (z.B. *kaufen* und *verkaufen*) steuert der Fokus die Verbwahl.

In der systemischen Grammatik (siehe Abschnitt 3.6) wird die konkrete Oberflächenform und die Wortwahl durch Merkmale gesteuert, deren Werte von den Anwendungsprogrammen bestimmt werden müssen. In diesem Grammatikmodell wird explizit versucht, die Intention des Sprechers und die Kommunikationssituation als Einflußgrößen der Oberflächenform zu repräsentieren.

3.4 Sprachliche Realisierung

In der KI sind mehrere Realisierungssysteme entstanden, deren Gemeinsamkeit darin besteht, daß sie Repräsentationen einer formalen Satzbeschreibungssprache in korrekt flektierte Oberflächensätze überführen. Dabei wird davon ausgegangen, daß die Wortwahl bereits stattgefunden hat. Motiviert sind diese Systeme dadurch, daß es durch sie möglich wird, Fragen der Auswahl und Textstruktur zu untersuchen, ohne auf die tatsächliche Ausgabe einer korrekt flektierten Oberfläche verzichten zu müssen.

Beim Entwurf des Systems MUMBLE [McDonald, 1983] ließ sich McDonald von Beobachtungen über menschliches Verhalten leiten. Sein wichtigstes Ziel ist es, der Geschwindigkeit menschlicher Sprecher nahezukommen, was in MUMBLE, zumindest für die angegebenen Beispiele, auch erreicht wird. Weiterhin arbeitet das System ohne Backtracking, d.h., daß einmal getroffene Entscheidungen zur Realisierung interner Strukturen durch Sprache nicht mehr zurückgenommen werden können, sondern nur durch weitere sprachliche Äußerungen erweiterbar sind. Dieses Determinismusprinzip hilft ihm, den Verarbeitungsaufwand zur Erzeugung einer Äußerung gering zu halten.

MUMBLE geht von einer Datenstruktur, der sog. *message*, aus, die ein Anwendungsprogramm (ein Experte in einer Domäne) ohne sprachliches Wissen erzeugt hat. Es wird gefordert, daß die *messages* disambiguiert sind.¹⁹ MUMBLE besteht aus einem Lexikon

¹⁹Die Ebene der *message* entspricht der präverbalen Bedeutungsstruktur, die im Abschnitt 2 vorgestellt worden ist.

und einer *linguistischen Komponente*. Das Lexikon ist domänenabhängig und enthält Informationen darüber, wie *messages* zu interpretieren sind, insbesondere, welche Wörter benutzt werden sollen.

Der Generierungsprozeß ist zweistufig. In der ersten Phase wird aus der *message* ein Syntaxbaum aufgebaut, der die Struktur des Oberflächensatzes repräsentiert. In der zweiten Stufe wird dieser Baum von links nach rechts traversiert, wobei die kategorialen Blätter des Baumes rekursiv durch Lexikoneinträge ersetzt werden und somit die Oberflächenstruktur erzeugt wird. Als Seiteneffekt der Abarbeitung des Baumes werden ein grammatischer Kontext aufgebaut und eine Liste der Diskurselemente erstellt, die zur Erzeugung von Anaphern notwendig ist, und es werden grammatische Restriktionen propagiert.

Eine Erweiterung, MUMBLE-86 [Meteer *et al.*, 1987], erleichtert den Zugang zu unterschiedlichen Anwendungsprogrammen durch eine Schnittstelle, an der für jede zu erzeugende Äußerung folgende Information erwartet werden:

- Die Abbildung jeder zu generierenden Einheit auf die zu erzeugende Struktur;
- die funktionalen Beziehungen zwischen den Einheiten z.B. Prädikation, Head, item Modifikator, Thema, etc., die die Organisation der Einheit im Text einschränken;
- die konkreten Wörter.

Damit obliegt allerdings dem Benutzer, die möglichen syntaktischen Strukturen zu kennen und die Abbildungsvorschriften von Konzepten seiner Wissensrepräsentation in Wörter festzulegen.

Das modulare und anpaßbare System SUTRA (Surface Transformations) [Busemann, 1988], das speziell für die deutsche Sprache entwickelt wurde, geht von Satzrepräsentationen aus, die durch einen vorgeschalteten beliebigen Verbalisierungsprozeß erzeugt werden. Zur Verbalisierung zählen Wortwahl, Wahl der Kennzeichnung, der Verbzeit, von Aktiv bzw. Passiv, des Satzmodus etc. Die Ausgangsstruktur von SUTRA wird in der Repräsentationssprache VS (Verbalisierte Struktur) dargestellt, deren Syntaxdefinition direkt zur Verarbeitung herangezogen wird. VS-Ausdrücke enthalten folgende Information:

- alle Wortstämme
- Genus, Modus und Tempus des Verbs
- thematische Rolle oder syntaktische Funktion eines Satzglieds
- Numerus der Nominalphrase
- Formulierung von Satzgliedern als Personal- oder Relativpronomen
- Komparation bei Adjektiven
- ggf. Elision von Satzgliedern
- ggf. Elision des Nomens einer Nominalphrase.

Um eine Ausgangsstruktur in einen Oberflächensatz zu transformieren, werden drei Wissensquellen benutzt:

- ein Wortlexikon
- eine Menge von Regeln zur Satzgliedstellung, sowie
- Flexionsparadigmen.

SUTRA wurde im Dialogsystem HAM-ANS und in anderen Systemen eingesetzt. Eine Erweiterung des Leistungsumfangs erfolgte im System FREGE, das objektorientiert implementiert wurde [Emele, 1987].

MUMBLE-S6, SUTRA und FREGE sind Front-End-Generatoren, die von vollständig spezifizierten Eingabestrukturen ausgehen. MUMBLE unterscheidet sich von den anderen Systemen vor allem dadurch, daß es keine linguistisch motivierten Eingabestrukturen voraussetzt. Dafür ist der Anpassungsaufwand des Lexikons sehr hoch. Durch allgemein gehaltene Schnittstellen wie VS kann eine weitestgehende Unabhängigkeit von vorzuschaltenden Generierungskomponenten sowie von Anwendungsprogrammen erzielt werden. Keiner der obigen Generatoren verfügt über eine linguistisch fundierte Grammatik wie z.B. LFG, HPSG oder GPSG. Die jeweiligen Grammatiken sind prozedural kodiert und erheben keinen theoretischen Anspruch. Die Implementation der Systeme begann auch vor der Entwicklung dieser Grammatiktheorien.

3.5 Wortwahl

Die Wahl von Inhaltswörtern (engl. *open class words*), bei der es sich um Abbildungen konzeptueller Strukturen auf lexikalische Semantik handelt, läuft beim Menschen mit extremer Geschwindigkeit ab. Im Durchschnitt äußern Menschen ca. 150 Wörter pro Minute, wobei diese Geschwindigkeit unter Zeitdruck verdoppelt werden kann. Vermutlich gibt es keinen anderen kognitiven Prozeß mit einer solch hohen Entscheidungsrate. Erstaunlich ist auch, daß diese Rate über lange Zeit gehalten werden kann und der Prozeß anscheinend sehr genau ist. Es wird angenommen, daß dieser schnelle Zugriff nur durch parallele Prozesse geleistet werden kann. Psycholinguistische Ergebnisse und theoretische Ansätze werden in [Levelt, 1989, Kap. 6] zusammengefaßt.

Dieser Abschnitt befaßt sich mit der Frage, nach welchen Kriterien und mit welchen Verfahren ein Generierungssystem Lexeme auswählen könnte, um eine gegebene Bedeutung in geschriebener Sprache zu verbalisieren. Bei allen Theorien der Wortwahl (engl. *lexical choice*) muß das *Konvergenzproblem* gelöst werden: es wird stets eine Entscheidung für *genau eine* Alternative getroffen. Wir unterscheiden folgende Aspekte der Wortwahl:

Definite Kennzeichnungen, Proformen: Mit welchen Wörtern kann ein Ereignis oder Objekt beschrieben werden, so daß es für den Hörer eindeutig identifizierbar ist? Wann ist die Verwendung von Proformen angemessen? (Hierauf wurde in Abschnitt-2 (S. 15ff) bereits eingegangen.)

Soziale Bewertungen: Welche Assoziationen ruft die Verwendung eines Wortes beim Hörer hervor? Am Unterschied zwischen *Putzfrau* und *Raumpflegerin* lassen sich Bewertungen und deren Konventionalisierung ersehen.

Kollokationen: Welche Kookkurrenzbeschränkungen zwischen Wörtern sind zu berücksichtigen?

Wahl von Inhaltswörtern (engl. *open class words*): Ist ein bestimmtes Transferereignis (Waren gegen Geld) durch *verkaufen*, *bezahlen*, *kaufen* oder *einnehmen* am besten zu beschreiben? Kennzeichnet man ein bestimmtes Lebewesen z.B. als *Knollenblätterpilz*, *giftigen Pilz* oder einfach als *Pilz*!

Wir gehen im folgenden zunächst auf Kollokationen und danach auf die Wahl von Inhaltswörtern ein. Kollokationen sind rein lexikalische Relationen zwischen einer Basis und einem Kollokator, wobei die Basis die Wahl des Kollokators beschränkt. Welche Kollokationen möglich sind, wird in den Lexikoneinträgen der Basis mithilfe lexikalischer Funktionen (vgl. etwa [Iordanskaja *et o./.*, 1991]) dargestellt. In (22) beschränkt das Nomen als Basis das Verb; soll das Verb zuerst festgelegt werden, ist zu beachten, ob es als Kollokator fungieren kann, falls eines seiner Argumente die Basis einer Kollokation ist.

(22) Pläne schmieden- Vorhaben * schmieden

Dieser Ansatz unterscheidet sich von dem im System DIOGENES [Nirenburg und Nirenburg, 1988] durch die explizite Trennung von Basis und Kollokator. In DIOGENES kann andererseits der lexikalische Kopf (z.B. das Verb) unabhängig von der Realisierung seiner Komplemente bestimmt werden.

Wir skizzieren nun die wichtigsten Techniken für die Bestimmung von Inhaltswörtern. Zwar wird in vielen Generierungssystemen konzeptuelles Wissen benutzt; die Modellierungskriterien für dieses Wissen sind aber unterschiedlich. Von einem theoretischen Standpunkt aus kann man das konzeptuelle Wissen anhand einer Theorie mentaler Kategorien hierarchisch darstellen. Anstelle dieser theoretischen Sicht tritt in der Praxis gewöhnlich eine am Zweck des jeweiligen Systems (u.U. ist dies gerade Verarbeitung von NL) orientierte Modellierungsweise. Es ist offensichtlich, daß die jeweilige Aufgabenstellung bei der Wahl von Inhaltswörtern von der jeweiligen Struktur des konzeptuellen Wissens abhängt.²⁰

Direkte Ersetzung: Diese oft verwendete Technik geht von einer Konzeptrepräsentation aus, die eine eindeutige Beziehung zu Lexemen gestattet (s.o.). Für ein gegebenes Konzept (z.B. PILZ) ist das Lexem (*Pilz*) eindeutig bestimmbar. Dieser Ansatz umgeht eigentlich das Wortwahlproblem.

Strukturersetzung: In der semantischen Ausgangsstruktur werden Teilstrukturen gesucht, die mit lexikalischen Einheiten in bestimmter Hinsicht übereinstimmen, und die gefundenen Lexeme ersetzen die Teilstrukturen. Das Verfahren ist beendet, wenn alle Elemente der Ausgangsstruktur durch lexikalische Einheiten ersetzt sind.

²⁰Zur Modellierung der Wissensbasis im LILOG-System siehe den Erfahrungsbericht von [Klose *et al.*, 1992].

Klassifikation: Das Ziel ist, das Lexem zu finden, das dem zu verbalisierenden Konzept am nächsten kommt. Bekannt sind Entscheidungsbäume [Goldman, 1975], in denen die möglichen Verbalisierungen eines Konzepts (z.B. INGEST) an den Blattknoten kodiert sind. Je nach Belegung verschiedener Merkmale der Rollenfüller (Etwa: Ist Agens ein Mensch? Ist der Aggregatzustand von Patiens „flüssig“?) erhält man unterschiedliche Lexeme (z.B. *essen/fressen, trinken/saufen, atmen*).

Strukturersetzung und Klassifikation können kombiniert werden, indem man die Klassifikationstechnik bei der Suche innerhalb des Struktursetzungsverfahrens verwendet. Alle vorgestellten Verfahren berücksichtigen nicht, daß die Wahl von Inhaltswörtern nicht nur vom propositionalen Gehalt der semantischen Repräsentation abhängt, sondern auch von dem vermuteten Wissen des Hörers, den Zielen des Sprechers, dem Kontext und den Konversationsmaximen.

Betrachten wir näher, wie diese Faktoren die Wortwahl beeinflussen. Um ein Objekt KBP-2 mit dem Ziel zu beschreiben, das Objekt als gefährlichen Pilz zu kennzeichnen, stehen dem Sprecher verschiedene Alternativen zur Verfügung. Sie ergeben sich aus dem Konzept des Objekts, KBP, und seinen Oberbegriffen, etwa PILZ und GIFTIGES-OBJEKT, sowie den diesen Konzepten assoziierten Wörtern (z.B. *Knollenblätterpilz, Pilz, giftig*). Der Sprecher kann also KBP-2 z.B. wie in (23) oder wie in (24) kennzeichnen. Die Angemessenheit der ersten Möglichkeit hängt davon ab, ob der Sprecher glaubt, daß der Hörer Knollenblätterpilze als giftig klassifiziert. Glaubte der Sprecher dies nicht, würde er mit (23) sein Ziel nicht erreichen und müßte auf (24) zurückgreifen. Glaubte er es aber, so muß er möglicherweise (23) verwenden, um falsche konversationale Implikaturen [Grice, 1975] zu vermeiden: Der Hörer folgert vielleicht (abhängig vom Kontext), daß KBP-2 kein Knollenblätterpilz ist, denn sonst hätte der Sprecher ihn als solchen gekennzeichnet. Grundlegende Vorschläge für eine Operationalisierung dieser Inferenzen finden sich bei [Reiter, 1990].

(23) Da ist ein Knollenblätterpilz.

(24) Da ist ein giftiger Pilz.

(25) Da steht ein Pilz.

Bei einem anderen Sprecherziel, etwa dem, den Hörer auf KBP-2 aufmerksam zu machen, erscheint keine der beiden obigen Varianten optimal. Naheliegender wäre eine Äußerung wie (25), bei der der Oberbegriff PILZ verbalisiert wurde. Psychologische Untersuchungen haben gezeigt, daß Menschen oft Wörter verwenden, die auf Basislevel-Kategorien (engl. *basic level classes*) verweisen [Rösch, 1978]. Basislevel-Kategorien werden von Kindern früh gelernt, auf sie reagieren Menschen in Reaktionszeit-Experimenten am schnellsten und sie werden von Menschen in „Nullkontexten“ verbalisiert. In unserem Beispiel ist PILZ die speziellste Basislevel-Kategorie, nicht aber KBP.

Weitere Kontextinformation, über Thema und Fokus, geht in die Wortwahl ein. Ein Autokauf etwa kann durch lexikalisch konverse Verben beschrieben werden (26)-(29). Bei allen Verben können sämtliche Mitspieler am Autokauf (Käufer, Verkäufer, Auto, Geld) verbalisiert werden, aber nicht alle sind obligatorisch. Auch wenn ein Mitspieler nicht

verbalisiert wird, kann seine Existenz im beschriebenen Ereignis gefolgert werden. Die Verben bringen unterschiedliche Mitspieler in *Perspektive*. Die Verbbestimmung muß also berücksichtigen, welche Objekte im Diskurs in Perspektive sind. [Jacobs, 1987] zeigt, wie in diesem Zusammenhang lexikalisches und konzeptuelles Wissen in ein- und demselben Wissensrepräsentationsformalismus bezogen werden kann.

(26) Hans hat ein Auto von Peter für 800 DM gekauft.

(27) Peter hat ein Auto an Hans für 800 DM verkauft.

(28) Hans hat 800 DM für ein Auto an Peter gezahlt.

(29) Peter hat 800 DM für ein Auto von Peter eingenommen.

Kontextwissen wird ebenfalls für die Wahl von Dimensionsadjektiven erforderlich. Allein aufgrund der definitiven Merkmale der Dimensionen etwa einer Stange kann man deren Maximale nicht beschreiben. Ob (30) oder (31) angemessen ist, hängt von ihrer kontextuell gegebenen Lage ab; nur wenn die Stange aufrecht steht, kann man (30) sagen.

(30) Die Stange ist 10 m hoch.

(31) Die Stange ist 10 m lang.

[Lang *et al.*, 1991] schlagen eine zweistufige propositionale semantische Repräsentation vor, bei der sprachunabhängige konzeptuelle und lexikalisch-semantische Entitäten aufeinander bezogen werden. U.a. werden die konzeptuellen Merkmale von räumlichen Objekten (STANGE hat eine maximale Achse) mit lexikalischen Restriktionen der Dimensionsadjektive kombiniert (*hoch* verlangt eine vertikale Orientierung der maximalen Achse, *lang* nicht).

In die Auswahl von Dimensionsadjektiven geht ebenfalls die räumliche Lokalisierung des Sprechers ein. Dies trifft ebenso für räumliche Präpositionen zu. (32) ist mehrdeutig: in seiner deiktischen Lesart ist die Relation zum Sprecher oder Hörer ausschlaggebend, in der intrinsischen Lesart die ausgezeichnete Seite des Referenzobjekts (hier: die Frontseite des Autos). Der Gebrauch räumlicher Präpositionen wurde im System CITYTOUR untersucht; den theoretischen Hintergrund und die Realisierung im System beschreibt [Retz-Schmidt, 1988].

(32) Der Ball ist vor dem Auto.

Wir wollen abschließend erwähnen, daß Wortwahl von arideren Entscheidungen abhängt. [Danlos, 1987, Kap. 3] zeigt den Zusammenhang mit der Reihenfolge der Informationspräsentation, der Satzsegmentierung und der Verwendung von Aktiv bzw. Passiv. Die Stellung eines Wortwahlsystems im Generierungsprozeß ist derzeit kontrovers.

3.6 Systemische Grammatik

Viele Beispiele zeigen, daß die Bestimmung der Form nicht unabhängig von der Bestimmung des Inhalts einer Äußerung gesehen werden kann. Auch die oben erwähnten modernen linguistisch fundierten Grammatiktheorien gehen von einer eindeutigen semantischen Repräsentation der Äußerung aus, der logischen Form, und erzeugen daraus einen

Oberflächensatz. Pragmatische Merkmale werden derzeit nicht berücksichtigt, wie z.B. die Tatsache, daß ein Sprecher, der Untergebener des Hörers ist, eine Aufforderung eher als Bitte denn als Befehl äußert. Ein im Bereich Sprachgenerierung sehr einflußreiches Gram-

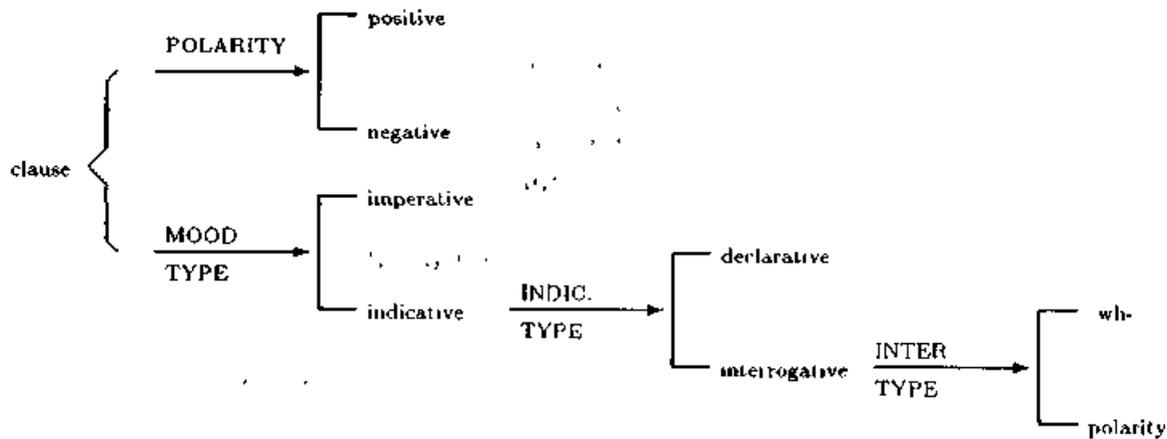


Abbildung 6: Ein Teil des Modus Netzwerkes

matikmodell ist die systemische Grammatik [Halliday, 1985]. Diese Grammatik beruht nicht auf der Mathematik oder der formalen Logik sondern auf der Anthropologie und Soziologie. Ihre Entwicklung wurde weniger durch Fragen der Grammatikalität oder des Erwerbs linguistischer Kompetenz vorangetrieben, sondern durch die Frage nach der Funktion von Sprache als sozialer Aktivität: *Was sind die sozialen Funktionen von Sprache?* und *Wie erfüllt Sprache diese sozialen Funktionen?*

In der systemischen Grammatik gibt es für die drei grundlegenden sprachlichen Mittel, Bedeutung zu konstruieren, verschiedene Ebenen der Beschreibung: Klassifikation (z.B. Wortart), Gruppierung (z.B. Wortsequenzen zu Konstituenten), Funktion (z.B. Beziehung von Elementen untereinander). Mit diesen Beschreibungsmitteln wird Sprache entlang verschiedener Dimensionen beschrieben: der Erfahrung/Erfassung der Welt (experiential), der logischen Zusammenhänge (logical), der personellen Beziehungen (interpersonal) und der textuellen Beziehungen (textual). Diese Dimensionen werden durch abstrakte Merkmale beschrieben und eine Oberflächenform ist das Ergebnis der Auswahl aus diesen abstrakten funktionalen Merkmalen.

Systemische Grammatiken werden durch Systemnetzwerke repräsentiert, die Merkmale und ihre möglichen Werte, sowie konjunktive (geschweifte Klammer) oder disjunktive Verknüpfungen (das U mit dem Pfeil) darstellen (siehe Abb. 6). Ein Satz wird z.B. dadurch bestimmt, daß sowohl ein Wert für die *Polarität* als auch ein Wert für den *Modus* angegeben wird. *Polarität* kann die Werte *positiv*, *negativ* annehmen und *Modus* die Werte *imperativ*, *indikativ*. Für jeweils einen dieser Werte muß man sich entscheiden. Die Grammatikalität einer Äußerung wird durch das Traversieren derartiger Netzwerke sichergestellt, die der Merkmalsauswahl dienen.

In der systemischen Linguistik wurde von vornherein versucht, alle Aspekte, die für die Erzeugung einer Äußerung von Bedeutung sind, mithilfe von Merkmalen und ihren möglichen

chen Werten zu beschreiben. Eine sehr große Grammatik des Englischen ist die NIGEL Grammatik, die am Information Sciences Institute entwickelt wurde und im PENMAN Generierungssystem eingesetzt wird [Matthiessen, 1985]. Die Generierung folgt dem grammatikgetriebenen Ansatz, d.h. eine Äußerung entsteht als Seiteneffekt der Traversierung der systemischen Netzwerke. Die Werte der einzelnen Merkmale werden von der Umgebung bestimmt. Wenn z.B. entschieden werden muß, ob der Wert des Merkmals *Modus imperativ* oder *indikativ* ist, so wird eine Anfrage an die semantische Repräsentation gestellt, ob es sich um einen Befehl handelt oder nicht. Für jedes Auswahlsystem sind derartige Auswahlprädikate definiert, die Tests anhand des Hintergrundprogramms durchführen, um die Merkmalswerte zu bestimmen. Wenn ein solches Prädikat zwischen definitem und indefinitem Artikel wählen muß, wird gefragt, ob der Kopf der NP auf ein generisches oder ein individuelles Konzept der Wissensbasis verweist und ob das Diskursmodell zeigt, daß das Objekt bereits vorerwähnt ist. Diese Art der Merkmalsbestimmung wird „inquiry semantics“ genannt und ermöglicht es der Grammatik, mit den Wissensquellen des Hintergrundprogramms zu kommunizieren.

Den vier Dimensionen systemischer Beschreibung entsprechen verschiedene Wissensquellen eines sprachverarbeitenden Systems:

experiential	- Welt-, Hintergrundwissen
logical	- semantische Repräsentation
interpersonal	- Diskurs-, Benutzermodell
textual	- Textstrukturregeln

Im System PENMAN ist das Weltwissen und die semantische Repräsentation in KL-ONE bzw. einer Erweiterung (LOOM) dargestellt und die inquiry semantics vermittelt zwischen Grammatik und der Wissensrepräsentation. Insgesamt zeichnen sich systemische Grammatiken durch eine holistische Sicht auf Sprache aus, wie sie von anderen Grammatikmodellen (noch?) nicht geteilt wird.

4 Generierung mit modernen Grammatikformalismen

Die linguistischen Syntaxtheorien der frühen Achtziger Jahre leiteten Bestrebungen zu ihrer Modularisierung und zu ihrer Unabhängigkeit von Verarbeitungsstrategien ein.

Modularisierung: Die deskriptive Definitionsweise der Linguistik spiegelt sich in einer deklarativen Repräsentationsweise von sprachlichem Wissen wider. Kontrollwissen (Parsing, Generierung), linguistischer Formalismus und Grammatik sollten voneinander getrennt definiert sein. Dies stellt eine Voraussetzung für den Austausch von Grammatiken dar, ohne die Kontrollstrategie zu verändern. Dies ermöglicht u.a. multilinguale Generierung, indem dasselbe Verfahren mit Grammatiken für Fragmente verschiedener Sprachen verwendet wird.

Bidirektionalität: Generierung und Parsing sollten mit demselben linguistischen Formalismus und derselben Grammatik realisiert werden können. Neben psycholinguistischen Argumenten sprechen ökonomische Gesichtspunkte gegen eine Duplizierung des grammatischen Wissens in einem NL System. Damit die Grammatik bidirektional verwendbar ist, muß sie modular im obigen Sinne sein. Man nennt sie dann oft *reversibel* oder *non-direktional*. Eine reversible Grammatik beschreibt *genau* die in einer natürlichen Sprache zulässigen syntaktischen Strukturen und die ihnen zugeordneten Bedeutungen. Eine weitergehende Frage ist, ob auch die Verfahren sich als Parametrisierungen eines einzigen, bidirektionalen Kontrollmechanismus effizient beschreiben lassen.

Effizienz: Zu einem Formalismus werden laufzeiteffiziente Verarbeitungsstrategien gesucht. Entscheidenden Einfluß übt die Formulierung der Grammatiken aus. In den meisten Fällen erweist sich die direkte Interpretation einer größeren, nach linguistischen Kriterien formulierten Grammatik als zu aufwendig. Daher werden die Grammatiken in der Vorverarbeitung so kompiliert, daß sie für gegebene Verfahren optimal verarbeitet werden können. Dies ergibt unterschiedliche Objektgrammatiken für Parser und Generator, wobei die jeweils zugrundeliegenden linguistisch motivierten Grammatiken identisch sein können.

Im folgenden befassen wir uns mit Generierungsverfahren für sehr verschiedene Grammatikformalismen. Gegenwärtig stehen constraintbasierte Ansätze im Mittelpunkt der Forschung und unter ihnen bildet die Head-Driven Phrase Structure Grammar (HPSG) einen Schwerpunkt. Mit der Entwicklung von HPSG [Pollard und Sag, 1987] verstärkt sich der bereits mit LFG vorhandene Trend zu lexikonzentrierten Theorien. Die meiste Information wird in den Lexikoneinträgen kodiert, während nur wenige Phrasenstruktureregeln verbleiben (HPSG sieht vier für das Englische vor). Lexikonzentrierte Theorien beruhen auf einer integrierten Semantikdarstellung (zu getrennter vs. integrierter Semantik vgl. 5.3).

Alle diese linguistischen Formalismen beschreiben die Syntax und ggf. eine oberflächen-nahe Semantik für NL Sätze. Die Generatoren können somit die Funktion der Realisierungskomponente in einem umfassenderen Generierungssystem ausfüllen.

Die Implementation als constraintbasiertes System ist nicht an diese durch linguistische Theorien induzierte Grenze gebunden; vielmehr wird neuerdings versucht, constraintbasierte Ansätze auch für semantische und pragmatische Prozesse zu entwickeln.

Im folgenden Abschnitt 4.1 wird inkrementelle Realisierung (mit Baumadjunktionsgrammatiken und Segmentgrammatiken) skizziert. Danach gehen wir in Abschnitt 4.2 ausführlich auf Generierung von GPSG-Strukturen aus einer getrennten Semantikdarstellung ein. Zwei weitere Abschnitte befassen sich mit integrierten Semantikdarstellungen; 4.3 beschreibt Generierung aus f-Strukturen in LFG und 4.4 die Steuerung durch den „semantischen Kopf“, die u.a für HPSG eingesetzt wird. Wir schließen mit einer Betrachtung über bidirektionale Verfahren (Abschnitt 4.5).

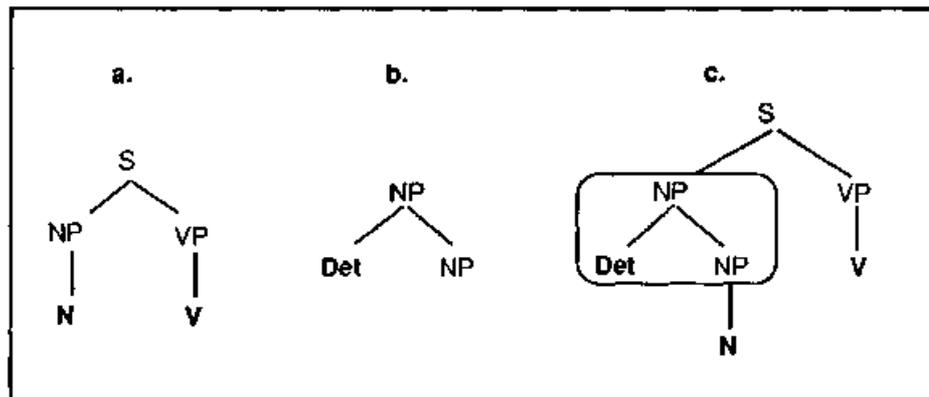


Abbildung 7: Initialer Baum (a), auxiliärer Baum (b) und Adjunktion beider (c).

4.1 Inkrementelle Generierung mit Baumadjunktionsgrammatiken (TAG)

Baumadjunktionsgrammatiken (engl. *Tree Adjoining Grammars* (TAG)) waren ursprünglich nicht als eine linguistische Theorie, sondern als ein Formalismus zur Baumerzeugung gedacht. Viele formale und linguistisch relevante Eigenschaften wurden inzwischen untersucht und in unterschiedlichen Dialekten von TAGs implementiert (einen Überblick gibt etwa [Joshi, 1986]).

Bei TAGs unterscheidet man zwei Mengen von *elementaren* Bäumen, *initiale* und *auxiliäre* Bäume. Initiale Bäume können als kontextfreie Ableitungsbäume aufgefaßt werden. Auxiliäre Bäume haben als Wurzel ein Nontenninal X, das auch an einer Tochter, dem Fußknoten, vorkommt. Sie können in einen initialen oder abgeleiteten Baum T eingesetzt werden, indem ein Nonterminal X in T ersetzt wird. Diese Operation heißt *Adjunktion* [vgl. Abb. 7. Um Endlosrekursionen zu vermeiden, müssen auxiliäre Bäume mindestens ein Terminalsymbol einführen.

Infolge der Adjunktionsoperation ist TAG schwach kontextsensitiv, was als angemessen zur Beschreibung von NL gilt.

Linguistisch motiviert ist die Einführung von Merkmalsunifikation in TAG. Dies erleichtert die Darstellung z.B. von Kongruenzphänomenen. Bei Adjunktion müssen Wurzel und Fuß des auxiliären Baumes bezüglich der vorher bestehenden Kospezifikationen unifizieren. Die Schwierigkeit besteht darin, unerwünschte Fehlschläge von Unifikationen innerhalb des adjungierten auxiliären Baumes zu vermeiden.

Analog zu GPSG wird zur flexiblen Behandlung von Wortstellung mit Varianten des ID/LP-Konzepts²¹ experimentiert. Ferner werden oft *lexikalisierte* TAGs benutzt, bei denen alle elementaren Bäume mindestens ein Terminalsymbol, den lexikalischen Head, einführen. Dadurch, daß in TAG elementare Bäume eine Tiefe größer als 1 haben dürfen,

²¹Darunter wird die Aufteilung der in kontextfreien Regeln enthaltenen Information auf zwei verschiedene Regelarten verstanden: *ID-Regeln* (ID = immediate dominance) drücken die Dominanzbeziehung zwischen der Mutter und den Töchtern aus; die *LP-Aussagen* (LP = linear precedence) definieren die grammatisch zulässigen Abfolgen der Töchter.

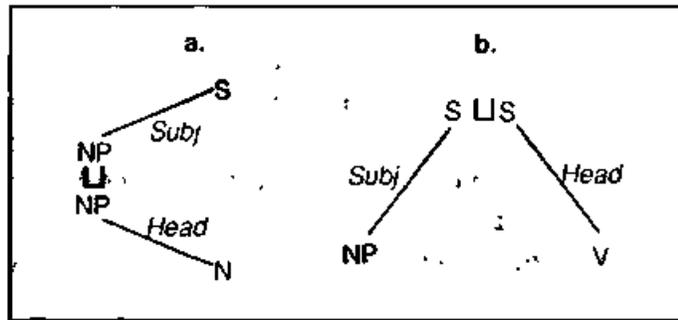


Abbildung 8: Konkatination (a) und Gabelung (b) in Segmentgrammatiken.

lassen sich manche sprachlichen Regularitäten durch einen einzigen Baum darstellen (z.B. Verb mit allen subkategorisierten Konstituenten). Dies ist kompakter als etwa bei Bäumen aufgrund von kontextfreien Phrasenstruktur-Regeln, führt aber zu Redundanz im Lexikon, falls die Bäume direkt kodiert werden.

Eine weitgehend redundanzfreie lexikalische Repräsentation in einem hierarchisch strukturierten Lexikon mit Vererbung würde eine Segmentierung elementarer Bäume in gemeinsame und unterschiedliche Anteile erfordern. Im unifikationsbasierten Formalismus der *Segmentgrammatik* [Kempen, 1987a; De Smedt, 1990] werden solche Segmente in der Form unärer lokaler Bäume (lokal heißt: der Tiefe 1) als Basiselemente angesetzt. Sie bestehen aus einer Wurzel- und einer Fußkategorie sowie aus einer beide verbindenden Relation. Segmente werden auf zwei Arten miteinander kombiniert (vgl. Abb. 8): durch Konkatination (Unifikation des Fußes des einen mit der Wurzel des ändern Segments) und Gabelung (engl. *furcation*, Unifikation beider Wurzeln). Subkategorisierungsinformation wird durch komplexe Ausdrücke im Lexikon dargestellt, die die Gabelung entsprechender Segmenttypen verlangen.

Segmentgrammatiken werden für inkrementelle Realisierungsverfahren eingesetzt [De Smedt, 1990], desgleichen unifikationsbasierte, lexikalisierte ID/LP-TAGs [Harbusch *et al.*, 1991]. Die Verfahren können in eine Kaskade verschiedener Generierungskomponenten nach der Inhaltsplanung und dem Textdesign als dritte Stufe integriert werden. Zum Vergleich von Segmentgrammatiken mit TAGs siehe [De Smedt, 1990, S. 51ff].

Als Eingabe für die Realisierung mit TAGs werden Lemmata (z.B. *Peter, Maria, helfen*) und funktionale Beziehungen zwischen ihnen (z.B. *agens(helfen) = Peter*) erwartet. Infolge der Lexikalisierung ist es unmittelbar möglich, zu jedem eintreffenden Lemma die Menge initialer Bäume zuzuordnen, die das entsprechende Lexem als lexikalischen Kopf besitzen. Die partiellen Strukturen werden miteinander kombiniert, wobei Realisierungen weiterer Lemmata durch Adjunktion eingefügt werden. Die resultierenden Strukturen werden danach durch LP-Regeln geordnet. Abschließend werden die Terminalsymbole durch eine Flexionskomponente flektiert und ausgegeben. Kombination, Anordnung und Flexion laufen ebenfalls inkrementell ab.

Im Unterschied hierzu beschreiben [McDonald und Pustejovsky, 1985] eine Verzahnung von Entscheidungs- und Realisierungsprozessen, indem sie, wie bei MUMBLE, determini-

stische Generierung zugrundelegen, die ausschließlich durch die Eingabestruktur gesteuert wird (engl. *message-driven*). Die Eingabestruktur spezifiziert u.a. rhetorische und pragmatische Kriterien, nach denen der Generator über die Formulierungen aus einer Menge semantisch äquivalenter Alternativen entscheidet. Er bestimmt, an welchen zulässigen Expansionsstellen (engl. *attachment points*) der syntaktischen Struktur weiteres Material eingefügt werden soll. Dort leitet er eine geeignete Adjunktionsoperation ein.

Lexikalisierte TAGs sind für inkrementelle Generierung gut geeignet. Die lexikalische Information wird bei Zugriff in die syntaktische Struktur integriert, da weiteres Material ggf. durch Adjunktion „zwischen geschoben“ werden kann. Durch die relativ komplexen elementaren Bäume werden oft genügend Constraints für die deterministische Ausführung nachfolgender Schritte erzeugt.

4.2 Generalisierte Phrasenstruktur-Grammatiken (GPSG)

GPSG ist keine lexikonzentrierte Theorie. Im Mittelpunkt steht die Definition eines zulässigen kontextfreien Ableitungsbaums. Eine Vielzahl unterschiedlicher Komponenten wirken gemeinsam, indem sie die Zulässigkeit lokaler Bäume beschränken. Die von einer GPSG-Grammatik beschriebenen Sätze sind diejenigen, zu denen Ableitungsbaume aus zulässigen lokalen Bäumen zusammengesetzt werden können. Mit GPSG wurde zum ersten Mal der Versuch unternommen, die Regelformate und die Effekte von Regeln eines Grammatikformalismus streng formal und deklarativ zu definieren [Gazdar *et al.*, 1985]. Wir skizzieren im folgenden die Wirkungsweise von einigen GPSG-Komponenten.

Zu den Komponenten in GPSG zählen ID-Regeln und LP-Aussagen. Ein Lexikoneintrag läßt sich als unäre Regel auffassen, deren rechte Seite aus einem Terminalsymbol besteht. Jeder zulässige lokale Baum ist durch einen Lexikoneintrag oder eine ID-Regel projiziert (d.h. er enthält mindestens ebensoviel Information wie diese). Knotenbezeichner sind *komplexe Kategorien*, d.h. Mengen aus Merkmal-Wert-Paaren, wobei Werte wiederum komplexe Kategorien sein dürfen (auf diese Weise kann Information über andere Kategorien dargestellt werden).²² Zulässige lokale Bäume gehorchen schließlich drei universellen (d.h. einzelsprachunabhängig gültigen) *Merkmal-Instantiierungs-Prinzipien* (im folgenden kurz: Prinzipien), die die Spezifikation von bestimmten Merkmalen an bestimmten Kategorien im Baum verlangen. Es folgen Beispiele für einen Lexikoneintrag eines transitiven Verbs, (33), eine ID-Regel, die transitive Verben einführt (34) und eine LP-Aussage, die das Verb im Nebensatz ans Ende rückt (35).²³

(33) V[subcat : trans] → *verabschied*

(34) S → V[subcat : trans], NP[cas : nom], NP[cas : acc]

(35) NP V[mainclause:-]

²²Durch geeignete Beschränkungen bleibt die Menge der Kategorien endlich, was für die Kontextfreiheit von GPSG notwendig ist

²³Man beachte, daß die Kategonensymbole (S, V, NP) nicht monadisch sind wie in üblichen kontextfreien Grammatiken sondern Abkürzungen für komplexe Kategonen.

Wie üblich wird strikt zwischen dem Formalismus (Syntax und Semantik von ID, LP und Prinzipien) und der Grammatik (Mengen von ID-Regeln, LP-Aussagen und Lexikon-einträgen) unterschieden. Eine GPS-Grammatik ist deklarativ dargestellt, kann also im Prinzip bidirektional verwendet werden, und ein Generator kann mit beliebigen Grammatiken arbeiten. Ein Beispiel für einen Strukturbaum zeigt Abb. 9b.²⁴

Um den linguistischen Formalismus für Generierung und Parsing zugänglich zu machen, war eine partielle Redefinition zwingend notwendig (Näheres siehe [Dusemann, 1992, Kap. 4]). Mit dem modifizierten Formalismus stellen sich einem Generierungsverfahren zwei Aufgaben:

- Konstruktion der syntaktischen Struktur nach Maßgabe der ID-Regeln
- Spezifizierung der Struktur aufgrund der Prinzipien und LP-Aussagen

Die Konstruktion erfolgt durch sukzessives Erweitern eines partiellen Baums um einen neuen Teilbaum durch Unifikation zweier Kategorien. Die Unifikation kann die Information in beiden Teilbäumen vermehren. Da ID-Regeln in ihren Kategorien gewöhnlich stark unterspezifiziert sind, kann die so aufgebaute Struktur als Skelett betrachtet werden, zu dem das Fleisch noch hinzugefügt werden muß. Dies erfolgt durch die Prinzipien, die die Kategorien eines lokalen Baums weitgehend spezifizieren und diesen zusammen mit den LP-Aussagen zulässig im Sinne der GPSG-Theorie machen.

Ein Generierungsverfahren kann die beiden Aufgaben nach folgender Strategie ausführen: Hat ein lokaler Baum noch nicht terminierte Töchter, wird er top-down durch eine ID-Regel expandiert, wobei die Kontrolle den neu eingefügten lokalen Baum fokussiert. Hat der fokussierte lokale Baum nur terminierte Blätter, wird er durch die Prinzipien weiter spezifiziert und durch die LP-Aussagen geordnet. Die Kontrolle fokussiert danach bottom-up den nächsten lokalen Baum.

An dieser Stelle ist auf das grundsätzliche Problem der Endlosrekursion bei der Top-Down-Generierung hinzuweisen. Eine strikte Links-Rechts-Verarbeitung wird bei einer Regel wie (36) nicht terminieren, da sie rekursiv anwendbar ist.

(36) $VP \rightarrow VP, AP$

(37) $AP \rightarrow VP, AP$

Eine verfeinerte Top-Down-Strategie wird die Expansion rekursiver Töchter erst nach der Verbalisierung der übrigen vornehmen. Dies genügt, um die in der Praxis häufig auftretende lokale Rekursion abzufangen; enthält die Grammatik jedoch nichtlokale Rekursion (z.B. Regeln (36) und (37)), ist das Problem gravierend.

Mit der genannten Strategie ist keine inkrementelle Verarbeitung möglich, da LP-Aussagen erst angewandt werden können, wenn die sie betreffenden Merkmale im Baum instantiiert sind. Dies ist erst nach Abschluß des Gesamtprozesses sichergestellt. Hierin

²⁴Das Zeichen '/' kürzt in GPSG das kategorienwertige Merkmal slash ab, das hier dazu dient, Constraints für die erste Konstituente des Deklarativsatzes (die Akkusativ-NP) gemäß Eigenschaften einer vom Hauptverb subkategorisierten Konstituente zu spezifizieren. Zur exakten Wirkungsweise und zum linguistischen Hintergrund verweisen wir auf [Gazdar *et al*, 1985] und [Busemann, 1992].

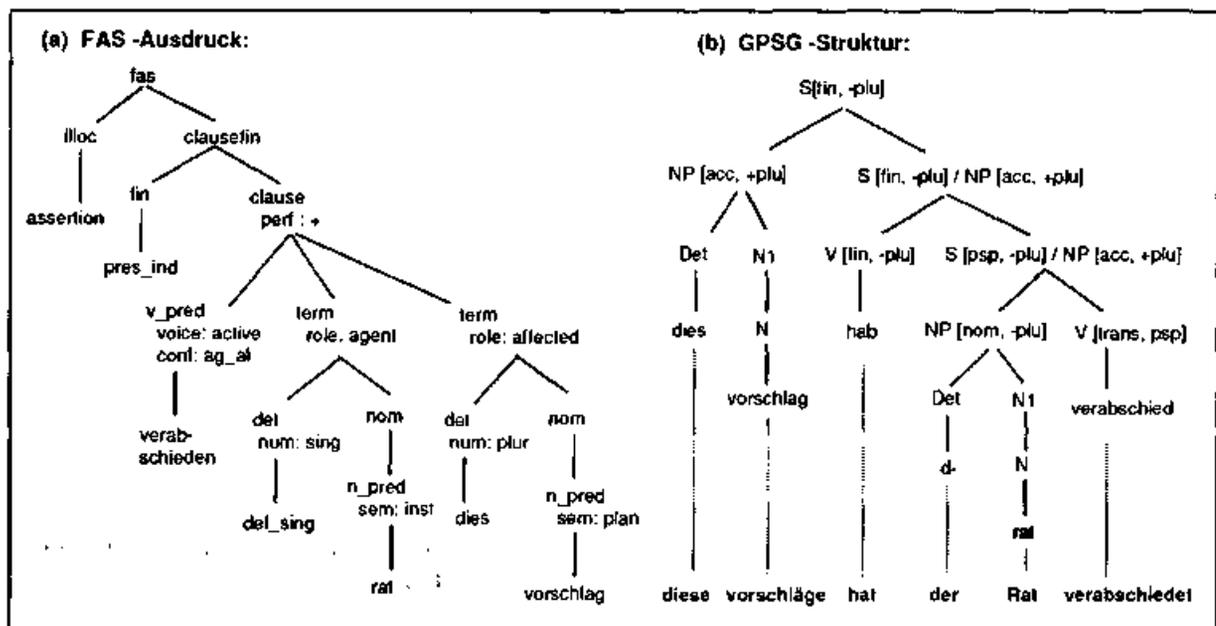


Abbildung 9: PAS-Ausdruck (a) und daraus generierte GPSG-Struktur (b).

liegt ein wichtiger Unterschied zur Verarbeitung mit TAGs, wo alle Kospezifikationen von Merkmalen in den Bäumen zur Definitionszeit festgelegt werden, was die lokale Entscheidbarkeit von LP-Aussagen garantiert.

Die getrennte Semantikepräsentation ist nicht Bestandteil der GPSG-Theorie, sondern hängt i.a. von möglichen Anwendungssituationen ab. Für das Berliner System für maschinelle Übersetzung (MÜ) wurde eine semantische Repräsentationssprache mit Namen FAS (Funktork-Argument-Strukturen) entworfen, die als (einzige) Transferebene zwischen Ausgangssprachlicher und Zielsprachlicher Satzrepräsentation fungiert und insofern der Ausgangspunkt der Generierung ist. FAS umfaßt u.a. Funktork-Argument-Beziehungen, Argumentrollen und semantische Merkmale. Ein Beispiel für eine FAS-Struktur gibt Abb. 9a. Die Syntax von FAS ist durch eine kontextfreie Grammatik mit komplexen Kategorien²⁵ definiert.

Mit vollständig spezifizierten FAS-Ausdrücken stehen dem Generator alle notwendigen Informationen zur eindeutigen Erzeugung eines Satzes zur Verfügung. Allerdings passen sie nicht direkt in den GPSG-Kontext, sondern müssen interpretiert und „übersetzt“ werden. Merkmale für den Satztyp (Hauptsatz, Deklarativsatz), Oberflächenkasus oder Numerus müssen noch berechnet und Hilfsverben auf der Grundlage von FAS-Merkmalsspezifikationen als GPSG-Strukturteile eingeführt werden.

Die Abbildung von lokalen FAS-Bäumen in GPSG-Strukturen erfolgt durch eine fest vorgegebene Traversalion des FAS-Ausdrucks (top-down, Tiefe zuerst, von links nach rechts), wobei zum einen die benötigte Information gesammelt und übersetzt wird und zum ändern

²⁵FAS-Kategorien sind analog zu Prolog-Tennen aufgebaut, unterscheiden sich also erheblich von der Syntax der GPSG-Kategorien.

die syntaktische Struktur nach Maßgabe der ID-Regeln aufgebaut -wird. Die Konstruktion der syntaktischen Struktur terminiert auch bei rekursiven Regeln, da sie durch den (endlichen) FAS-Ausdruck gesteuert wird. Die Abbildung beruht auf Pattern-Action-Regeln, die separat definiert sind als Tripel, bestehend aus einem *FAS-Pattern* der Tiefe 1, einer Menge von *informationsgenerierenden Aktionen* und einer Menge von *strukturbildenden Aktionen*.

Informationsgenerierende Aktionen benutzen einen Zwischenspeicher, um zeitweise FAS-Information abzuspeichern und diese zu einem späteren Zeitpunkt in die GPSG-Struktur einzubringen. So ist z.B. die Information, aus der der Oberflächenkasus berechnet wird, am Lexikoneintrag des Verbs zugänglich, das für bestimmte kasustragende Konstituenten subkategorisiert. Diese Konstituenten werden jedoch erst generiert, wenn die betreffenden FAS-Teilausdrücke bei der Traversierung erreicht sind.

Strukturbildende Aktionen wählen eine von sehr wenigen in der Pattern-Action-Regel aufgeführten ID-Regeln. Somit werden nur die lokal relevanten ID-Regeln betrachtet, was die Anzahl der auszuprobierenden ID-Regeln maßgerecht und drastisch beschränkt. Der syntaktische Teilbaum wird von der strukturbildende Aktion in die GPSG-Struktur eingefügt.

Das Beispiel (38) zeigt eine Pattern-Action-Regel (in Prolog-Notation) mit einem Pattern, das auf lexikalische lokale FAS-Bäume paßt, die eine Konfiguration semantischer Rollen des Typs [agent, affected] aufweisen und aktives Verbgenus haben. In diesem Fall soll die Agens-Rolle durch den Kasus Nominativ realisiert werden und die Affected-Rolle durch Akkusativ. Dieses wird mithilfe der informationsgenerierenden Aktion `put_store` zwischengespeichert. Schließlich wird ein GPSG-Lexikoneintrag erzeugt und in die GPSG-Struktur eingefügt.

```
(38) pa_rule( [v_pred(conf:ag-af,voice:active),[verabschieden]],
              [put_store(agent,nom),put_store(affected,acc)] ,
              [call_lex("V(subcat:trans) --> verabschied")]) •
```

In (39) wird auf die zwischengespeicherte Information wieder zugegriffen (`remove_store` löscht sie gleichzeitig). Das Pattern paßt, auf einen term mit einer semantischen Rolle, die an die Variable `Role` gebunden wird, und einer Anzahl von Töchtern, deren erste ein definitiver Determiner im Plural ist. Die erste informationsgenerierende Aktion holt die zur Rolle entsprechende Kasusinformation aus dem Zwischenspeicher und bindet sie an die Variable `Case`. Die informationsgenerierende Aktion `set_gpsg_features` übersetzt Kasus- und Numerusinformation in eine GPSG-Kategorie, die durch die strukturbildende Aktion in die GPSG-Struktur eingebracht wird. Die strukturbildende Aktion ruft eine ID-Regel auf, die eine NP erzeugt und führt den lokalen GPSG-Baum in die syntaktische Struktur ein.

```
(39) pa_rule([term(role:Role),[det(def:+,numrplur)|J] ,
              [remove_store(Role, Case),
               set_gpsg_features( [plu,cas] , [+ ,Case] )] ,
              [call_id("NP --> Det, N1")]).
```

Offenbar setzt dieses Verfahren eine strikte Reihenfolge voraus, in der die lokalen FAS-Bäume verarbeitet werden. In FAS stehen Funktoreri (z.B. v_preds) stets links von ihren Argumenten (z.B. $term$). Daher ist die Links-Rechts-Traversierung des FAS-Ausdrucks geeignet, diese Voraussetzung zu schaffen.

Dieser Ansatz kann auf andere Semantikdarstellungen übertragen werden. Zu berücksichtigen ist, daß jede Änderung der ID-Regeln in der GPS-Grammatik eine Anpassung der Pattern-Action-Regeln erfordert, und ebenso jede Änderung an FAS. Das Verfahren ist detailliert in [Busemann, 1992] beschrieben.

4.3 Lexikalisch-Funktionale Grammatik ()

LFG gehört zu den ersten lexikonzentrierten Grammatiktheorien. NL Sätze werden durch eine kontextfreie Grammatik beschrieben (c-Struktur), deren Kategorien mit einer Menge von Gleichungen annotiert sind. Diese definieren eine Abbildung auf eine funktionale Ebene (f-Struktur), die Funktor-Argument-Beziehungen expliziert.

Oftmals erweist sich aufgrund von Verarbeitungsgesichtspunkten eine Reformulierung des verwendeten Formalismus als unumgänglich. Für LFG in der Fassung von [Kaplan und Bresnan, 1982] stellt sich als theoriespezifisches Problem, aus einer f-Struktur, die in der Form eines gerichteten, azyklischen Graphen (DAG) u.a. relationale syntaktische und semantische Information enthält, einen syntaktischen Konstituentenstrukturbaum (c-Struktur) zu erzeugen, der die Endkette beschreibt. Da die Theorie der LFG keine Abbildung von f-Strukturen in c-Strukturen vorsieht, sondern nur von c-Strukturen in f-Strukturen, erscheint nach [Block, 1987] LFG besser für Parsing geeignet als für Generierung: Um eine Endkette zu generieren, muß eine c-Struktur erst erzeugt und danach getestet werden, ob sie mit der Ausgangs-f-Struktur verträglich ist. Ein solches Verfahren ist zwar theoriekonform, aber grundsätzlich ineffizient, da zum einen die Generiere-und-Teste-Strategie unvermeidbar ist und zum ändern nicht zu jeder c-Struktur eine f-Struktur existieren muß. Block kommt zu dem Schluß, daß die in [Kaplan und Bresnan, 1982] beschriebene Theorie der LFG für Generierung inadäquat sei.

[Wedekind, 1986] gibt ein Ableitungskonzept für LFG an, das die simultane Beschreibung von c- und f-Strukturen ermöglicht und insofern neutral bezüglich der Verarbeitungsrichtung ist. Eine Ableitung ist bei Wedekind eine Folge von Tripeln bestehend aus einer partiellen c-Struktur, einer partiellen f-Struktur und einer Abbildung von Knoten der c-Struktur in Teil-DAGs der f-Struktur (vgl. Abb. 10). Die c-Struktur des Starttripels besteht aus dem Knoten mit Etikett S; die c-Struktur des letzten Tripels einer Ableitung beschreibt die Endkette.

Auf dieser Grundlage können aus f-Strukturen durch Unifikation deterministisch Endketten generiert werden. Die Grundidee (vgl. [Momma und Dörre, 1988; Wedekind, 1988]) besteht in der simultanen Erzeugung von l- und c-Strukturen auf der Grundlage von Regeln, die Korrespondenzen zwischen beiden Strukturen definieren. Um eine effiziente Kontrollstrategie zu verwirklichen, muß die Konstruktion der c-Struktur durch die Ausgangs-f-Struktur bestimmt werden. Ziel des Generierungsverfahrens ist die Rekonstruktion der Ausgangs-f-Struktur, womit quasi als „Seiteneffekt“ eine c-Struktur (und eine Endkette)

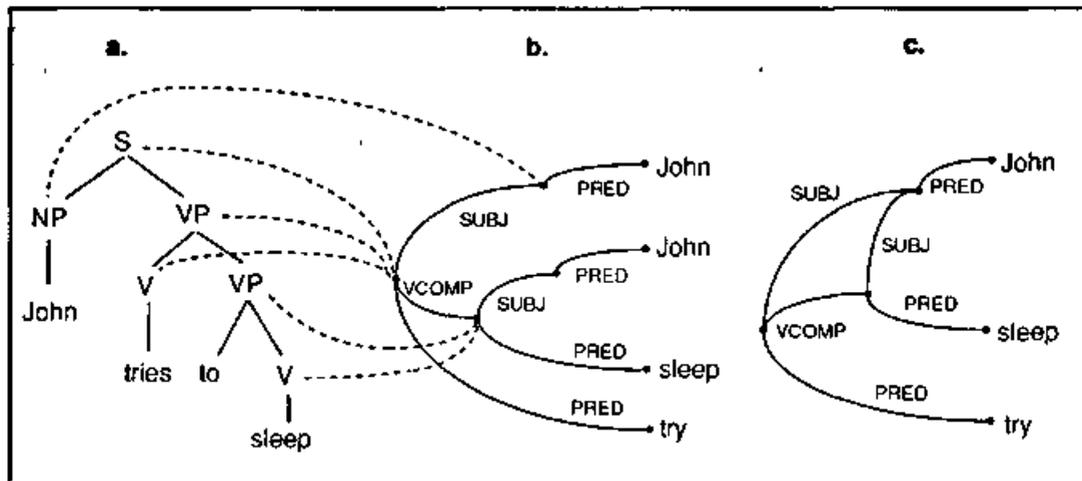


Abbildung 10: c-Struktur (a), entfaltete f-Struktur (b) und Ausgangs-f-Struktur (c)

generiert wird. Bei der Rekonstruktion ist darauf zu achten, daß

- alle Information in der Ausgangs-f-Struktur erzeugt wird („completeness“) und
- keine weitere Information hinzukommt („coherence“).

Die erste Bedingung ist erfüllt, wenn alle mehrfachen Dominanzen von Teil-DAGs („reentrancies“) und ferner alle Pfade der Ausgangs-f-Struktur abgeleitet werden. Dazu wird der DAG der Ausgangs-f-Struktur entfaltet, d.h. in einen Baum transformiert, wobei alle mehrfachen Dominanzen durch gleiche Teilstrukturen ersetzt werden (Abb. 10b). Ferner werden alle bereits abgeleiteten Pfade dieser Struktur nach Maßgabe der Regeln markiert. Die zweite Bedingung wird erfüllt, wenn die Ausgangs-f-Struktur die rekonstruierte f-Struktur subsumiert.

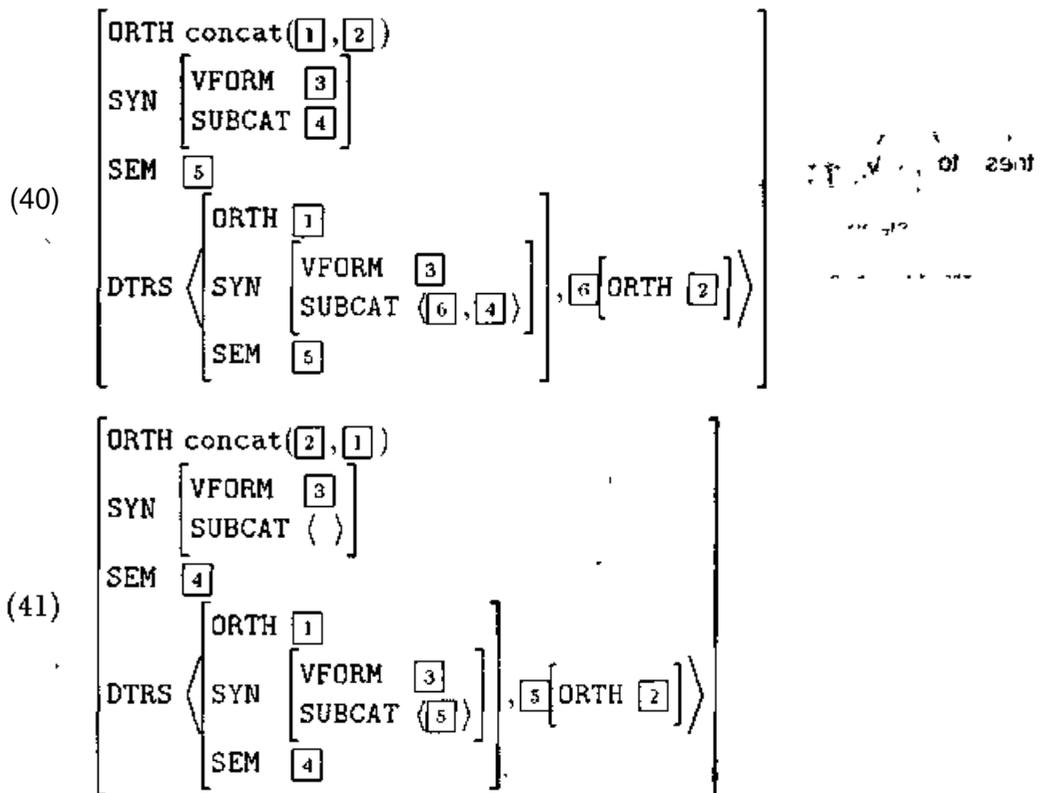
Die Ableitung der c-Struktur erfolgt dem Ableitungskonzept zufolge top-down. Da das Verfahren strukturgetrieben arbeitet und die Ausgangs-f-Struktur zusammenhängend rekonstruiert wird, ist sichergestellt, daß keine Endlosrekursion auftritt.

Das LFG-System ist modular in dem Sinne, daß mit verschiedenen Grammatiken generiert werden kann ([Mammaund Dörre, 1988] nennen Grammatiken für deutsch, französisch und englisch). Ursprünglich wurde nicht angestrebt, ein- und dieselbe Grammatik für Generierung und Parsing zu verwenden, aber doch einen gemeinsamen Kern an Regeln.

4.4 Steuerung durch den „semantischen Kopf“

In lexikonzentrierten linguistischen Theorien kommen in jedem Lexikoneintrag und jeder Grammatikregel komplexe Kategorien mit orthographischer, syntaktischer und semantischer Information vor, die mittels Koreferenz zueinander in Beziehung stehen. Koreferenz zweier Strukturen ist in den folgenden Beispielen durch Gleichheit der in Kästchen geschriebenen Zahlen dargestellt.

Das subcat-Merkmal ist, anders als in GPSG, eine Liste. Auf der tiefsten Ebene im Strukturbaum, der lexikalischen, enthält sie Kategorien für alle Komplemente („ungesättigter“ Zustand), während auf jeder höheren Ebene, auf der ein Komplement eingeführt wird (in Regeln, vgl. (40) und (41)), die betreffende Kategorie aus der Liste entfernt ist. An der höchsten Kategorie im Strukturbaum ist die Liste leer („gesättigter“ Zustand).



Im Unterschied zu Lexikoneinträgen haben Regeln ein DTRS-Attribut, in dem die Töchter gelistet sind. In Regeln wird im wesentlichen festgelegt, wie die subkategorisierten Elemente als Konstituenten auftreten. Fast alle Information über die Komplemente stammt aus den lexikalischen Kategorien. Daher können die Regeln so allgemein formuliert werden. Regel (40) erwartet eine Tochter mit nichtleerer subcat-Liste, deren erstes Element koreferent ist zur zweiten Tochter der Regel (Index [5]). Der Funktionsaufruf von concat unter ORTH bewirkt, daß die Endkette der ersten Tochter vor der der zweiten geäußert wird. Diese Regel dient etwa zur Realisierung von Verb und Objekt. Regel (41) erwartet eine Tochter mit einelementiger subcat-Liste, deren Element zur zweiten Tochter koreferent ist. Das ORTH-Merkmal fordert hier jedoch eine andere Wortstellung. Diese Regel wird z.B. benutzt, um Subjekts-NP und VP zu erzeugen.

Anhand dieser beiden Regeln kann mit den Lexikoneinträgen in (42) und (43) durch Unifikation die Ableitung für den Satz *Peter helps Mary* in (44) konstruiert werden.

$$(42) \left[\begin{array}{l} \text{ORTH } \langle \textit{helps} \rangle \\ \text{SYN} \left[\begin{array}{l} \text{VFORM } \textit{fin} \\ \text{SUBCAT } \langle \text{SEM } \boxed{2} \rangle, \left[\begin{array}{l} \text{SYN } \left[\begin{array}{l} \text{PER } 3 \\ \text{NUM } \textit{sg} \end{array} \right] \\ \text{SEM } \boxed{1} \end{array} \right] \rangle \end{array} \right] \\ \text{SEM } \textit{help}(\boxed{1}, \boxed{2}) \end{array} \right]$$

$$(43) \text{ a. } \left[\begin{array}{l} \text{ORTH } \langle \textit{peter} \rangle \\ \text{SYN } \left[\begin{array}{l} \text{PER } 3 \\ \text{NUM } \textit{sg} \end{array} \right] \\ \text{SEM } \textit{p} \end{array} \right] \quad \text{b. } \left[\begin{array}{l} \text{ORTH } \langle \textit{mary} \rangle \\ \text{SYN } \left[\begin{array}{l} \text{PER } 3 \\ \text{NUM } \textit{sg} \end{array} \right] \\ \text{SEM } \textit{m} \end{array} \right]$$

$$(44) \left[\begin{array}{l} \text{ORTH } \langle \textit{peter}, \textit{helps}, \textit{mary} \rangle \\ \text{SYN } \left[\text{VFORM } \textit{fin} \right] \\ \text{SEM } \textit{help}(\textit{p}, \textit{m}) \\ \text{DTRS} \left\langle \begin{array}{l} \left[\begin{array}{l} \text{ORTH } \langle \textit{helps}, \textit{mary} \rangle \\ \text{SYN } \left[\begin{array}{l} \text{VFORM } \textit{fin} \\ \text{SUBCAT } \langle \boxed{1} \rangle \end{array} \right] \\ \text{SEM } \textit{help}(\textit{p}, \textit{m}) \end{array} \right] \\ \left[\begin{array}{l} \text{ORTH } \langle \textit{helps} \rangle \\ \text{SYN } \left[\begin{array}{l} \text{VFORM } \textit{fin} \\ \text{SUBCAT } \langle \boxed{2}, \boxed{1} \rangle \end{array} \right] \\ \text{SEM } \textit{help}(\textit{p}, \textit{m}) \end{array} \right] \end{array} \right\rangle, \left[\begin{array}{l} \text{ORTH } \langle \textit{mary} \rangle \\ \text{SYN } \left[\begin{array}{l} \text{PER } 3 \\ \text{NUM } \textit{sg} \end{array} \right] \\ \text{SEM } \textit{m} \end{array} \right] \right\rangle, \boxed{1} \left[\begin{array}{l} \text{ORTH } \langle \textit{peter} \rangle \\ \text{SYN } \left[\begin{array}{l} \text{PER } 3 \\ \text{NUM } \textit{sg} \end{array} \right] \\ \text{SEM } \textit{p} \end{array} \right] \right\rangle \end{array} \right]$$

Das listenbasierte Subkategorisierungskonzept macht eine strikte Top-Down-Verarbeitung schwierig. Man müßte mit einer leeren Liste beginnen, deren endgültige Länge zunächst nicht bekannt ist. Bei Einführung von Konstituenten durch die Regeln werden entsprechende Kategorien an die Liste angehängt, bis der lexikalische Kopf eingeführt wird. Dann erst stellt sich heraus, ob die eingeführten Konstituenten in den Subkategorisierungsrahmen des lexikalischen Kopfes passen. Dieses Verfahren ist auf Versuch und Irrtum gegründet und daher ineffizient.

Das listenbasierte Subkategorisierungskonzept erlaubt durchaus eine bottom-up vorgehende Generierungsstrategie, die von der Subkategorisierungsspezifikation des lexikalischen Kopfes ausgeht und kontrolliert, welche Konstituente auf der nächsthöheren Ebene der syntaktischen Struktur eingeführt werden darf. Dabei wird jedoch nicht die Semantik in der vorgegebenen Ausgangsstruktur berücksichtigt, so daß bei einem schließlich fälligen Abgleich Indeterminismus auftritt.

Diesem Mangel kann man durch den Begriff des *semantischen Kopfes* abhelfen. Auf ihm beruht die sogenannte „semantic-head-driven strategy“ [Shieber *et al.*, 1990], die für den integrierten Semantikansatz als die fortgeschrittenste gilt. Sie zielt darauf ab, daß vom aktuellen Startknoten ein Pfad im Analysebaum zu einem Knoten K existiert derart, daß alle

Knoten auf dem Pfad mit derselben semantischen Information etikettiert sind, nicht aber die Töchter von K. K wird als „semantischer Kopf“ bezeichnet. Von ihm ausgehend, werden sowohl Top-Down-Schritte als auch Bottom-Up-Schritte vorgenommen. Zunächst werden seine Töchter rekursiv top-down expandiert; danach wird bottom-up entlang des Pfades durch Einführung von Grammatikregeln die Reduktion zur Wurzel hin vorgenommen, wobei auf jeder Ebene wiederum die restlichen Töchter der Regeln top-down expandiert werden.

Der Algorithmus beruht auf einer Unterteilung der Grammatikregeln in zwei Mengen. Die sogenannten *Kettenregeln* zeichnen sich durch gleiche semantische Information an der Mutter und mindestens einer Tochter aus. Bei den restlichen Regeln und insbesondere bei den Lexikoneinträgen ist das nicht der Fall. Bei den Regeln auf S. 48 handelt es sich um Kettenregeln, während der Lexikoneintrag (42) keine Kettenregel ist.

Ein Bottom-Up-Schritt erfolgt ausschließlich aufgrund einer Kettenregel, während ein Top-Down-Expansionsschritt mit einer der übrigen Regeln bewerkstelligt wird.²⁶ Anhand der Ausgangsstruktur (45) bestimmt sich der Pfad zum semantischen Kopf aufgrund der Kategorien in den Kettenregeln. Ein semantischer Kopf mit gleicher semantischer Information ist die Mutter in (42). Von diesem Knoten ausgehend wird zuerst top-down das Terminalsymbol *helps* generiert. Dann wird bottom-up die Regel (40) angewandt. Die zweite Tochter wird top-down als Objekt generiert. Danach folgt die Reduktion zur Wurzelkategorie mit Regel (41), und deren restliche Tochter wird als Subjekt generiert. Als Resultat erhält man den Satz *Peter helps, Mary* (vgl. auch (44)).

$$(45) \left[\begin{array}{l} \text{SYN} \left[\text{SUBCAT} \{ \} \right] \\ \text{SEM } \text{help}(p, m) \end{array} \right]$$

Damit dieses Verfahren möglichst deterministisch arbeitet, kommt es darauf an, semantische Köpfe eindeutig zu identifizieren. Im vorliegenden Fall wird von einem Vollformenlexikon ausgegangen, so daß noch weitere Formen von *help* infrage kamen. Bei einer „falschen“ Wahl würde Backtracking erst bei der Generierung der Subjekts-NP einsetzen, da *Peter* lexikalisch als NP in der dritten Person Singular markiert ist. [Shieber *et al.*, 1990] merken an, daß dieses Problem durch die Verwendung eines Stammformenlexikons und einer getrennt operierenden Flexionskomponente gelöst werden kann.

Eine weitere Voraussetzung für Determinismus ist ein möglichst beschränktes Spektrum an Reduktionsmöglichkeiten. Die Gefahr von Sackgassen, aus denen keine weitere Reduktionsmöglichkeit mehr zur Wurzel führt, kann durch geeignete Kodierung der Grammatiken in der Praxis vermieden oder beschränkt werden. Um diese Quelle des Nondeterminismus grundsätzlich auszuschließen, muß das Verfahren vor jedem Reduktionsschritt prüfen, ob es anschließend noch einen Pfad zur Wurzel gibt.

²⁶Bei einer Kettenregel wäre die Expansionsstelle per definitionem nicht der semantische Kopf.

4.5 Bidirektionale Verarbeitung

Wie zu Beginn des Abschnitts 4 erwähnt wurde, unterscheidet man die Reversibilitätseigenschaft von Grammatiken von der Bidirektionalität der Verarbeitungsprozesse. Im einzelnen nennt [Neumann, 1991] vier Stufen von Bidirektionalität, die ein NL System aufweisen kann (siehe auch [Reithinger, 1992, S. 43f]):

1. Aus einer gemeinsamen Wissensquelle werden verarbeitungsspezifische Repräsentationen kompiliert. Die Prozesse sind durch unterschiedliche Verfahren implementiert.
2. Die Prozesse verwenden eine gemeinsame Wissensquelle und sind durch unterschiedliche Verfahren implementiert.
3. Die Prozesse verwenden eine gemeinsame Wissensquelle und gemeinsame Basisoperationen. Sie sind durch unterschiedliche Verfahren implementiert.
4. Die Prozesse verwenden eine gemeinsame Wissensquelle und sind durch ein einziges, je nach Verarbeitungsrichtung parametrisierbares Verfahren implementiert.

Es ist zu beachten, daß diese Beschreibungen auf verschiedenartige Wissensquellen und Prozesse eines (komplexen) NL Systems verweisen können. Wir beschränken uns hier auf die Grammatik und die sie betreffenden Prozesse, d.h. aus Generierungssicht auf die Systemleistungen, die wir unter Realisierung subsumieren. Ein weitergefaßter, aktueller Überblick findet sich in [Strzalkowski, 1992].

Als Beispiel für die erste Stufe dient die Kompilation von LFG-ähnlichen Grammatiken in spezielle Parsing- bzw. Generierungsgrammatiken, wie z.B. von [Horacek und Pyka, 1988] vorgeschlagen. Ohne Kompilation der Grammatik verfährt das Berliner MÜ-System (vgl. Abschnitt 4.2). Da die Prozesse für die GPSG-Prinzipien einen gemeinsamen Teilprozeß verwenden, ist das System bidirektional gemäß der dritten Stufe.

Wir wenden uns nun der Frage zu, wie Parsing- und Generierungsverfahren in einer einheitlichen Systemarchitektur als Ausprägungen einer einzigen Verarbeitungsstrategie betrachtet werden können (Bidirektionalität der vierten Stufe). Die an der Universität Stuttgart entwickelte Sprache TFS (für Typed Feature Structures) erlaubt die Definition getypter Grammatiken (insbesondere von HPSG) [Emele und Zajac, 1990]. Eine Merkmalsstruktur vom Typ *sign-type* etwa ist durch eine bestimmte Menge von Attributen definiert, deren Werte wiederum getypte Merkmalsstrukturen sind (vgl. (46)). Ein speziellerer Typ wie *phrasal-sign-type* nutzt Vererbung im Typverband aus, indem die Definition von *sign-type* mit den lokalen Attributen unifizieren muß.

$$(46) \quad \textit{sign-type} \equiv \left[\begin{array}{l} \text{ORTH } \textit{orth-type} [] \\ \text{SYN } \textit{syntax-type} [] \\ \text{SEM } \textit{semantics-type} [] \end{array} \right]$$

$$(47) \quad \textit{phrasal-sign-type} \equiv \textit{sign-type} \left[\text{DTRS } \textit{daughters-type} [] \right]$$

In TFS ist als einzige Verarbeitungsstrategie die vollständige Typexpansion durch Unifikation definiert. Wenn als Eingabe eine Merkmalsstruktur vom Typ *phrasal-sign-type* vorliegt, werden sämtliche uninstantiierten Attribute durch Typexpansion ergänzt. Ist in der Eingabe ORTH spezifiziert, nicht aber SYN und SEM, so handelt es sich um eine Parsingaufgabe. Ist andererseits in der Eingabe SEM spezifiziert, nicht aber SYN und ORTH, so erfolgt Generierung. In jedem Fall wird die gleiche, vollständig spezifizierte Merkmalsstruktur erzeugt. Für das System ist folglich irrelevant, ob es parst oder generiert.

Wie viele Systeme, die moderne Formalismen verwenden, geht auch TFS davon aus, daß Grammatikwissen unabhängig vom Diskurskontext einer Äußerung verarbeitet werden kann. Die Adäquatheit dieser Annahme für diskursbasierte NL Systeme stellt [Appelt, 1987] infrage. [Neumann und van Noord, 1992] schlagen vor, Parsing und Generierung als zwei ineinander verzahnte Aufgaben zu betrachten, die auf einer einzigen parametrisierbaren Deduktionsmaschine basieren. Die Verzahnung ermöglicht es, den Parser zur Unterstützung der Generierung heranzuziehen und umgekehrt. Dieser Vorschlag wird gegenwärtig im Projekt „Bidirektionale Linguistische Deduktion“ an der Universität des Saarlandes ausgearbeitet.

■

Literatur

- [Allen, 1983] J. F. Allen. Recognizing intentions from natural language utterances. In M. Brady und R. C. Berwick (Hg.), *Computational Models of Discourse*, Kapitel 2, Seite 107-164. MIT Press, Cambridge, Mass., 1983.
- [Appelt, 1985] D. E. Appelt. *Planning English Sentences*. Cambridge University Press, Cambridge, 1985.
- [Appelt, 1987] D. E. Appelt. Bidirectional Grammars and the Design of Natural Language Generation Systems. In Y. Wilks (Hg.), *Theoretical Issues in Natural Language Processing*, Seite 185-191. New Mexico State University, Las Cruces, NM., January 7-9 1987.
- [Bateman, 1990] J. A. Bateman. Upper Modeling: organizing knowledge for natural language processing. In *Proceedings of the 5th International Workshop on Language Generation*, Pittsburgh, PA. 1990.
- [Block, 1987] R. Block. Can a 'parsing grammar' be used for natural language generation? The negative example of LEG. In *Proceedings of the First European Natural Language Generation Workshop*, Abbey de Royaumont. January 23-26 1987.
- [Brown und Dell, 1987] P. Brown und G. S. Dell. Adapting production to comprehension: The explicit mention of Instruments. *Cognition Psychology*, (19):441-472, 1987.
- [Busemann, 1988] Stephan Busemann. Surface Transformations During the Generation of Written German Sentences. In D. D. McDonald und L. Bolc (Hg.), *Natural Lan-*

- guage Generation Systems*, Seite 98-165. Springer, Berlin, New York, 1988. Symbolic Computation.
- [Busemann, 1992] Stephan Busemann. *Generierung natürlicher Sprache mit Generalisierten Phrasenstrukturgrammatiken*. Springer, Berlin, Heidelberg, 1992. IFB Bd. 313.
- [Carberry, 1983] S. Carberry. Tracking User Goals in an Information-Seeking Environment. In *Proceedings of the 3rd AAI*, Washington. 1983.
- [Cohen *et al.*, 1990] P. R. Cohen, J. L. Morgan und M. E. Pollack (Hg.). *Intentions in Communication*. A Bradford Book, MIT Press, System Development Foundation Benchmark Series, Cambridge, MA., 1990.
- [Dale *et al.*, 1990] Robert Dale, Chris Mellish und Michael Zock (Hg.). *Current Research in Natural Language Generation*. Academic Press, London, 1990.
- [Dale *et al.*, 1991] Robert Dale, Eduard Hovy, Dietmar Rosner und Oliviero Stock (Hg.). *Aspects of Automated Natural Language Generation. Proc. 6th International Workshop on Natural Language Generation, Trento, Italy, April 1992*. Springer-Verlag, Berlin, Heidelberg, New York, 1991. Lecture Notes in Artificial Intelligence No. 587.
- [Danlos, 1987] L. Danlos. *The Linguistic Basis of Text Generation*. Cambridge University Press, Cambridge, 1987.
- [De Smedt, 1990] K. J. M. J. De Smedt. *Incremental Sentence Generation*. Doktorarbeit, Katholieke Universiteit Nijmegen, Nijmegen Institute for Cognition Research and Information Technology, 1990. Report 90-01.
- [Deutsch, 1976] W. Deutsch. *Sprachliche Redundanz und Objekt Identifikation*. Doktorarbeit, Universität Marburg, 1976.
- [Dobes und Novak, 1991] Z. Dobes und H.-J. Novak. From Knowledge Structures to Text Structures. In O. Herzog und C.-R. Rollinger (Hg.), *Text Understanding in LILOG*, Seite 670-684. Springer, Lecture Notes in Artificial Intelligence, 546, 1991.
- [Emele und Zajac, 1990] Martin Emele und Remi Zajac. Typed Unification Grammars. In *Proc. 13th COLING-90*, Helsinki, Seite 293-298. 1990.
- [Emele, 1987] Martin Emele. FREGE-Ein objektorientierter FRont-End-GENERator. In K. Morik (Hg.), *GWAI-87. 11th German Workshop on Artificial Intelligence*, Berlin, New York, Seite 64-73. Springer, 1987. IFB Bd. 152.
- [Fay, 1980] D. Fay. Transformational errors. In V. A. Fromkin (Hg.), *Perception and production of fluent speech: Slips of the tongue, ear, pen, and hand*. Academic Press, New York, 1980.
- [Fromkin, 1980] V. A. Fromkin (Hg.). *Errors in linguistic performance. Slips of the tongue, ear, pen, and hand*. Academic Press, New York, 1980.

- [Garret, 1976] M. F. Garret. Syntactic Processes in Sentence Production. In R. J. Wales und E. Walker (Hg.), *New Approaches to Language Mechanisms*. North Holland, Amsterdam, 1976.
- [Gazdar et al., 1985] G. Gazdar, E. Klein, G. Pullum und I. Sag. *Generalized Phrase Structure Grammar*. Basil Blackwell, London, 1985.
- [Goldman, 1975] N. M. Goldman. Conceptual Generation. In R. C. Schank (Hg.), *Conceptual Information Processing*. North Holland, Amsterdam, 1975.
- [Good und Butterworth, 1980] D. A. Good und B. Butterworth. Hesitancy as a conversational resource: Some methodological implications. In H. W. Dechert und M. Raupach (Hg.), *Temporal variables in speech*. Mouton, The Hague, 1980.
- [Grice, 1975] H.P. Grice. Logic and Conversation. In Peter Cole und Jerry L. Morgan (Hg.), *Syntax and Semantics, Vol. 3: Speech Acts*, Seite 41-58. Academic Press, San Diego, New York, Berkeley, 1975.
- [Halliday, 1985] M. A. K. Halliday. *An Introduction to Functional Grammar*. Edward Arnold, London, 1985.
- [Harbusch et al., 1991] Karin Harbusch, Wolfgang Finkler und Anne Schauder. Incremental Syntax Generation with tree Adjoining Grammars. Research Report RR-91-25, DFKI, Kaiserslautern, Saarbrücken, Oktober 1991.
- [Herrmann, 1983] T. Herrmann. *Speech and Situation: A psychological conception of situated speaking*. Springer, Heidelberg, 1983.
- [Hoepfner et al., 1984] W. Hoepfner, H. Marburger und K. Morik. Talking it Over: The Natural Language Dialog System HAM-ANS. In Leonard Bolc (Hg.), *Cooperative Interactive Systems*. Springer, Berlin, New York, 1984. Symbolic Computation.
- [Horacek und Pyka, 1988] Helmut Horacek und Claudius Pyka. Facets of Knowledge about Natural Language Syntax: Representation and Use in Parsing and Generation. In W. Hoepfner (Hg.), *Künstliche Intelligenz. Proc. GWAI-88, 12. Jahrestagung*, Berlin, New York, Seite 130-139. Springer, 1988. IFB Bd. 181.
- [Horacek und Zock, 1992] Helmut Horacek und Michael Zock (Hg.). *New Concepts in Natural Language Generation: Planning, Realization, and Systems*. Frances Pinter, London, 1992.
- [Hovy, 1988a] E. H. Hovy. Generating language using a phrasal lexicon. In D. D. McDonald und L. Bolc (Hg.), *Natural Language Generation Systems*, Kapitel 10, Seite 353-384. Springer-Verlag, Berlin, 1988.
- [Hovy, 1988b] E. H. Hovy. *Generating Natural Language under Pragmatic Constraints*. Lawrence Erlbaum, Hillsdale, NJ., 1988. n* •

- [Hovy, 1990a] E. H. Hovy. Parsimonious and Profligate Approaches to the Question of Discourse Structure Relations. In *Proc. of the 5th International Workshop on Language Generation*, Seite 128-136, 1990.
- [Hovy, 1990b] E. H. Hovy. Unresolved Issues in Paragraph Planning. In R. Dale, C. Mellish und M. Zock (Hg.), *Current Research in Natural Language Generation*, Kapitel 2, Seite 17-46. Academic, London, 1990.
- [Iordanskaja et al., 1991] Lidija Iordanskaja, Richard Kittredge und Alain Polguere. Lexical selection and paraphrase in Meaning-text Theory. In Cecile Paris, Swartout William und William Mann (Hg.), *Natural Language in Artificial Intelligence and Computational Linguistics*, Dordrecht, Seite 292-312. Kluwer, 1991.
- [Jacobs, 1987] P.S. Jacobs. KING: a knowledge-intensive natural language generator. In G. Kempen (Hg.), *Natural Language Generation*, Seite 219-230. Martinus Nijhoff Publishers, Dordrecht, Boston, Lancaster, 1987. NATO ASI Series E Nr. 135.
- [Joshi, 1986] A. K. Joshi. An Introduction to Tree Adjoining Grammars. Technical Report No. MS-CIS-86-64, Department of Computer and Information Science, University of Pennsylvania, August 1986.
- [Kaplan und Bresnan, 1982] Ronald Kaplan und Joan Bresnan. Lexical Functional Grammar: A Formal System for Grammatical Representation. In Joan Bresnan (Hg.), *The Mental Representation of Grammatical Relations*, Kapitel 4, Seite 173-281. MIT Press, Cambridge, MA., 1982.
- [Kaplan, 1983] R. M. Kaplan. Co-operative response from a portable natural language System. In M. Brady und R. C. Berwick (Hg.), *Computational Models of Discourse*, Seite 167-208. MIT Press, Cambridge, Mass., 1983.
- [Kempen und Hoenkamp, 1987] G. Kempen und E. Hoenkamp. An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11:201-258, 1987.
- [Kempen, 1978] G. Kempen. Sentence construction by a psychologically plausible formula-tor. In R. Campbell und P. Smith (Hg.), *Recent Advances in the Psychology of Language. Vol. 2: Formal and experimental approaches*. Plenum Press, New York, 1978.
- [Kempen, 1987a] G. Kempen. A framework for incremental syntactic tree formation. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy, Seite 655-660. August 23-28 1987.
- [Kempen, 1987b] G. Kempen (Hg.). *Natural Language Generation*. Martinus Nijhoff Publishers, Dordrecht, Boston, Lancaster, 1987. NATO ASI Series E Nr. 135.
- [Klose et al., 1992] Gudrun Klose, Ewald Lang und Thomas Pirlein (Hg.). *Ontologie und Axiomatik der Wissensbasis von LILOG*. Springer, Berlin, Heidelberg, 1992. IFB Bd. 307.

- [Lang *et al.*, 1991] E. Lang, K.-U. Carstensen und Geoff Simmons. *Modelling Spatial Knowledge on a Linguistic Basis*. Springer, Berlin, Heidelberg, 1991. Lecture Notes in Artificial Intelligence No. 481.
- [Levelt, 1982] W. J. M. Levelt. Linearization in describing spatial networks. In S. Peters und E. Saarinen (Hg.), *Processes, beliefs, and questions*. Reidel, Dordrecht, 1982.
- [Levelt, 1989] W. J. M. Levelt. *Speaking: From Intention to Articulation*. A Bradford Book, MIT Press, Cambridge, MA., 1989.
- [Mangold, 1986] R. Mangold. *Sensorische Faktoren beim Verstehen überspezifizierter Objektbenennungen*. Peter Lang, Frankfurt, 1986.
- [Mann *et al.*, 1982] W. C. Mann, M. Bates, B. J. Grosz, D. D. McDonald, K. R. McKeown und W. Swartout. Text Generation: The State of the art and the literature. *American Journal of Computational Linguistics*, 8:62-69, 1982.
- [Mann und Moore, 1982] W. C. Mann und J. Moore. Computer generation of multiparagraph English text. *American Journal of Computational Linguistics*, 7:17-29, 1982.
- [Mann und Thompson, 1988] W.C. Mann und S.A. Thompson. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8(3):243-281, 1988.
- [Mann, 1988] W. C. Mann. Text Generation: The Problem of Text Structure. In D. D. McDonald und L. Bolc (Hg.), *Natural Language Generation Systems*, Seite 47-68. Springer, New York, 1988.
- [Matthiessen, 1985] C. Matthiessen. The Systemic Framework in Text Generation: Nigel. In J. Benson und W. Greaves (flg.), *Systemic Perspectives on Discourse. Volume 1*. Ablex, Norwood, NJ., 1985.
- [McCoy, 1986] K. F. McCoy. The ROMPER System: Responding to Object-Related Misconceptions using Perspective. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, New York, Seite 97-105. June 1986.
- [McDonald und Bolc, 1988] David D. McDonald und Leonard Bolc (Hg.). *Natural Language Generation Systems*. Springer, Berlin, New York, 1988. Symbolic Computation.
- [McDonald und Pustejovsky, 1985] D. D. McDonald und J. D. Pustejovsky. TAGs as a Grammatical Formalism for Generation. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, Chicago, Illinois, Seite 94-103. July 1985.
- [McDonald, 1983] D. D. McDonald. Natural Language Generation as a Computational Problem: An Introduction. In M. Brady und R. C. Berwick (Hg.), *Computational Models of Discourse*. MIT Press, Cambridge, Mass., 1983.

- [McDonald, 1987] D. D. McDonald. Natural language generation. In S. C. Shapiro (Hg.), *Encyclopedia of Artificial Intelligence*, Seite 642-655. John Wiley and Sons, Chichester, 1987.
- [McKeown und Swartout, 1988] K. R. McKeown und W. R. Swartout. Language generation and explanation. In M. Zock und G. Sabah (Hg.), *Advances in Natural Language Generation, Volume 1*, Kapitel 1, Seite 1-52. Pinter Publishers Ltd., London, 1988.
- [McKeown, 1985] K. R. McKeown. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge University Press, Cambridge, 1985.
- [Meter *et al.*, 1987] Marie W. Meter, David D. McDonald, S.D. Anderson, D. Forster, L.S. Gay, A.K. Huettner und Penelope Sibun. MUMBLE-86: Design and Implementation. Report No. 87-87, University of Cambridge, COINS, Cambridge, Mass., 1987.
- [Meter, 1988] M. W. Meter. The implications of revisions for natural language generation. In *Proceedings of the Fourth International Natural Language Generation Workshop*, Catalina Island. July 17-21 1988.
- [Meter, 1990] M. W. Meter. The "Generation Gap". The Problem of Expressibility inText Planning. Report No. 7347, Bolt, Beranek and Newman Inc., Cambridge, Mass., February 1990.
- [Momma und Dörre, 1988] Stefan Momma und Jochen Dörre. Generation from f-Structures. In E. Klein und J. van Benthem (Hg.), *Categories, Polymorphism and Unification*. Universität Edinburgh, Centre for Cognitive Science and Universität Amsterdam, Institute for Language, Logic and Information, 1988.
- [Moore und Paris, 1989] J.D. Moore und C.L. Paris. Planning Text for Advisory Dialogues. In *Proc. Conf. of the 27th Annual Meeting of the ACL*, Vancouver, Seite 203-211. 1989.
- [Moore, 1980] R. C. Moore. Reasoning about Knowledge and Action. Technical Report No. 191, SRI International Artificial Intelligence Center, 1980.
- [Neumann und van Noord, 1992] Günter Neumann und Gertjan van Noord. Reversibility and Self-Monitoring in Natural Language Generation. In Tomek Strzalkowski (Hg.), *Reversible Grammars and Natural Language Processing*. Kluwer, 1992. im Erscheinen.
- [Neumann, 1991] Günter Neumann. A Bidirectional Model for Natural Language Processing. In *Proc. 5th Conf. of the European Chapter of the ACL*, Berlin, Seite 245-250. 1991.
- [Nirenburg und Nirenburg, 1988] S. Nirenburg und I. Nirenburg. A framework for lexical selection in natural language generation. In *Proceedings of the 12th International Conference on Computational Linguistics and the 24th Annual Meeting of the Association for Computational Linguistics*, Budapest, Hungary, Seite 471-475. August 22-27 1988.

- [Novak, 1987] H.-J. Novak. *Textgenerierung aus visuellen Daten: Beschreibungen von Straßenszenen*. Springer, Berlin/Heidelberg/New York, 1987.
- [Olson, 1970] D. R. Olson. Language and Thought: Aspects of a Cognitive Theory of Semantics. *Psychological Review*, (77):257-273, 1970.
- [Paris et al., 1991] Cecile Paris, William Swartout und William Mann (Hg.). *Natural Language Generation in Artificial Intelligence and Computational Linguistics*. Kluwer, Dordrecht, 1991.
- [Pechmann, 1984] T. Pechmann. *Überspezifizierung und Betonung in referentieller Kommunikation*. Doktorarbeit, Universität Mannheim, 1984.
- [Pollard und Sag, 1987] Carl J. Pollard und Ivan A. Sag. *Information-Based Syntax and Semantics. Vol. 1*. Univ. of Chicago Press, Chicago, 1987. Center for the Studies of Language and Information, CSLI Lecture Note No. 13.
- [Reiter, 1990] Ehud B. Reiter. Generating Descriptions that Exploit a User's Domain Knowledge. In Robert Dale, Chris Mellish und Michael Zock (Hg.), *Current Research in Natural Language Generation*, Seite 257-286. Academic Press, 1990.
- [Reithinger, 1992] Norbert Reithinger. *Eine parallele Architektur zur inkrementellen Generierung multimodaler Dialogbeiträge*. Infix, Sankt Augustin, 1992.
- [Retz-Schmidt, 1988] Gudula Retz-Schmidt. Various Views on Spatial Prepositions. Bericht 33, Sonderforschungsbereich 314, Universität des Saarlandes, 1988.
- [Rosch, 1978] Eleanor Rosch. Principles of Categorization. In E. Rosch und B. Lloyd (Hg.), *Cognition and Categorization*, Seite 27-48. Lawrence Erlbaum, Hillsdale, 1978.
- [Rosner, 1987] Dietmar Rosner. Generation of Content vs. Generation of Form: A Review of Recent Work in the SEMSYN Project. In K. Morik (Hg.), *GWAI-87. 11th German Workshop on Artificial Intelligence*, Berlin, New York, Seite 308-314. Springer, 1987. IFB Bd. 152.
- [Schegloff, 1987] E. A. Schegloff. Between macro and micro: Contexts and other connections. In J. Alexander, B. Giesen, R. Munch und N. Smelser (Hg.), *The Micro-Macro Link*. University of California Press, Berkeley, 1987. „•, . . . , . . . \
- [Shieber et al, 1990] Stuart M. Shieber, Gertjan van Noord, Robert C. Moore und Fernando C. N. Pereira. A Semantic Driven Generation Algorithm for Unification-Based Formalisms. *Computational Linguistics*, 16(1):30-42, 1990.
- [Sidner, 1983] C. L. Sidner. Focusing in the Comprehension of Definite Anaphora. In M. Brady und R. C. Berwick (Hg.), *Computational Models of Discourse*, Seite 267-330. MIT Press, Cambridge, MA., 1983. • • ' • '«'».

- [Strzalkowski, 1992] Tomek Strzalkowski (Hg.). *Reversible Grammars and Natural Language Processing*. Kluwer, 1992. im Erscheinen.
- [Thompson, 1977] H. Thompson. Strategy and Tactics in Language Production. In W. A. Beach, S. E. Fox und S. Philosoph (Hg.), *Papers from the Thirteenth Regional Meeting of the Chicago Linguistics Society*. Chicago, April 14-16 1977.
- [Van Dijk, 1985] T. A. Van Dijk. *Handbook of Discourse Analysis*. Academic Press, London, 1985.
- [Wahlster *et al.*, 1983] W. Wahlster, H. Marburger, A. Jameson und S. Busemann. Over-answering Yes-No Questions: Extended Response to NL Questions in a Vision System. In *Proc. 8th IJCAI-83*, Seite 643-646, 1983.
- [Wedekind, 1986] Jürgen Wedekind. A Concept of Derivation for LEG. In *Proc. 11th COLING-86*, Bonn, Seite 487-489. 1986.
- [Wedekind, 1988] J. Wedekind. Generation as structure driven derivation. In *Proceedings of the 12th International Conference on Computational Linguistics and the 24th Annual Meeting of the Association for Computational Linguistics*, Budapest, Hungary, Seite 732-737. August 22-27 1988.
- [Zock und Sabah, 1988] Michael Zock und Gerard Sabah (Hg.). *Advances in Natural Language Generation. An Interdisciplinary Perspective. Vol. 1, 2*. Frances Pinter, London, 1988.

Index

- Absicht
 - kommunikative, 8, 11, 13, 15, 20, 28, 31
- Antwort
 - direkte, 14, 23
 - indirekte, 23
- Appelt, D.E., 27, 28
- Architektur, 20
 - integriert, 19
 - interagierend, 10
 - kaskadiert, 5
 - sequentiell, 19
 - zweistufig, 4
- ARGOT, 24-26
- Artikulation, 9, 10
- Auswahlproblem, 20

- Baumadjunktionsgrammatik, 39-42
- Bedeutungsstruktur, 20
 - Konstruktion der, 13
 - konzeptuelle, 9
 - vorsprachliche, 8, 10-12, 14

- COOP, 23

- DIOGENES, 34
- Diskursbereich, 6, 7, 21, 24
- Diskurshistorie, 14
- Diskursmodell, S, I I , 1(i), 38
- Diskursstruktur, 3
- Diskurstyp, 15, 22, 26
- Diskursziel, 15

- Entscheidungsprozeß, 4, 6, 19

- Formulator, 9-12, 20
- FREGE, 33

- generation gap, 30
- Generierung
 - formbestimmende, 4
 - inhaltsbestimmende, 4
 - inkrementelle, 10, 40
 - semantic-head-driven, 49
- GEOTEX, 19, 27
- GPSG, 7, 42-46
- Grammatik
 - bidirektionale, 7, 39, 51
- Grice'sche Maximen, 22
- Grice, H.P., 22

- HAM-ANS, 23, 25, 26, 33
- Hovy, E.H., 30
- How-to-Say-It, *siehe* Generierung, formbestimmende
- HPSG, 7, 39, 51

- Informationsstruktur, 14
- Inhaltsbestimmung
 - planbasiert, 26, 27
 - schemabasiert, 26
- Intention, *siehe* Absicht, kommunikative

- KAMP, 19, 27
- Kempen, G., 4
- Kontext
 - außersprachlicher, 13
 - sprachlicher, 13
- Konzeptualisierer, 4, 8

- Levelt, W.J.M., 8, 10, 12, 17
- Lexikon, 18
 - mentales, 9
- Lexikoneintrag, 18, 42
- LEG, 7, 46-47
- Linearisierung, 17

- Makroplanung, 8, 13, 17
- Mann, W.C., 29
- McDonald, D.D., 4
- McKeown, K.R., 26, 27, 30
- Mikroplanung, 8
- monitoring, 8, 9
- MUMBLE, 19, 31, 32

NAOS, 19, 27
NIGEL, 38

PAULINE, 19
PENMAN, 38
Perspektive, 9, 12, 14
Pläne, 25, 29
POPEL, 19, 20
Präsupposition, 23

Realisierung, 4, 31, 39
Referenz, 16
Rhetorical Structure Theory, 29
rhetorisches Prädikat, 26

Segmentgrammatik, 39, 41
SEMTEX, 19, 27
Sprache
 geschriebene, 33
 gesprochene, 5, 8, 10, 20
 interne, 9
Sprechakt, 13, 25, 30
Sprechermodell, 8
stonewalling effect, 23
SUTRA, 32, 33
systemische Grammatik, 36-38

TAG, *siehe* Baumadjunktionsgrammatik
TEXT, 19, 26, 27
Textgenerierung, 6, 26
Textsorte, *stehe* Diskurstyp
Texttyp, *siehe* Diskurstyp
TFS, 51
Thompson, S., 29
TRACK, 24

What-to-Say, *siehe* Generierung, inhalts-
 bestimmende
Wortwahl, 33