Diskursplanung und Planrealisierung

Helmut Horacek
Universität des Saarlandes
FB 14 Informatik
D-66041 Saarbrücken, Deutschland
horacek@cs.uni-sb.de

21. März 2001

1 Einleitung

Menschen produzieren Texte in verschiedener Form, und diese unterscheiden sich je nach Gegebenheit der vorliegenden Situation oft erheblich hinsichtlich Inhalt, Detail, und Ausdrucksweise. Um diese Fähigkeit zumindest für aufgabenorienterte Texte auf einem Rechner nachzubilden, werden relevante Aspekte einer Situation modelliert, in der die Produktion einer Äußerung angebracht ist, und es wird versucht durch schrittweises Anreichern und Konkretisieren einer ursprünglich abstrakten und oft vagen Spezifikation einer kommunikativen Absicht zu einer vollständigen Beschreibung eines Diskursbeitrags zu kommen und im weiteren zu einer adäquaten sprachlichen Umsetzung dieser Beschreibung. Die Rolle der Diskursplanung, die den im wesentlichen gestalterischen Anteil eines solchen Verfahrens ausmacht, besteht in diesem Kontext im Ausdrücken dieser *kommunikativen Absicht* in Form eines *Textplans*. Die Umsetzung dieses Plans erfordert dann eine Überführung der darin befindlichen Spezifikationen in Text, und zwar mit den vorhandenen linguistischen Mitteln. Betrachten wir als Beispiel dazu eine Wegauskunft. Die kommunikative Absicht, einem Auskunftsuchenden Kenntnis über den Weg von einem Ort A zu einem Ort B zu verschaffen, bedarf bei nichttrivialer Wegstrecke zu ihrer Umsetzung einer Folge sorgfältig geplanter und realisierter Äußerungen. Darin werden die Abschnitte, aus denen sich ein solcher Weg zusammensetzt, in geordneter Weise und geeigneter Ausführlichkeit angeführt, je nach situativen Gegebenheiten und Informationsbedürfnis des Adressaten, wie etwa in:

Das ist ein kurzer Weg. Sie gehen von hier weiter bis zur Kirche und biegen dort links ab. Dann kommen Sie an einem großen Haus, dem Rathaus, vorbei und erreichen schließlich B.

Um einen solchen Text zu erzeugen, muß der gesamte Weg mental in genügend kleine, leicht faßliche Abschnitte zerlegt und es müssen dazu Beschreibungen gebildet werden. Diese Beschreibungen sollten dann in kohärenter Weise zusammengesetzt und zu besseren Orientierung an einigen Stellen mit Referenzen auf auffällige Objekte, sogenannte Landmarken, versehen werden. Schließlich ist eine angemessene sprachliche Umsetzung des Plans in Form von Text erforderlich, von der man plausibel annehmen kann, daß dem Auskunftsuchenden mit der erzeugten Beschreibung gedient ist.

Im weiteren Verlauf dieses Kapitels wollen wir Verfahren zur automatischen Planung solcher Diskurse und deren Umsetzung in Form von Text näher untersuchen, meist durch Beispiele von Wegauskünften, fallweise aber auch aus anderen Bereichen erläutert. Dazu beschreiben wir zunächst die Aufgaben der Diskursplanung und der Planumsetzung. Danach betrachten wir die Funktionalität dieser Prozesse auf einer etwas formaleren Basis. Im Hauptteil werden die grundlegenden Methoden der Diskursplanung und Planumsetzung vorgestellt, wesentliche Annahmen dieser Methoden und daraus resultierende Probleme diskutiert, sowie ihre Eignung charakterisiert. Abschließend wird die Rolle der wichtigsten mit Generierung befaßten linguistischen Theorien erläutert.

2 Aufgaben der Planung und Umsetzung

Generell besteht die Aufgaben der Diskursplanung und der Umsetzung dieses Plans darin, eine bestimmte kommunikative Absicht bei gegebenem Kontext und Annahmen über den Adressaten zu interpretieren, daraus die Spezifikation eines Textplans zu bilden, und schließlich daraus mittels zur Verfügung stehender linguistischer Ressourcen einen Text zu erzeugen, von dem plausibel angenommen werden kann, daß sein Äußern die kommunikative Absicht erfüllt. [1] Um die dabei anfallenden Planungsaufgaben in einem formalen System zu erfüllen, wird eine der Domäne und Anwendung angemessene Wissensbasis herangezogen und in zweierlei Hinsicht verarbeitet:

- Es werden solche Elemente der Wissensbasis selektiert, die für die Erfüllung der kommunikativen Absicht als ausreichend relevant erachtet werden. Dieser Aspekt wird als *Inhaltsfestlegung* bezeichnet.
- Die selektierten Informationen, die in der Wissensbasis in kontextneutraler Weise repräsentiert sind, müssen in eine zusammenhängende,

[1] Es kann dabei keine Erfolgsgarantie geben, da menschliche Kommunikation inhärent fehlerhaft ist.

kohärente, den Bedürfnissen des Adressaten angemessene Struktur gebracht werden. Dieser Aspekt wird als *Inhaltsorganisation* bezeichnet.

Wegen der großen Zahl und unterschiedlichen Art von Maßnahmen, die bei der Planung und Planumsetzung berücksichtigt werden müssen, wird der gesamte Verarbeitungsprozeß meist in drei Abschnitte gegliedert:

- Die Festlegung, welche Informationen kommuniziert werden, wie sie angeordnet und wie die einzelnen Teile rhetorisch zueinander in Beziehung gesetzt werden, macht in gewissem Sinne die Planung im Großen aus. Dieser Abschnitt wird als *Textplanung* bezeichnet.

- Basierend auf der durch die Textplanung aufgebaute Grobstruktur wird eine feinere Strukturierung vorgenommen, die Satzebene betreffend. Dazu gehören hauptsächlich die Kompaktifizierung von Aussagen mit gemeinsamen Teilen (Aggregation), Bestimmung der Satzgrenzen und der Topik-Fokus Struktur, Erzeugung von Referenzausdrücken und Wahl der linguistischen Ressourcen zum Ausdrücken konzeptueller Inhalte (Lexikalisierung). Dieser Abschnitt wird als *Satzplanung* bezeichnet.

- Schließlich werden die abstrakten Satzrepräsentationen in Text umgesetzt. Dieser Abschnitt wird als linguistische *Oberflächenrealisierung* bezeichnet.

Dabei besteht das Zwischenergebnis nach der Diskursplanung in einer teilgeordneten Struktur von Spezifikationen satzartiger Aussagen, mit zusätzlichen Angaben für die rhetorischen Bezüge zwischen diesen Elementen. In der Satzplanung werden diese Strukturen um Spezifikationen für Referenzausdrücke, Konnektive, Topik-Fokus Struktur und Rolle in der Diskursstruktur (Haupt- oder Nebensatz sowie deren Reihenfolge) angereichert, sowie durch Maßnahmen wie Aggregierung und Lexikalisierung teilweise umstrukturiert.

Sowohl die Art des Zusammenspiels zwischen Inhaltsfestlegung und -organisation und auch die Einbindung und Abfolge einzelner Maßnahmen der Satzplanung ist jedoch nicht weiter festgelegt. Da sie sich in einer Reihe von Fällen gegenseitig bedingen, ist eine opportunistische Verzahnung dieser beiden Aspekte bzw. der unterschiedlichen Maßnahmen in einem Planungsprozeß aus theoretischer Sicht motiviert. Dem tragen auch verschiedene aufwendige, nicht-sequentielle Architekturtypen Rechnung wie integrierte, interaktive und revisionsbasierte. In implementierten Systemen und auch in methodischen Verfahren wird hingegen aus praktischen Erwägungen meist eine strikte Abfolge — zunächst Inhaltsfestlegung, dann -organisation, dann Maßnahmen zur Satzplanung — eingehalten, mit Ausnahme der Wahl von

```
    Kommunikatives Ziel
             |
  -----------------------
  |     Textplaner      |
  -----------------------
             |
          Textplan
             |
 -----------------------
 |       Satzplaner    |
 -----------------------
             |
 Abstrakte Spezifikation
             |
---------------------------
| Oberflaechenrealisierer |
---------------------------
             |
            Text
```

Abbildung 1: Die 'Standard'-Generieungsarchitektur

Beschreibungselementen zur Referenzgenerierung (gemäß der mittlerweile als "Standardarchitektur" geltenden Aufteilung [Rei94], siehe Abbildung 1). Eine der wenigen Ausnahmen von diesem Vorgehen bildet das System EES (*E*xplainable *E*xpert *S*ystems, [Moo89]), in dem Teile der Erklärungsinhalte erst in Abhängigkeit von dem Zustand des bereits gebildeten, noch partiellen Textplans ausgewählt und in diesen eingefügt werden. In Bezug auf Maßnahmen im Rahmen der Satzplanung wird deren Reihenfolge und Verzahnung meist opportunistisch in der Weise festgelegt, die der sich die spezifischen Anforderungen des Anwendungssystems am leichtesten umsetzen lassen.

Wie auch immer die Organisation des gesamten Planungsprozesses (Diskurs- und Satzplanung) aufgebaut ist, es muß die Gewährleitung der Umsetzbarkeit des erzeugten Textplans mit den vorliegenden linguistischen Mitteln sichergestellt sein (das Problem der "generation gap" [Met92] muß gelöst werden). Dieses Problem entsteht grundlegend dadurch, daß die Ausgangsspezifikationen für die Textgenerierung im allgemeinen nicht-linguistischer Natur sind. Deshalb muß im Rahmen von Text- und Satzplanung, beim Übergang von nicht-linguistischen zu linguistischen Elementen, die Zusammensetzbarkeit der ausgewählten linguistischen Elemente gesichert sein — inbesonders Typkompatibilitäten, wie Bedingungen an die Füller von Kasusrahmen. So kann etwa ein *Beschluß*, der zu einem *wichtigen* Ergebnis geführt hat, nur in nominalisierter Form in einer Phrase ausgedrückt werden ("wir haben einen *wichtigen Beschluß* gefaßt") und nicht als Vorgang (* "wir haben *wichtig beschlossen*").

An vielen Stellen der Diskurs- und Satzplanung bestehen Wahlmöglichkeiten in Bezug auf Ausführlichkeit, Ausdrucksweise oder Anordnung von Textteilen. Dabei sind jedoch die Gründe für die Auswahl zwischen möglichen Varianten im allgemeinen oft noch wenig verstanden. Meist werden Präferenzen eher auf pragmatisch motivierte Entscheidungen und spezifische Domänenkonventionen zurückgeführt und nicht auf rhetorisch bedingte Faktoren. Im allgemeinen erweisen sich die linguistischen Kriterien eher als notwendig um offensichtlich ungeeignete Varianten von Textspezifikationen zu vermeiden, aber nicht als hinreichend, um Präferenzen in einem bestimmten situativen Kontext abzuleiten (siehe [Kit91]).

Für die Wegplanung bedeutet die Inhaltsbestimmung die Festlegung eines Pfades vom Ausgangsort zum Zielort, inklusive der an den entsprechenden Stellen einbezogenen Landmarken und sonstiger ergänzender Information wie etwa Übersichtshinweise. Die Inhaltsorganisation ergibt sich in Bezug auf die Reihenfolge von Aussagen bereits aus der Folge von Wegsegmenten. In der Satzplanung können allerdings Beschreibungen von mehreren Segmenten fallweise koordiniert und ergänzende Informationen, etwa zu Landmarken, geeignet eingefügt werden. Wahlmöglichkeiten bestehen unter anderem darin festzulegen, welche Teilstrecken explizit genannt und welche koordiniert werden, welche Landmarken zur Referenz herangezogen werden und wie sie beschrieben werden, und welche begleitenden Hinweise in die eigentliche Wegbeschreibung eingefügt werden. Dabei kann durch eine Verzahnung mit der Inhaltsorganisation eine motiviertere Inhaltsauswahl, etwa bei Entscheidungen zu Übersichtshinweisen, unterstützt werden, wenn Informationen über Segmentkoordinationen bereits bei der Inhaltsauswahl zur Verfügung stehen.

3 Funktionalität des Planungsprozesses

Die Beschreibung einer Situation, in der die Erzeugung einer Äußerung oder eines Textes angebracht ist, kann im Sinne von [Rei00] als Quadrupel $< k, c, u, d >$ aufgefaßt werden, wobei k die Wissensbasis, c die kommunikative Absicht, u das Adressatenmodell und d die Diskurshistorie darstellen. Im Rahmen dieses Buchkapitels beschränken wir uns im wesentlichen auf die Rolle der ersten beiden Komponenten, da Adressaten- und Dialogmodellierung in eigenen Kapiteln beschrieben werden. Wir verweisen nur gelegentlich bei der Auswahl zwischen alternativen Möglichkeiten auf Annahmen aus einem Adressatenmodell oder auf einer Elemente einer Diskurshistorie, die die damit verbundenen Entscheidungen motivieren.

Bei der Spezifikation einer kommunikativen Absicht und der damit verbundenen Inhalte der Wissensbasis können Diskrepanzen zwischen den Annahmen eines Diskursplanungsverfahrens und den vorliegenden Gegebenheiten einer Anwendungseinbettung auftreten. Während Diskursplanungsverfahren meist vereinfachend annehmen, daß die Propositionen, auf denen sie arbeiten, linguistisch motiviert aufgebaut sind und in Haupt- oder Nebensätzen ausgedrückt werden können, können in Hintergrundsystemen eine solche Absicht und speziell die damit verbundenen Inhalte in sehr unterschiedlicher, oft nicht-linguistischer Form vorliegen. Dabei kann es sich unter anderem um Tabellen mit numerischen Daten, Relationen einer Datenbank, oder Propositionen aus einem wissensbasierten System handeln. Je nach Inhalt der Wissensbasis muß also unter Umständen eine Interpretation von nicht-linguistischen Daten vorgenommen werden, durch die ihnen Entsprechungen in Form von Propositionen mit sprachlich ausdrückbaren Prädikaten zugeordnet werden. Die Definition solcher Zuordnungen kann durch eine Kombination von Auswahl, Zusammenfassung und Inferenzen auf der Basis der vorliegenden Daten erfolgen und ist in hohem Maße eine anwendungs- und domänenabhängige Aufgabe, deren Systematik noch wenig entwickelt ist. Diese Umstrukturierungen, die die Bildung von durch lexikalisches Material ausdrückbaren Propositionen aus den Elementen der Wissensbasis zum Ziel haben, werden auch als Lexikalisierung im *eigentlichen Sinn* bezeichnet, während die Auswahl zwischen alternativen Lexemen (Wortwahl) als Lexikalisierung im *engeren* Sinn angesehen wird. Diese Kategorisierung geht auf Analysen im RAGS-Projekt (*R*eference *A*rchitecture for *G*eneration *S*ystems) zurück [Cah98]. Einige Beispiele bei der Konstruktion von linguistisch motivierten Strukturen aus Repräsentationen von Leistungssystemen, deren Ergebnisse oder Verhalten mit einem Generierungssystem sprachlich beschrieben werden soll, findet man in [Rei00]. Darüberhinaus sind einige Arten von Umformungen linguistisch und taxonomisch motiviert und führen zur Verbesserung der Inhaltsstrukturierung — die *M*eaning *T*ext *T*heorie (siehe das vorletzte Teilkapitel) stellt hierzu einen variantenreiches Instrumentarium zur Verfügung, wobei allerdings der Aufwand zur Erstellung des lexikalischen Materials und der Berechnungsaufwand zur Umstrukturierung der Spezifikationen extrem hoch sind.

Um der Annahme von Diskursplanungsverfahren in Bezug auf Eingabespezifikationen gerecht zu werden, erfolgt oft *vor* der Diskursplanung die Lexikalisierung im eigentlichen Sinn, während *nach* der Diskursplanung nur noch die Auswahl von Lexemen vorgenommen wird. Diese umfaßt die Wahl zwischen Synonymen, Alternativen von Ausdrücken mit unterschiedlichen Perspektiven (wie etwa "kaufen" und "verkaufen") und eventuellen Nominalisierungen (wie etwa "der Kauf") oder ähnliche Variationen, die den gewünschten Informationsgehalt in vergleichbaren Strukturen ausdrücken. Diese Aufteilung der Lexikalisierung ist systematisch, und sie ermöglicht eine einfachere Textplanung. Sie legt jedoch viel Verantwortung auf die Lexikalisierung im eigentlichen Sinne, so daß die erzeugten Propositionen im Zusammenhang oft sprachlich verbesserungsbedürftig sind, was sich dann in Texten äußert, die als schematisch und hölzern empfundenen werden. Wir werden anhand von Beispielen von Wegbeschreibungen sehen, welche Verbesserungen durch Umstrukturierung der Information in Varianten linguistisch motivierter Repräsentationen und sprachliche Umformulierung erzielt werden können, wobei auch klar werden sollte, daß das Bilden der entsprechenden Propositionen ohne den Kontext der Diskursplanung nicht motiviert erfolgen kann.

Die Ergebnisse der Diskursplanung, die sich in einer strukturellen Zusammensetzung von satzartigen Propositionen, meist einer rhetorischen Baumstruktur, manifestieren, sind in mehrerer Hinsicht zu unspezifisch, um eine typische Oberflächenrealisierungskomponente anzusteuern. Zu den dazu erforderlichen Maßnahmen gehört in jedem Fall die Wahl von Lexemen für die in den Propositionen vorkommenden Prädikate. In der eingangs angegebenen Wegauskunft etwa wird die (hypothetische) Fortbewegung des Auskunftssuchenden einmal durch "geben" und einmal durch "vorbeikommen" ausgedrückt, in diesem Fall in Abhängigkeit von der relativen Position der Landmarken. Eine weitere Aufgabe besteht in der Erzeugung von Referenzausdrücken für die in den Propositionen durch Variable repräsentierten Objekte. Die erzeugten Phrasen können Pronominalisierungen sein, einfache Referenzen mittels Objektkategorien oder umfangrechere Beschreibungen (etwa "ein großes Haus, das Rathaus"). Schließlich erfolgt die Bildung von expliziten Satzstrukturen aus den diesbezüglich neutralen Propositionen, mit Festlegungen für Haupt- und Nebensatz, sowie Koordination von Strukturen mit partiell übereinstimmenden Teilen. In der eingangs angegebenen Wegauskunft etwa wird die Beschreibung des zweiten Satzes "Sie gehen von hier weiter bis zur Kirche und biegen dort links ab." aus ursprünglich zwei Propositionen gebildet, wobei wegen Subjektsgleichheit das Vorkommen von "sie" im zweiten Hauptsatz elidiert wird. Das Ergebnis der Satzplanung kann von unterschiedlicher Art sein, je nach Anforderungen der verwendete Oberflächenrealisierungskomponente. Mögliche Varianten sind abstrakte syntaktische Beschreibungen, lexikalisierte Kasusrahmen sowie Mischformen, auch mit festgelegten Textteilen an einzelnen Stellen dieser Strukturen.

Betrachten wir nun die Aufgabe einer Wegbeschreibung aus einer etwas formaleren Sicht. Dafür könnten die kommunikative Absicht c und die Wissensbasis k (hier ist nur jener Teil angegeben, der für den vorliegenden Fall relevant ist) in einer nicht-linguistischen Form wie folgt aussehen: [2]

- $c = Describe(Route, A, B)$
- $k = \{at(User, A), Path(A, C), Path(C, D), Path(D, B),$
  $Left(Path(A, C), Path(C, D)), Straight(Path(C, D), Path(D, B)),$
  $Type(C, church), Type(D, town-hall), Property(D, big), ...\}$

Die Spezifikation der kommunikativen Absicht besagt in diesem Beispiel, daß eine Beschreibung des Wegs von A nach B zu produzieren ist. Die Wissensbasis enthält unter anderem den momentanen Standort des Adressaten der Beschreibung ($at(User, A)$), Teilstrecken ($Path(A, C)$), Richtungsbezüge zwischen angrenzenden Teilstrecken ($Left(Path(A, C),\ Path(C, D))$) und Eigenschaften von Landmarken ($Type(D, town - hall)$, $Property(D, big)$).

Für Zwecke der Diskursgenerierung müssen die Spezifikationen der Wissensbasis zunächst in linguistisch motivierte Prädikate überführt werden, die an Stelle der abstrakten Pfadeigenschaften für den Kommunikationszweck geeignete Prädikationen ausdrücken wie (hypothetische) Fortbewegungen einer Person (hier: des Auskunftssuchenden). Weitergehende Umstrukturierungen der Prädikationen der Wissensbasis sind in diesem Fall zum Unterschied von anderen Domänen und Anwendungssystemen nicht nötig. Dieser Schritt könnte als Lexikalisierung im eigentlichen Sinne aufgefaßt werden, und das Resultat würde etwa wie folgt aussehen:

$\{be(User, at(A)),\ go(User, from(A), to(C)),$
$go(User, from(C), to(D)),\ go(User, from(D), to(B)),$
$turn(User, at(C), direction(left)),\ turn(User, at(D), direction(none)),$
$church(C), town - hall(D), big(D)...\}$

In der weiteren Verarbeitung wird sich allerdings zeigen, daß Umstrukturierungen innerhalb dieser linguistisch motivierten Prädikationen (etwa Koordination mehrerer Prädikationen, Anreichern mit Konnektiven, diversifizierte Wortwahl) unerläßlich sind und die Spezifikationen des Texts kommunikativ adäquater zu gestalten, etwa zur Vermeidung stereotyper Wiederholungen.

Wir wollen davon ausgehen, daß diese Interpretation der Elemente der Wissensbasis im Zusammenhang mit der Diskursplanung vorgenommen wird,

[2] Um klar zwischen formalen und lexikalischen Spezifikationen unterscheiden zu können, wählen wir für die Prädikate der Wissensbasis englische Namen und für die sprachlichen Ausdrücke deutsche.

inklusive der Umstrukturierung zur Verbesserung der Informationsaufteilung. Während die Eingangsspezifikationen typischerweise ungeordnet sind oder zu anderen Zwecken strukturierte Informationen beinhalten (etwa zur Pfadsuche für die Wegauskunft), manifestiert sich das Ergebnis der Diskursplanung in einer baumartigen Textstruktur, in der die Blattknoten die einzelnen Aussagen des zu erzeugenden Texts enthalten (*messages* im Sinne von [Rei00]) und die Zwischenknoten Diskursrelationen darstellen, die die Bezüge dieser Aussagen beschreiben.

Eine Textstruktur für den in der Einführung angegebenen Text könnte dann wie in Abbildung 2 angegeben aussehen. Um diese Struktur zu erzeugen, werden zunächst die Elemente des Pfades von Ausgangsort A nach Zielort B als Sequenz von Fortbewegungen des Auskunftsuchenden ausgedrückt — darin manifestiert sich der Aspekt der Textorganisation. Zusätzlich werden kommunikativ motivierte Operationen (opportunistisch) angewandt, die diese Struktur umformen. Um dem Hörer das Verständnis der Wegbeschreibung zu erleichtern, wird eine kurze Charakterisierung des Gesamtwegs als BACKGROUND der gesamten Wegbeschreibung vorangestellt, und das Erreichen des Zielorts wird explizit angegeben ($be(User, at(B))$ wird mit einer EVIDENCE Relation versehen und eingefügt). Diese Informationsanreicherung kann als ein Teil der Inhaltsauswahl angesehen werden. Die am Ende der Strecke von A nach C nötige Richtungsänderung ergänzt mittels einer ELABORATION Relation die Beschreibung dieser Strecke. Die Beschreibung der folgenden zwei Teilstrecken zwischen denen keine Richtungsänderung stattfindet ($go(User, from(C), to(D))$ und $go(User, from(D), to(B))$) wird zu einer Strecke zusammengefaßt und durch die Beschreibung der dabei passierten Landmarke ausgedrückt ($pass(User, by(D))$). Die Rechtfertigung dieser Ausdruckswahl liegt darin, daß aus dem Vorbeigehen an einem Ort (hier: D) auch geschlossen werden kann, daß dazu zunächst der Weg von dem vorher genannten Ort zurückgelegt werden muß. Dies ist eine kontextuell motivierte Inferenz, die dem Hörer zugetraut wird, und solche Situationen kommen in natürlichen Dialogen häufig vor. Ähnliches gilt für den Weg zu dem danach genannten Ort. Diese Umformungsschritt können eher als ein Beispiel für Lexikalisierung im eigentlichen Sinn aufgefaßt werden denn als Aggregation — obwohl der Aggregationsbegriff von manchen Autoren sehr weit gefaßt wird — da eine Änderung des Prädikats ("pass" statt "go") vorgenommen wird und sich auch eine modifizierte Argumentstruktur ergibt (nur die unterwegs passierte Landmarke anstelle von Start- und Zielpunkt).

Um aus dieser Struktur eine für Oberflächengenerierung geeignete Zwischenrepräsentation zu erzeugen, müssen die Aussagen an den Blattknoten, insbesondere die darin vorkommenden Prädikate, lexikalisch ausgedrückt und mit einer linguistischen Umsetzung der Diskursrelationen (Satzbildung und

```
|   relation: BACKGROUND                                            |
|   short(path(A,B))                                                |
|                                                                   |
|   |   relation: EVIDENCE                                      |   |
|   |                                                           |   |
|   |   |   relation: SEQUENCE                              |   |   |
|   |   |                                                   |   |   |
|   |   |   |   relation: ELABORATION                   |   |   |   |
|   |   |   |   go(User,from(A),to(C))                  |   |   |   |
|   |   |   |   turn(User,at(C),direction(left))        |   |   |   |
|   |   |                                                   |   |   |
|   |   |   pass(User,by(D))                                |   |   |
|   |                                                           |   |
|   |   be(User,at(B))                                          |   |
```

Abbildung 2: Textstruktur für den Beispieltext

Reihenfolge) kombiniert werden, und es müssen Referenzen zu den Objekten (User, A, B, C und D) gebildet werden. Die lexikalische Umsetzung erfolgt weitgehend direkt ("go(from,to)" etwa wird als "weitergehen von ... bis ..." ausgedrückt, "turn(at,direction)" als "abbiegen" und "pass(by)" als "vorbeikommen an"). Die einzelnen Aussagen werden dann einfach gemäß ihrer chronologischen Reihenfolge als Hauptsätze aneinandergereiht, wobei aus der SEQUENCE Relation ein zeitliches "dann" als explizite Markierung eingefügt wird. Die Herausstellung der BACKGROUND Information als eigener Satz ist rhetorisch gut motivierbar. Bei den folgenden Aussagen wird eine Koordination jeweils zweier Hauptsätze vorgenommen, mit Elision des Subjekts im jeweils zweiten Teil — dies ist in gewissem Sinn opportunistisch, es gibt aber auch ähnlich gute Ausdrucksmöglichkeiten. Bei der Bilding von Referenzausdrücken wird der Bezug auf den Benutzer ("sie"), den Standort ("hier"), und auf die zweite Referenz auf C ("dort") mit Pronomen vorgenommen. Für die Vorkommen der Landmarken C (das erste Vorkommen) und D werden Nominalphrasen erzeugt bestehend aus einer Kategorie ("die Kirche") bzw. eine komplexere Beschreibung ("großes Haus, das Rathaus"), was durch die Annahme motiviert ist, daß die Beschreibung "Kirche" ausreichend zur Identifikation ist, während dies bei "Rathaus" nicht sicher zutrifft.

Die Ergebnisse der Satzplanung können in verschiedenartigen Repräsentation ausgedrückt werden (wie etwa abstrakte kasus-rahmenartige Repräsentationen, lexikalisierte Kasusrahmen, oder abstrakte syntaktische Teilbeschreibungen). Die Entscheidung für eine dieser Beschreibungsarten hängt von der

```
| head:    | sein                          | |
| subject: | head: pronoun                 | |
|          | type: path                    | |
| object:  | head: | Weg |                 | |
|          | features: | definite: false | | |
|          | modifier: | head: | kurz |  | | |
| and
| | head:          | weitergehen         |       | |
| | subject:       | head: pronoun       |       | |
| |                | type: human         |       | |
| |                | referent: user      |       | |
| | from-location: | head: pronoun       |       | |
| |                | type: location      |       | |
| |                | preposition: von    |       | |
| |                | referent: A         |       | |
| | to-location:   | head: | Kirche |    |       | |
| |                | definite: true      |       | |
| |                | preposition: bis+zu |       | |
| | head:      | abbiegen        |               | |
| | subject:   | head: elided    |               | |
| | location:  | head:  pronoun  |               | |
| |            | referent: C     |               | |
| | direction: | head: | links | |               | |
| and
| | head:     | vorbeikommen                        | | |
| | subject:  | head: pronoun                       | | |
| |           | type: human                         | | |
| |           | referent: user                      | | |
| | time:     | head: pronoun                       | | |
| |           | type: time                          | | |
| | location: | head: | Haus |                      | | |
| |           | definite: false                     | | |
| |           | preposition: an                     | | |
| |           | modifier:   | head: | gross |   | | | |
| |           | apposition: | head: | Rathaus | | | | |
| |           |             | definite: true    | | | |
|                                                     |
| | head:     | erreichen          |                  | |
| | subject:  | head: elided       |                  | |
| | adjunct:  | head: schliesslich |                  | |
| | location: | head: | B |        |                  | |
```

Abbildung 3: Abstrakte syntaktische Beschreibung des Beispieltexts

Art des Zusammenspiels zwischen Textplanung und Lexikalisierung ab. Die Realisierung dieses Zusammenspiels ist theoretisch noch unzureichend geklärt; sie erfolgt meist opportunistisch durch die Gegebenheiten der Anwendung motiviert (siehe auch die Diskussion der Alternativen in [Rei00]). In den in Abbildung 3 angegebenen Strukturen wurden abstrakte syntaktische Beschreibungen gewählt.

In der Folge wollen wir Verfahren näher betrachten, mit denen der Aufbau einer Textstruktur ermöglicht oder zumindest wesentlich gesteuert wird, sowie Verfahren zu verschiedenen Aspekten der Satzplanung und Oberflächenrealisierung.

4 Methoden zur Diskursplanung

Zur Umsetzung der Aufgaben der Diskursplanung sind zwei grundlegende Methoden entwickelt worden:

- Textschemata, die erste Methode zur Textplanung, eingeführt von McKeown [McK85b, McK85a]

- Schrittweise Planung mit Diskursrelationen, zuerst operationalisiert für Relationen der RST [Man91] von Hovy [Hov93]

In der Folge beschreiben wir zunächst die Funktionalität jeder dieser Methoden, und vergleichen dann ihre Vor- und Nachteile. Die Anwendungsbeschränkungen und einige Ansätze zur Überwindung dieser Beschränkungen behandeln wir in einem separaten Kapitel.

4.1 Textschemata

Schemata sind vordefinierte Repräsentationen von stereotypen Strukturen, die Texten zugrundeliegen, die etwa der Größe eines Absatzes entsprechen. Ein Schema ist als ein Muster definiert, das Bedingungen für den Inhalt und die Reihenfolge von Haupt- und Nebensätzen innerhalb eines Absatzes definiert. Durch die geeignete Spezifikation solcher Schemata — McKeown hat ihre Versionen auf der Basis einer Reihe empirischer Analysen erstellt — kann die erforderliche Kohärenz unterstützt werden. Die Elemente eines Schemas sind rhetorische Prädikate, die die Art von Information charakterisieren, die an der entsprechenden Stelle im Schema angeführt werden kann, und die mit der Semantik der Elemente der Wissensbasis abgeglichen werden müssen. Eine gewisse Variabilität in der Abfolge von Sätzen wird erreicht durch Optionalität einzelner Elemente, Alternativen von rhetorischen Prädikaten, sowie Wiederholungen. Zusätzlich können Schemata auch rekursive Definitionen enthalten, wobei anstelle von Referenzen zu rhetorischen Prädikaten Referenzen auf andere Schemata stehen.

Als erste Anwendung in dem System TEXT hat McKeown durch Analyse gemeinsamer Muster verschiedener Korpora vier Diskursstrategien identifiziert und Schemata dafür ausgearbeitet: *constituency*, *identification*, *attributive*, und *compare and contrast*. Diese Strategien wurden zur Beantwortung von Fragen über die Struktur einer Datenbank ausgearbeitet, wobei jedes Schema für eine oder mehrere Klassen von Anfragen anwendbar ist. Drei Klassen von Anfragen waren vorgesehen: a) Anfragen zur Definition eines Konzepts, b) Anfragen zum Vergleich zweier Objekte und c) Anfragen zur Beschreibung verfügbarer Information.

Das *constituency* Schema etwa, das für Definitionen und Beschreibungen von Objekten anwendbar ist, besteht aus vier Schritten:

1. *Identifiziere* das zu definierende Konzept als einer generischen Klasse zugehörig oder präsentiere *attributive* Information darüber

2. Präsentiere die Bestandteile (*Konstituenten*) dieses Konzepts

3. Präsentiere *charakteristische Information* über seine Konstituenten

4. Präsentiere *attributive* oder *analoge* Information über das Konzept

In dieser informellen Darstellung beziehen sich die kursiv geschriebenen Phrasen auf die Semantik der rhetorischen Prädikate (wie *Identity*, *Constituency*, *Attributive*, und *Analogy*).

Die Instanziierung eines Schemas erfolgt mittels Funktionen, die mit den rhetorischen Prädikaten assoziiert sind. Diese Funktionen wählen Information aus der Wissensbasis aus, die zu dem rhetorischen Prädikat paßt und bildet daraus eine Proposition. Die mit dem Prädikat *Identification* verbundene Funktion etwa liefert für ein zu definierendes Objekt eine Proposition, die das übergeordnete Konzept dieses Objekt, seine definierenden Attribute, sowie dieses Objekt selbst enthält. Diese Instanziierung der rhetorischen Prädikate erfolgt schrittweise für das gesamte Schema, gemäß der Folge von rhetorischen Prädikaten in dem Schema. Die Auswahl zwischen alternativen Fortsetzungen, wenn solche an bestimmten Stellen in einem Schema möglich sind, erfolgt auf der Basis von Kriterien über den *Aufmerksamkeitsfokus*, der sich auf jenes Objekt in einem Satz bezieht, auf das der Äußernde besondere Aufmerksamkeit legt. Zur Umsetzung wurden Bedingungen, wann sich der Aufmerksamkeitsfokus von einem Satz zum nächsten ändern kann, wie sie von [Gro86] identifiziert wurden, durch Prioritäten ausgedrückt. Auf Grund dieser Prioritäten wird versucht, falls möglich, den Aufmerksamkeitsfokus auf ein eben in den Diskurs eingeführtes Objekt zu verschieben, andernfalls wird der aktuelle Aufmerksamkeitsfokus beibehalten und nur, falls dies auch nicht möglich ist, auf einen früheren Fokus zurückgegriffen. Darüberhinaus wird der Aufmerksamkeitsfokus auch herangezogen, um zwischen Alternativen bei der Realisierung der Propositionen auszuwählen.

In der Zwischenzeit wurden die Techniken der Verwendung von Schemata erweitert und verfeinert. Dazu gehören die Entwicklung weiterer Schemata, wie etwa das *process* Schema [Par88], die Anpassung an Benutzerannahmen [Par87], und die Einbettung in Multimediaumgebungen [Fei90].

In unserer Anwendung, Wegbeschreibung, könnten die folgenden zwei Schemata einfache Formen von Domänentexten beschreiben (dabei drücken 0,1 Optionalität, + Wiederholung und | eine Alternative aus), wie in unserem eingangs angeführten Beispiel:

`Schemaname: Gesamtweg`	`Schemaname: Wegabschnitt`
`attributive(Weg)`^{0,1}	`define(Richtung) |`
`describe(Wegabschnitt)`$^{+}$	`define(Fortbewegung)`
`identify(Ziel)`	`attributive(Landmarke)`^{0,1}

Das Schema *Gesamtweg* definiert eine Wegbeschreibung als eine Folge von Wegabschnittsbeschreibungen, der optional ein charakteristisches Attribut des Gesamtwegs vorangeht, und die mit der Identifikation des Zielorts abgeschlossen wird. Die Wegabschnitte selbst, wie in dem Schema *Wegabschnitt* definiert, können durch eine Richtungsänderung oder eine Fortbewegung beschrieben werden, optional ergänzt durch eine Landmarke. Diese Aufteilung erweist sich als pragmatisch leicht handhabbar, weicht aber leicht von der rhetorisch organisierten Textstruktur (Abbildung 2) ab.

Pragmatische Faktoren wie Erfordernisse an Ausführlichkeit und Explizitheit bestimmen die Auswahl bei Optionen in einem Schema bei dieser Anwendung in viel stärkerem Maße als bei den Beschreibungen von Datenbankkonzepten. Die geeignete Verwendung des Schemas *Gesamtwegbeschreibung* etwa gestaltet sich unterschiedlich in Abhängigkeit von verschiedenen Eigenschaften wie Weglänge und Einfachheit. Bei kurzen oder weitgehend geraden Wegen kann man sich auf Wegabschnittbeschreibungen beschränken, während bei längeren Wegen eine kurze Wegcharakterisierung vorweg meist für den Auskunftsuchenden hilfreich ist. Bei unübersichtlichem Ausgangsort oder schwer identifizierbarem Zielort erweisen sich Beschreibungen dieser Orte als erforderlich. Gleiches gilt für die Beschreibung von Landmarken im Verlauf des Wegs bei einer *Wegabschnittbeschreibung*. Es zeigt sich demnach in diesem Anwendungsbereich, daß die Auswahl zwischen Alternativen weitgehend durch Kontext und Konventionen bestimmt sind, während rhetorische Kriterien vorrangig die Komposition der einzelnen Aussagen regeln.

Bei der bereits oben eingeführten Spezifikation von kommunikativer Absicht c und Wissensbasis k könnten diese beiden Schemata in einer Situation d — Frage nach dem Weg von A nach B — zur Erzeugung eines Diskursplans führen, der in späteren Phasen der Generierung zu dem eingangs angegebenen Text umgesetzt werden kann.

Für unser Beispiel wollen wir annehmen, daß der Adressat ein vollständige Beschreibung bei ihm unbekannter Umgebung wünscht. Deshalb erfolgt die Instanziierung der optionalen Wegcharakterisierung ($attributive(Weg)$, das erste rhetorische Prädikat im Schema Gesamtweg) durch *short(path(A,B))*. Das folgende rhetorische Prädikat, $describe(Wegabschnitt)$, ist die Referenz zu einem Schemata, zu dem die Wissensbasis Material für drei Instanziierung enthält (die Wegabschnitte A nach C, C nach D und D nach A). Die Anordnung der daraus erzeugten Propositionen erfolgt am besten chronologisch nach domänenspezifischen Kriterien — dies entspricht auch allgemeinen Fokuskriterien — bei Erreichen des Ziels eines Abschnitts wird ein neuer Fokus eingeführt, der im folgenden Abschnitt durch den gleichen Start beibehalten wird. Die einzelnen Abschnitte können dann unterschiedlich behandelt werden, gemäß den Alternativen der Definition des Schemas Wegabschnitt. Der erste Abschnitt (von A nach C) wird durch eine Fortbewegung beschrieben, zusätzlich mit einer Landmarke versehen (die Kirche). Die nächsten beiden Abschnitte werden zusammengefaßt — es findet bei D keine Richtungsänderung statt — und durch die Richtungungsänderung bei C sowie eine daraufNfolgende Beschreibung der Landmarke D realisiert. Diese Zusammenfassung erfolgt opportunistisch während der Schemainstanziierung und verbessert wesentlich die Qualität der Textspezifikation durch inhaltliche Kompaktheit und Variation des Ausdrucks. Schließlich wird die Option zum Ausdrücken des letzten Prädikats im Schema, Zielcharakterisierung, wahrgenommen. Die Instanziierungen der Schemata sehen dann wie folgt aus:

Instanziiertes Schema	rhetorisches Element
`short(path(A,B))`	`attributive(path)`
`go(User,from(A),to(church))`	`define(movement)`
	`attributive(landmark)`
`turn(User,at(church),direction(left))`	`define(direction)`
`pass(User,by(big & building & town-hall))`	`attributive(landmark)`
`be(User,at(B))`	`identify(goal)`

Eine geeignete sprachliche Umsetzung der Folge von Prädikationen in dem instanziierten Schema könnte dann zur Erzeugung des eingangs angeführten Beispieltexts führen. Dabei ist zu beachten, daß die Prädikationen selbst nur eine Folge von Aussagen darstellen, während sich die Relationen zwischen ihnen, die die Gestaltung des Texts mitbestimmen, nur in den Schemadefinitionen manifestieren.

4.2 Schrittweise Planung mit Diskursrelationen

Eine alternative Möglichkeit zur Beschreibung des gesamten Aufbaus kohärenter Diskurse, wie sie durch Schemata definiert wird, besteht in der formalen Beschreibung ihrer elementaren Teile, der Diskursrelationen. Diskursrelationen verbinden zwei oder mehrere Textsegmente zu einer Einheit, die wiederum mittels einer Diskursrelationen in Bezug gesetzt werden kann zu anderen elementaren oder zusammengesetzten Textsegmenten. Eine in dieser Weise zusammengesetzte Struktur bildet dann einen Diskursstrukturbaum, wobei die inneren Knoten und insbesondere die Wurzel Diskursrelationen repräsentieren, und die Endknoten elementare Aussagen (Sprechakte). Um zu gewährleisten, daß ein solcher Baum kohärent aufgebaut ist, müssen an die Diskursrelationen Bedingungen geknüpft werden, die die Zusammensetzung von Textteilen regeln, in Abhängigkeit von der Art der Relation. So wird zum Beispiel gefordert, daß die einzelnen Textteile, die mittels einer Relation SEQUENCE zu einer zeitlichen Sequenz zusammengebaut werden sollen, in der ihrer Position in der Sequenz entsprechenden chronologischen Ordnung stehen.

Zur Beschreibung des Aufbaus und anderer Eigenschaften von Diskursen sind in der Linguistik eine Reihe von Theorien vorgeschlagen worden, die alle die im vorigen Absatz beschriebenen grundlegenden Konzepte (Diskursrelationen, rekursiver Aufbau eines Diskursstrukturbaums, Kohärenzkriterien) beinhalten und sich in subtilerer Weise etwa bei der Taxonomie der Diskursrelationen voneinander unterscheiden. Eine der bekanntesten dieser Theorien, die in der Diskursplanung häufig eingesetzt wird, ist die Rhetorical Structure Theory (RST) [Man91], auf die wir uns in der Folge beziehen wollen, wobei die beschriebenen Techniken jedoch auch allgemeiner einsetzbar sind.

In RST drücken Diskursrelationen Bezüge zwischen aneinander grenzenden Textsegmenten aus. Manche dieser Bezüge beschreiben untereinander gleichberechtigte Relationen, (*Multi-Nucleus* Relationen), während die meisten Relationen Bezüge zwischen einem dominanten Textsegment (dem *Nucleus*) und einem untergeordneten (dem *Satellite*) festlegen. Die Bedeutung der Relationen wird durch ihren Effekt im Diskurs und durch Bedingungen an den Inhalt von Nucleus und Satellit angeben. Die Relationen können in hierarchischen Taxonomien organisiert werden, je nach Spezialisierung der mit den Relationen verbundenen Bedeutung (siehe auch [Mai93]).

Um Diskursrelationen zu operationalisieren, werden sie als Planoperatoren über Adressateninferenzen und Sprecherzielen ausgedrückt, so daß das Zusammensetzen mehrerer Textsegmente damit gesteuert werden kann. Anforderungen an Nucleus und Satellit werden als semantische Bedingungen an die zu kombinierenden Inhalte formuliert. Außerdem können die an Stelle von Nucleus oder Satellit befindlichen Textsegmente durch Einfügen zusätzlicher Inhalte mittels anderer Relationen erweitert werden (diese Stellen werden *growth points* genannt).

Die häufig vorkommenden Relationen SEQUENCE, ELABORATION, RESULT und BACKGROUND können als Operatoren wie folgt ausgedrückt werden (N steht für das Nucleus- und S für das Satellit Textsegment):

Relation	SEQUENCE ELABORATION RESULT BACKGROUND
Effekt:	Der Sprecher und der Hörer glauben beide, daß obige Relation zwischen den Eingabeentitäten besteht (z.B., daß N zeitlich/räumlich vor S kommt oder daß S Detailinformation zu N angibt)
Bedingungen: (Nucleus und Satellit)	Es besteht folgende Relation zwischen den Eingabeentitäten: N kommt zeitlich/räumlich vor S (Sequenz) S ist eine Detailinformation über N (Elaboration) S ist eine Konsequenz von N (Result) S ist Hintergrundinformation zu N (Background)

Diese Operatoren können nun zur schrittweisen Planung von Diskursen verwendet werden. Dies kann in unterschiedlicher Weise erfolgen, je nach Art des Vorliegens der zu vermittelnden Information. In für den Einsatz von Planungsverfahren mit Diskursstrukturrelationen typischen Umgebungen liegen Informationen vor in Form einer Menge von elementaren Aussagen. Durch Testen anwendbarer rhetorischer Relationen, jeweils auf Paare dieser Aussagen, können bei erfüllten Bedingungen einer Relation schrittweise umfangreichere Strukturen bottom-up aufgebaut werden. Eine andere Vorgehensweise besteht in der expliziten Vorgabe eines Ziels, das durch eine Aussage erfüllt werden kann. Durch Anwendung der Operatoren wird dieses Ziel dann solange zerlegt, bis die einzelnen Teilziele durch elementare Aussagen erfüllt werden können. Diese Strategie ist auch für unsere Anwendung die geeignetere, da die elementaren Aussagen über Wegabschnitte erst gebildet werden

```
                           BACKGROUND
                               /\
                           n  /  \  s
                             /    \
                      EVIDENCE   Describe(property(path))
                           / \
                       n  /   \  s
                         /     \
                  SEQUENCE    Describe(at(destination))
                       / \
                   n  /   \  s
                     /     \
            ELABORATION   Describe(Pass(D))
                   / \
                n /   \  s
                 /     \
Describe(Path(A,C))   Describe(left-turn(at(C)))
```

Abbildung 4: Textstrukturbaum mit rhetorischen Relationen für den Beispieltext

müssen, durch Zerlegen des gegebenen Diskursziels, die Beschreibung des Gesamtwegs. Dies erfolgt in gewissem Sinn hierarchisch, wobei zunächst eine Grobstruktur der Wegbeschreigung gebildet und später in ihren Teilen verfeinert wird. Dazu können Teile der Beschreibungen in einzelnen Abschnitten, gebildet mittels mehrfacher Anwendung des SEQUENCE-Operators, ergänzt werden, etwa mittels einer ELABORATION durch Bezug auf Richtungsänderungen oder Landmarken und zusätzlich können Erläuterungen über den Weg als Gesamtes zum besseren Verständnis integriert werden, wie BACKGROUND Information oder eine Anmerkung über das Erreichen des Ziels als EVIDENCE.

Konkret wird zunächst das ursprüngliche Ziel, "Beschreibe den Weg von A nach B" näher untersucht. Wenn es sich hierbei nicht um einen genügend kurzen Abschnitt handelt (ein anwendungsspezifisch auswertbares Kriterium), der einfach beschrieben werden kann, wird dieses Ziel in eine Sequenz zerlegt, und zwar in die Teilziele "Beschreibe den Weg von A nach C" und "Beschreibe den Weg von C nach B". Da beide Abschnitte selbst keine Richtungsänderung beinhalten, wollen wir davon ausgehen, daß sie nicht weiter zerlegt werden. Die Beschreibungen der Teilabschnitte können nun, durch

Eigenschaften der Wegabschnitte bestimmt, als Fortbewegung ("Sie gehen von hier bis zur Kirche" für $path(A, C)$) und ("Dann kommen Sie an einem großen Haus, dem Rathaus, vorbei" für $path(C, B)$) verbalisiert werden. Der erste Abschnitt sollte, wegen der anschließenden Richtungsänderung, durch diese ergänzt werden, realisiert durch eine ELABORATION. Dadurch wird der Diskursstrukturbaum an einem seiner Blattknoten erweitert. Schließlich kann Einfügen von EVIDENCE (Erreichen des Ziels) und BACKGROUND (Charakteristische Eigenschaft des Gesamtwegs), Information in Bezug auf die gesamte Wegstrecke hinzugefügt werden. Zum Unterschied von der vorigen Erweiterung erfolgen diese Einfügungen an inneren Knoten des Diskursstrukturbaums, im speziellen Fall sogar an der Wurzel. Der somit entstandene Diskursstrukturbaum sieht dann wie in Abbildung 4 aus. Diese Struktur ist eine notationelle Variante des als Merkmalsstruktur angegebenen Textplans im vorigen Kapitel, wobei hier allerdings an den Blattknoten Teilzielspezifikationen anstelle von Satzteilspezifikationen stehen.

Die elementaren Teilpläne an den Blattknoten werden wie an einzelnen Fällen oben beschrieben, durch Satzteile (Haupt- oder Nebensätze) ausgedrückt. Zur kohärenten Verbindung dieser Teile werden die rhetorischen Relationen herangezogen, sowohl zur Entscheidung zwischen Haupt- und Nebensatz und zur Wahl von Konnektiven. Die Relation SEQUENCE wird beispielsweise durch Verbindung von zwei Hauptsätzen mit "und" oder durch zwei getrennte Hauptsätze realisiert, wobei der zweite mit "dann" eingeleitet wird.

Darüberhinaus gibt es einige rhetorisch motivierte Kriterien der kontextuell geeignete Realisierung der einzelnen Textsegmente [Sco90]. Diese Kriterien enthalten generelle und eigentlich offensichtliche Regularien, wie Nähe der durch eine rhetorische Relation verbundenen Satzteile, bevorzugte Aggregation von Propositionen je größer die Anzahl gemeinsamer Konstituenten ist, oder Vermeidung von Ambiguitäten, deren Einhaltung allerdings nicht immer leicht in einem formalen Generierungssystem zu erkennen ist. Andere Kriterien sind leichter formal zu überprüfen, etwa daß Einbettungen auf ELABORATION Relationen beschränkt bleiben sollen, wobei jedoch der Satellit, der eingebettet wird, wiederum keine LIST Relation sein sollte. Wie bereits an früherer Stelle angedeutet, machen diese Kriterien deutlich, daß sie zwar notwendig sind in dem Sinne, daß sie einige offensichtlich unerwünschte Realisierungen ausschließen, aber nicht hinreichend um wirklich qualitativ gute Ausdrucksvarianten zu ermöglichen.

4.3 Eignung der Verfahren

Im Gebiet der Sprachgenerierung wurden Textschemata und Planen mit Diskursrelationen einige Zeit als alternative Verfahren in einer gewissen Konkurrenzsituation angesehen. Mittlerweile hat sich die Ansicht durchgesetzt, daß Schemata in gewissem Sinne auskompilierte Zusammensetzungen von Diskursrelationen sind.

In Bezug auf die Aufgaben der Textplanung leisten Schemata und Planung mit Diskursrelationen vorrangig die Organisation von zu präsentierenden Inhalten, durch die Reihenfolge der im Laufe der Schemainstanziierungen gebildeten Propositionen. Bezüglich der Inhaltsauswahl sind beide Verfahren in gewissem Sinne neutral. Die Alternativen, Optionalitäten und Wiederholungen in den Schemadefinitionen bzw. die partielle Reihenfolge der Aussagen in den Diskursstrukturbäumen lassen im Prinzip verschiedene Teilmengen der maximal auf diese Definitionen passenden Inhalte der Wissensbasis zu. Die implizite Defaultstrategie bei der Schemainstanziierung ist jedoch die Auswahl aller mit den Schemaspezifikationen matchenden Elemente der Wissensbasis. Ähnliches gilt für Planen mit Diskursrelationen. Eine echte Auswahl kann demnach entweder durch eine Vorselektion von Teilen der Wissensbasis oder durch zusätzliche, eventuell kontextabhängige Selektionsmechanismen bei der Schemainstanziierung erfolgen.

In noch stärkerem Maß als bei den Schemata gibt es eine große Zahl von Möglichleiten, wie eine bestehende Diskursstruktur erweitert werden kann. [Mar97] beschreibt ein rein technisches Verfahren, alle Lösungen zu generieren. Die in einer Anwendung gewählten Varianten in den Schematas bzw. in zur Erweiterung einer partiellen Diskursstruktur oder zum Weglassen von Information sind oft mehr durch Adressateneigenshaften und Domänenkonventionen bestimmt als durch generelle rhetorische Kriterien [Kit91]. Die Reihenfolge bei Schemainstanziierungen und Textplanexpansionen erfolgt am besten opportunistisch, da mit größere Strukturen mehr Information über Eignungskriterien enthalten. Praktische Textplanungsverfahren arbeiten jedoch meist strikt top-down, und chronologisch bei Schemata.

Generell erwiesen sich jedoch Schemata durch die explizite Festlegung der möglichen Alternativen und des Gesamtaufbaus für manche Verwendungen als zu unflexibel, da sie es nicht erlauben, die ihrem Aufbau und entsprechenden Instanziierungen zu Grunde liegenden Absichten darzustellen und darüber zu Schlüsse zu ziehen. Die generell, allgemein akzeptierte Einschätzung der beiden Verfahren ist, daß inkrementelles Planen mit Diskurs flexibler ist als der Einsatz von Schemata, aber auch schwieriger zu kontrollieren. In der Praxis kann man manchmal kombinierte Ansätze finden, wobei Schemata für weniger variantenreiche Textteile eingesetzt werden und inkrementelle Planung für die variantenreicheren. Ein typisches Beispiel für eine solche Mischform sind standardisierte Geschäftsbriefe.

4.4 Annahmen von Planungsverfahren

RST und darauf basierende Planungsverfahren treffen einige vereinfachende Annahmen, die die Ausdrucksmöglichkeiten beschränken und für stereotype Erscheinungsformen automatisch erzeugter Texte mitverantwortlich sind.

- Textsegmente sind immer Satzteile (Haupt- oder Nebensatz), aber keine Nominal- oder Präpositionalphrasen
- Es besteht die Forderung nach (maximal) einer einzigen, prominenten Relation zwischen je zwei Propositionen
- Alle Spezifikationen werden entsprechend der vorliegenden Textstruktur explizit ausgedrückt

Die erste Annahme ist weniger linguistisch, sondern nur pragmatisch motiviert (laut Bill Mann, dem Mitautor von RST). Eines der wenigen Verfahren, das versucht NPs, PPs und Satzteile als mögliche Varianten bei der Realisierung von Textsegmenten zuzulassen, ist [Dob93]. Dabei erhöht sich jedoch der Aufwand, Auftreten von "generatio gaps" bei der Kombination von Phrasen zu vermeiden.

Die zweite Annahme der Planungsverfahren ist unter Umständen problematisch. Es können jedoch nicht nur mehrere solcher Relationen bestehen — etwa eine oder mehrere informationelle (d.h., semantische), wie SEQUENCE und CAUSE, zusätzlich zu einer intentionalen wie MOTIVATION. Manchmal sind sogar die zugrunde liegenden informationellen und intentionalen Strukturen unterschiedlich. In dem von [Moo92] angegebenen Beispiel

```
a) Come home by 5:00.
b) Then we can go to the hardware store before it closes.
c) That way we can finish the bookshelves tonight.
```

ist, aus intentionaler Sicht, b) eine MOTIVATION für a), und c) wiederum eine MOTIVATION für b). Aus informationeller Sicht sieht es hingegen genau umgekeht aus: a) ist eine CONDITION für b), und b) ist wiederum eine CONDITION für c). Während sich zwar die Ansicht durchgesetzt hat, daß im Prinzip beide Repräsentationen notwendig sind [Mos96], ist es nicht klar, ob innerhalb eines Systems beide Strukturen explizit aufgebaut werden müssen.

Bezüglich der dritten Annahme gibt es nur einige wenige Ansätze, die Strukturänderungen behandeln zwischen rhetorischen Relationen aus der Wissensbasis (oder entsprechend den Informationen der Wissensbasis) und den Relationen der Propositionen in der aufgebauten Textstruktur. Dies tritt oft auf im Zusammenhang mit Aufzählungen. Dabei können mehrere Propositionen, die in einer bestimmten Relation zu einer anderen stehen (etwa PART-WHOLE oder ELABORATION) zu einer PRESENTATION SEQUENCE zusammengefaßt werden [Mai95]. Andererseits wird eine Expansion eines Elements einer Aufzählung bevorzugt an entfernterer Stelle genannt (vor oder nach der Aufzählung) um eine PRESENTATION SEQUENCE kompakt zu halten [Hor98], wobei der Skopus dieser Expansion unpräzise wird und inhaltlich erschlossen werden muß. In inferenzreichen Domänen kommt es häufig vor, daß einzelne Sachverhalte implizit bleiben [Hor98]. So muß der Adressat bei der Äußerung *a) Some extra copies of the Spring 1984 issue of AI Magazine are available in the library. b) This issue includes a "Research in Progress" report on AI research at ISI.* erkennen, daß b) als MOTIVATION zu a) gemeint ist. Um die Äußerung in dieser Form zu generieren, muß ein formales System allerdings über eine detailliertere Wissensbasis verfügen, da der Adressat weder in a) noch in b) explizit genannt ist, die Semantik von MOTIVATION jedoch einen Bezug zu dem Adressaten erfordert. Deshalb muß aus der Wissensbasis ein solcher Bezug herstellbar sein, etwa daß der Adressat die Ausgabe des AI Magazines kaufen und lesen soll. Das Problem aus Generierungssicht besteht dann darin, diese Propositionen in der Präsentation implizit zu lassen, da eine vollständig explizite Äußerung als redundant empfunden wird. In welchen Situationen jedoch der Redundanzaspekt dominiert, und in welchen die Erschließbarkeit der Information schwierig ist, bedarf einer ausgiebigen Domänenmodellierung und ist bisher nur in einigen Experimentalsystemen behandelt worden.

Die Techniken der Planung mit Diskursrelationen sind schließlich auch auf Multimediaumgebungen erweitert worden [And94]. Dabei können die mit den Blattknoten assoziierten kommunikativen Teilziele entweder durch Textteile oder durch graphische Elemente realisiert werden, je nach Eignung der Medien. In der Anwendung zur Beschreibung der Funktionalität einer Espressomaschine etwa wird BACKGROUND Information zur Lokalisierung von Teilen der Maschine durch einen Bildausschnitt angegeben, während ein Hinweis zur Wassertemperatur sprachlich ausgedrückt wird. Die Koordination der Medien erweist sich dabei als eine wichtige Aufgabe, die mit cross-medialen Referenzausdrücken unterstützt wird.

5 Satzplanungsverfahren

In diesem Teilkapitel werden kurz die wesentlichsten Verfahren zur Bewältigung von Aufgaben der Satzplanung vorgestellt. Die gewählte Reihenfolge impliziert dabei keineswegs einen präferierten Ablauf in einem umfassenden Satzplanungsprozeß.

5.1 Lexikalisierung

Für die in den Propositionen des Textplans aufscheinenden Prädikate müssen lexikalische Elemente der Zielsprache (Lexeme, grammatische Funktionen und Features) gewählt werden, die den Informationsgehalt dieser Prädikate ausdrücken. Dieser Vorgang, die Lexikalisierung im engeren Sinn, muß in einer kompositionalen Weise erfolgen, so daß die einzelnen lexikalischen Elemente eine korrekte Struktur in der Zielsprache bilden, und die lexikalischen Ausdrücke müssen dem pragmatischen Kontext angemessen sein. Dies ist in Bezug auf Kompositionalität oft nicht einfach, da die Abbildung zwischen Prädikaten im Textplan und lexikalischen Elementen nicht immer eine 1:1-Abbildung ist. Weiters müssen auch einzelsprachspezifische Besonderheiten beim Zusammentreffen von Wörtern mit kontextabhängiger Bedeutung berücksichtigt werden (sogenannte *Kollokationen*). Zum Beispiel wird im Deutschen ein Mensch, der viele Zigaretten konsumiert, als 'starker Raucher' bezeichnet; die Entsprechungen des Adjektivs differieren jedoch sowohl im Englischen 'heavy (wörtlich: schwer) smoker' als auch im Französischen 'grand (wörtlich: groß) fumeur'.

Je nach Reichhaltigkeit der Entsprechungen zwischen Prädikaten im Textplan und lexikalischen Elementen stehen oft mehrere Alternativen bei der Lexikalisierung zur Verfügung. Die Auswahl zwischen diesen wird durch verschiedene Einflußfaktoren gesteuert:

- *Konnotationen* von Wörtern müssen berücksichtigt werden. So ist etwa das Bevorzugen von 'Köter' gegenüber dem neutralen 'Hund' nur dann gerechtfertigt, wenn eine negative Einstellung des Sprechers gegenüber dem Tier ausgedrückt werden soll. Ähnliches gilt für die bevorzugte Verwendung von Basiskategorien — 'Pudel' wird nur dann statt 'Hund' verwendet, wenn die Information über die Rasse in der Situation wesentlich ist.

- Die *Perspektive* auf Sachverhalte kann durch die Wortwahl verändert werden. Mit der Wahl von 'kaufen' bzw. 'Kauf' wird die Perspektive des Erwerbs auf den Käufer gelegt, mit 'verkaufen' bzw. 'Verkauf' auf den Verkäufer.

- Ein kompensatorischer Effekt besteht zwischen *Wiederholungen* und *knappen Ausdrucksformen*. Während die erste Variante auf verbesserte Verständlichkeit abzielt, ermöglicht die zweite kürzere Texte. Manche Formen von Wiederholungen können in Laufe weiterer Verarbeitung kompakter ausgedrückt werden (durch Aggregierung, siehe nächstes Teilkapitel).
- Ein generelles Kriterium für die Auswahl zwischen Lexemen bilden die Eigenschaften der *kontextuellen Einbettung*, die den Adressaten der Äußerung und die Stellung im Diskurs umfassen. Dazu gehören die Vermeidung von Wörtern, die dem Adressaten nicht bekannt sind, und Konsistenz mit textuellen Eigenschaften des vorangegangenen Diskurses.

Verfahren zur Lexikalisierung sind meist musterbasierte Abbildungen zwischen zwei Repräsentationsebenen, den konzeptuellen Beschreibungen aus Ausgangsspezifikationen (hier: dem Textplan) und den lexikalischen Strukturen als Ziel, wie in [Nog91]. Dabei sind viele Entsprechungen elementweise definiert, manchmal wird eine Zusammensetzung von konzeptuellen Elementen auf ein Lexem abgebildet (etwa bei semantisch komplexen Wörtern) oder ein konzeptuelles Element auf eine sprachliche Struktur (etwa bei idiomatischen Ausdrücken). Die Kompositionalität ist auf den jeweiligen Reprs̈entationsebenen definiert; zusammenhängende Strukturen auf der konzeptuellen Ebene werden ohne Überlappungen und Lücken zerlegt, so daß zur Abbildung jedes dabei entstehenden Teils mindestens ein Muster definiert ist. Die Bilder dieser Muster auf der Seite der Zielsprache müssen dann gemäß semantischen und grammatischen Erfordernissen zusammensetzbar sein; das heißt unter anderem, Selektionsrestriktionen und Typbedingungen von Phrasenköpfen müssen durch die jeweils abhängigen Strukturen erfüllt sein. Falls dies an einer Stelle unzutreffend ist, muß eine andere Kombination von Abbildungen auf lexikalische Elementen genommen werden.

Die Regelung der Zusammensetzbarkeit ist gerade bei komplexeren Abbildungen eine Herausforderung. Ein Verfahren mit einem besonders reichhaltigen Repertoire an Ausdrucksmöglichkeiten, insbesondere bei geschachtelten Nominal- und Adjektivphrasen ist [Elh97]. [Hor96] zeigt, wie metonymische Relationen je nach Zielsprache implizit oder explizit ausgedrückt werden können. [Ste96] ist in der Lage, recht unterschiedliche Aufteilungen des Informationsgehalts auf die gewählten Lexeme vorzunehmen. Dadurch kann etwa aus derselben konzeptuellen Repräsentation das im Englischen bessere 'he swam across the river' (er schwamm über den Fluß) und das im Französischen bessere 'il a traverse la riviere a la nage' (er hat den Fluß schwimmend überquert) erzeugt werden.

Bei dem Beispiel zur Wegbeschreibung liegen solch komplizierte Fälle nicht vor. Die konzeptuellen Prädikate werden im wesentlichen auf geeignete Lexeme 1:1 abgebildet. Eine Ausnahme ist das Prädikat 'left-turn', das durch das Verb 'abbiegen' zusammen mit der Ergänzung 'links' ausgedrückt wird.

5.2 Aggregation

Unter Aggregation versteht man im Rahmen der Sprachgenerierung das Erzeugen von Satzrepräsentationen aus satzteilartigen (englisch: 'clauses') semantischen Repräsentationen. Dabei werden insbesondere Koordinationsphänomene berücksichtigt, so daß Redundanzen, die sich durch mehreren semantischen Repräsentationen gemeinsame Teile ergeben, kompakt ausgedrückt werden können. Relevante Phänomene sind die Koordination von Konstituenten und verschiedene Arten von Konstituententilgungen. Koordination von Konstituenten kann erfolgen, wenn sich zwei Aussagen nur in einem der beteiligten Objekte unterscheiden, wie in der Erzeugung des Satzes "Hans mag Maria und Anna" aus den Aussagen "Hans mag Maria" und "Hans mag Anna". Dabei ist jedoch eine mögliche Interferenz mit lexikalischer Semantik bei bestimmten Verben zu berücksichtigen, da in Sätzen wie "Hans und Maria gehen auf eine Party" die kollektive Lesart präferiert wird, die unter Umständen nicht den getrennten Aussagen der Ausgangsrepräsentation entspricht. Bei Konstituententilgungen werden nach Van Oirsouw [vO87] einige Regeln für Tilgungen bei Gleichheit unterschieden, wie *Gapping* in "Hans ißt Fisch und Willi ϕ Reis", *Rechtsknotentilgung* in "Hans fing ϕ und Willi aß den Fisch", *Verbphrasentilgung* in "Hans ißt Fisch und Willi ϕ auch", *Konstituentenreduktion* in "Hans ißt Fisch zu Mittag und ϕ Reis am Abend". Ein automatisches Verfahren muß somit nicht nur Aussagen geeignet zusammenfügen, sondern beim Durchführen der Tilgungen diese sprachlichen Phänomene berücksichtigen, was existierende Systeme nur zum Teil tun.

Ein Verfahren, das einige Aggregationsregeln basierend auf einer kleinen empirischen Studie erstellt hat, wird in [Dal96] beschrieben. Wir wollen in diesem Zusammenhang ein neueres, systematisches Verfahren etwas näher betrachten, das die meisten der oben genannten Phänomene berücksichtigt [Sha98a, Sha98b]. Das Verfahren setzt sich aus vier Phasen zusammen:

1. Zunächst werden die Propositionen, die die semantischen Repräsentationen beinhalten, sortiert, wobei allerdings pragmatische und kontextuelle Einschränkungen, wie rhetorische Relationen zwischen einzelnen Aussagen, berücksichtigt werden. Die Sortierung erfolgt nach Graden der Ähnlichkeit der Propositionen, beginnend mit jenem Argument, das die kleinste Anzahl von Werten in allen Propositionen hat. Dadurch

wird eine Sortierung erreicht, bei der sich zwei unmittelbar aufeinander folgende Propositionen möglichst wenig unterscheiden.

2. Als nächstes werden sequentiell jeweils zwei aufeinanderfolgende Propositionen verglichen, die identischen Elemente zunächst nur markiert, und daraus eine gemeinsame Aussage gebildet. Dadurch ergibt sich gegenüber anderen Ansätzen, die gleichzeitig Tilgungsoperationen einführen, eine erhöhte Flexibilität.

3. Diese Markierung erfolgt gemäß parameterisierbarer Komplexitätsgrenzen; so werden zwar mehrere Aussagen, die sich nur in den Werten eines Arguments unterscheiden, zu einer koordinierten Aussage zusammengefaßt, während nur maximal zwei Aussagen, die sich in den Werten mehrerer Argumente unterscheiden. Dieses Vorgehen führt zur Festlegung von Satzgrenzen.

4. Schließlich erfolgt die Tilgung redundanter Elemente. Dazu werden syntaktisches Wissen und Information über die Reihenfolge von Konstituenten herangezogen. Dadurch können Fälle von Vorwärtstilgung, wie in "On Monday, Al re-stocked coffee and [on Monday,] [Al] removed wrotten milk" unterschiedlich zu "Al re-stocked coffee [on Monday,] and [Al] removed wrotten milk on Monday" behandelt werden (Beispiele aus [Sha98a]).

In unserer Wegbeschreibung spielen solche Aggregationsoperationen keine große Rolle. Die Festlegung der Reihenfolge der Sätze und der Satzgrenzen ist von Besonderheiten der Domäne beeinflußt. Die Hintergrundinformation wird zuerst ausgedrückt, als eigener Satz. Darauf werden die beiden Aussagen über die gerade Fortbewegung und die Richtungsänderung zusammengefaßt, wobei im zweiten Teil das Subjekt getilgt werden kann. Ebenso wird mit der Aussage über die Landmarke und der Evidenz über das Erreichen des Ziels verfahren. Es zeigt sich hier, daß diese Zusammenfassungen eher opportunistisch durch Aggregationsmöglichkeiten und vom Nutzen der Information — Kenntnis der Weglänge — motiviert sind als von der darunterliegenden rhetorische Struktur.

5.3 Referenzausdrücke

Die Propositionen im Textplan enthalten Referenzen auf Objekte, die mindestens einen von zwei grundlegenden Zwecken erfüllen sollen: *Identifikation* des Objekts (durch eine *referentielle* Beschreibung) bzw. *Information über* ein Objekt (durch eine *attributive* Beschreibung). Für die Realisierung dieser Aufgaben stehen mehrere sprachliche Ausdrucksmittel zur Verfügung, mit unterschiedlicher Eignung je nach Kontext. Dazu gehören Pronomen, Eigennamen, indefinite und definite Beschreibungen, letztere sind auch in reduzierter Form möglich. Bei attributiven Referenzen werden spezifizierte Informationen in sprachliche Ausdrücke überführt und im Zusammenhang mit dem zu beschreibenden Objekt genannt. Dies erfolgt opportunistisch im Kontext des gesamten Texts und ist somit eher eine Aufgabe fr̈ die Lexikalisierung und, bei umfangreicher Information, auch für die Textplanung. Die Generierung referentieller Beschreibungen hingegen erfordert die Erzeugung eines sprachlichen Ausdrucks, von dem plausibel angenommen werden kann, daß der Adressat den intendierten Referenten dadurch eindeutig identifizieren kann. Da dieses Objekt intern einfach durch eine zugeordnete Variable identifiziert ist, erfordert das Finden einer geeigneten referentiellen Beschreibung somit in der Bestimmung zusätzlicher Angaben.

Die einfachste Referenz auf ein Objekt ist jene mit einem Pronomen. Die Entscheidung allerdings, ob eine solche Referenz im Einzelfall adäquat ist oder nicht, kann je nach Kontext sehr heikel und subtil sein. Daher ist für automatische Verfahren ein konservatives Vorgehen wie in [Rei00] zu empfehlen, durch das Pronomen möglichst selten in inadäquaten Situationen erzeugt werden, dafür öfters kein Pronomen gewählt wird, obwohl es in den entsprechenden Situationen gut wäre. Das einfache Kriterium besagt, daß ein Pronomen nur dann zu wählen ist, wenn seine grammatischen Eigenschaften (Person, Numerus und Geschlecht) nur auf den intendierten Referenten und auf kein anderes Objekt im selben oder im unmittelbar vorangegangenen Satz passen.

Für nicht-pronominale referentielle Beschreibungen sind mehrere ähnliche Verfahren angegeben worden, wobei die neueren meist Verfeinerungen und Erweiterungen der grundlegenden Methode in [Dal95] sind. Die Auswahl der Beschreibungselemente erfolgt nach drei komplementären Anforderungen:

- Der Referenzausdruck muß *adäquat* sein in dem Sinne daß er genügend Informationen zur Identifikaiton des intendierten Referenten enthält.
- Der Referenzausdruck muß *effizient* sein, also keine für die Aufgabe unnötige Information enthalten.
- Der Referenzausdruck muß *sensitiv* in Bezug auf die Kenntnisse und Fähigkeiten des Addressaten sein.

Speziell die geeignete Interpretation des Effizienzkriteriums hat die Forscher einige Jahre beschäftigt. Zunächst wurde versucht, möglichst minimale

Beschreibungen zu erzeugen, was sich jedoch als sehr rechenaufwendig herausstellte. Motiviert durch die empirische Erkenntnis, daß Menschen selbst oft redundante Deskriptoren in Referenzausdrücken verwenden, wurden einfachere Verfahren vorgeschlagen, die psychologischen Erkenntnissen besser Rechnung tragen. Das grundlegende Vorgehen ist dabei wie folgt: Es werden die Mengen aller potentiell mit dem intendierten Referenten verwechselbaren Objekte im aktuellen Diskursabschnitt bestimmt. Desweiteren werden alle Deskriptoren des intendierten Referenten gesammelt und nach domänenspezifischen Präferenzen in eine Liste sortiert. Dann wird die zu erzeugende Beschreibung aus Elementen dieser Liste aufgebaut. Dabei wird für jeden Deskriptoren inkrementell geprüft, ob er auf mindestens eines der potentiell mit dem intendierten Referenten verwechselbaren Objekte nicht zutrifft, auf das die bisher erzeugte Beschreibung hingegen paßt. Ein solcher Deskriptor schließt somit die Verwechslung mit mindestens einem anderen Objekt aus und wird daher zu der Beschreibung hinzugefügt. Dieses Verfahren wird solange wiederholt, bis alle potentiell mit dem intendierten Referenten verwechselbaren Objekte ausgeschlossen werden können (der Erfolgsfall) oder keine Deskriptoren mehr zur Verfügung stehen (dann ist mit diesen Mitteln auch keine identifizierende Beschreibung möglich).

In unserem Wegbeschreibungstext kommen mehrere pronominale Referenzen und zwei referenzielle Beschreibungen vor. Dabei sind die pronominalen Referenzen auf die auskunftsuchende Person mit 'Sie' mit dem oben skizzierten, einfachen Verfahren zur Erzeugung pronominaler Referenzen erklärbar, wegen der Eindeutigkeit der grammatischen Person. Die Erzeugung der lokalen Referenzen, 'hier' und 'dort', bedarf spezieller Algorithmen mit Wissen über räumliches Schließen. Um die Anwendung des oben skizzierten Verfahrens zur Erzeugung referenzieller Beschreibungen für die beiden Landmarken zu erklären, die sich in den Ausdrücken 'Kirche' bzw. 'ein großes Haus, das Rathaus' manifestiert, sind einige Annahmen zugrundezulegen. Als Attribute stehen in der Deskriptorenliste die Kategorien der Landmarken, 'Kirche' bzw. 'Rathaus', an vorderster Stelle, gefolgt von Größe als einer der auffälligsten Eigenschaften von Gebäuden. Zur Identifikation der Kirche kann angenommen werden, daß das Nennen der Kategorie zur Identifikation ausreicht — wenn sie die Gestalt einer typischen Kirche hat und sich keine weitere Kirche in der Nähe befindet. Der Fall des Rathauses ist problematischer. Da nicht einfach die Kategorie des Gebäudes genannt wird, ist davon auszugehen, daß der Auskunftssuchende dieses Rathaus nicht kennt. Also wird nur die allgemeinere Kategorie, Haus, genannt, und das nächstauffällige Attribut, die Größe, angeführt. Das weitere Anführen der Bezeichnung 'Rathaus' läßt sich dann nur durch die mit der quantitativen, vagen Beschreibung 'groß' verbundenen Unsicherheit erklären, in der Hoffnung, daß dem Auskunftsuchenden die Identifikation durch Heranziehen von stereotypem Wissen über Rathäuser und hypothetischem Schließen in der lokalen Situation erleichtert wird. Diese Überlegeungen reichen jedoch weit über die konzeptuelle Abdeckung aller existierender Referenzgenerierungsagorithmen hinaus. Es zeigt sich somit, daß die stark abstrahierenden Verfahren für Realweltsituationen mit komplexen Objekten unzureichend sind, vor allem wegen der Nichtmodellierung des Interpretationsspielraums von Attributen und den daraus folgenden Unsicherheiten.

6 Verfahren zur Oberflächenrealisierung

Sobald die Verfahren der Satzplanung angewendet wurden, ist die durch die Textplanung erzeugte Baumstruktur in eine logische Form überführt worden. Dann sind Verfahren zur Oberflächenrealisierung anwendbar, die eine in logischer Form vorliegende Satzsemantik in eine Sequenz von flektierten Wörtern überführt. Es gibt zwei prinzipielle Verfahren für diese Aufgabe: strukturgesteuerte und lexemgesteuerte. Wir betrachten sie der Reihe nach.

6.1 Strukturgesteuerte Oberflächenrealisierung

In strukturgesteuerten Verfahren erfolgt die Realisierung mittels einer deklarativen Grammatik, die im Idealfall unabhängig von der Interpretationsrichtung ist. Wenn die Grammatik allerdings im Stil parsing-orientierter Verarbeitungstechniken einfach top-down oder bottom-up eingesetzt wird, gibt es für die Generierung zwei fundamentale Probleme:

- Linksrekursion kann zu Endlosschleifen führen.
- Das Zusammenspiel zwischen Interpretation von Grammatikregeln und Zugriff auf lexikalische Information kann ohne ausgeklügelte Kontrolltechniken sehr ineffizient werden.

Endlosschleifen treten bei top-down Verarbeitung auf im Zusammenhang mit wechselseitiger Rekursion oder bei bestimmten Konstruktionen wie Abspaltung eines Verbpartikels aus einer VP. Eine solche Regel (Abspaltung eines Verbpartikels) ist linguistisch sinnvoll, sie könnte jedoch mehrmals hintereinander angewendet werden, was zu einer nicht intendierten Übergenerierung führt. Das grundlegende Problem besteht darin, daß die Wohlfundiertheit des Generierungsprozesses auf lexikalischer Information beruht, die für top-down Verfahren nicht zugänglich ist.

Im Gegensatz zu der top-down Verarbeitung haben bottom-up Verfahren kein Problem mit Endlosschleifen. Sie leiden jedoch bei der typischen Abarbeitung der rechten Seite der Grammatikregeln, von links nach rechts, an Ineffizienz, da diese Stratigie an der bei der Analyse vorhandenen Information, dem Eingabezeichenkette, orientiert ist, und nicht an der logischen Form, der Ausgangsrepräsentation bei der Generierung. Daraus ergeben sich eine Vielzahl von Indeterminismen, zum Beispiel bei der Generierung einer Nominalphrase in Bezug auf Kasus und Numerus, da diese erst vom nachfolgenden Verb festgelegt werden. Deshalb wurden Verarbeitungsstrategien entwickelt, die sich an der semantischen Struktur orientieren, die sogenannten kopf-gesteuerten Generierungstechniken [Shi90]. Sie basieren auf einer systematischen Gemeinsamkeit der logischen Form zwischen Mutterknoten und einem der Tochterknoten, dem *Pivot*, und setzen semantisch monotone Grammatiken voraus. Im wesentlichen werden zunächst top-down Pivot Knoten expandiert, solange bis ein Zugriff auf lexikalische Information erfolgen kann. Dann werden sie, soweit erforderlich, bottom-up mit Mutterknoten verbunden. Schließlich werden mit der nun verfügbaren lexikalischen Information, top-down alle nicht-Pivot Knoten expandiert. Dieses Verfahren löst zwar das Effizienzproblem, trifft dafür aber starke Annahmen über die Grammatik.

6.2 Lexikalisch gesteuerte Oberflächenrealisierung

Im Gegensatz zu den strukturgesteuerten Verfahren ist die Eingabe für die lexikalisch gesteuerten keine voll strukturierte logische Form, sondern eine Sammlung lexikalischer Elemente, wobei die semantischen Relationen durch geeignet instanziierte Variable ausgedrückt sind. Dies entspricht inhaltlich weitgehend einer logischen Form, nur die Skopusbeziehungen sind wegabstrahiert. Verschiedene Algorithmen sind zur Erzeugung von Sätzen aus solchen Strukturen vorgeschlagen worden, mit unterschiedlichem Erfolg in Bezug auf Effizienz. Sie reichen von 'shake-and-bake' Generierung [Whi92], über Chart Generierung [Kay96], bis zu einem semilexilaischen [Car99], dem neuesten und effizientesten Verfahren.

'Shake-and-bake' Generierung kombiniert die lexikalischen Elemente der Eingabe in jeder möglichen Reihenfolge und testet die grammatische Wohlgeformtheit der jeweiligen Ergebnisse mit einem shift-reduce Parser. Da dieses Verfahren keinerlei Information über die Eingabesemantik ausnutzt, ist sein exponentielles Verhalten nicht verwunderlich.

Chart Generierung adaptiert ein aus dem Parsing als effizient bewährtes Verfahren und vermeidet die wiederholte Berechnung von Kombinationen lexikalischer Elemente. Zunächst werden alle Lexikoneinträge bestimmt, die von der Eingabe subsumiert werden, auf die Agenda gesetzt, von wo sie nacheinander in die Chart eingetragen werden. Dabei werden Interaktionen zwischen Einträgen der Chart berücksichtigt, das heißt, es werden Zusammensetzungen von Charteinträgen gebildet, die einer Grammatikregel entsprechen, und sie werden in die Chart eingefügt. Ein erfolgreicher Abschluß des Verfahrens ist dann sichergestellt, wenn eine Zielstruktur, meist ein Satz, gebildet wurde, und dabei alle Elemente der Eingabe genau einmal verwendet wurden. Chart Generierung erzielt signifikante Verbesserungen gegenüber 'Shake-and-bake' Generierung, hat aber Effizienzprobleme mit Kombinationen von intersektiven Modifikatoren. Auch die von Kay vorgeschlagene Verbesserung, daß nur maximale Strukturen mit anderen kombiniert werden, hilft nur teilweise, da dadurch das Bilden aller Teilmengen von Modifikatoren nicht vermieden werden kann.

Das Problem der Chart Generierung mit intersektiven Modifikatoren führte zu der Verbesserung in dem semilexikalischen Verfahren. Dieses ist ein zweistufiges Verfahren, in dem zunächst Chart Generierung ohne intersektive Modifikatoren durchgeführt wird, und diese anschließend mit Adjunktion sukzessive eingefügt werden. Die Korrektheit dieses Vorgehens ergibt sich daraus, daß Adjunktion von intersektiven Modifikatoren höchstens zu einer Spezialisierung, aber nie zu einer Änderung syntaktischer Kategorien führt, und somit durch eine solche Adjunktion niemals zusätzliche Grammatikregeln anwendbar werden. Bei der Adjunktion von Modifikatoren werden Bedingung über ihre Reihenfolge berücksichtigt, und Wahlmöglichkeiten bezüglich der Reihenfolge werden kompakt durch Pointerstrukturen ausgedrckt, so daß Teilstrukturen nicht mehrfach erzeugt werden müssen.

Das semilexikalische Verfahren wurde für umfangreiche Grammatiken des Englischen getestet und zeigte signifikate Laufzeitverbesserungen gegenüber reiner Chart Generierung, zwischen einem Faktor drei für normale Sätze, bis zu Faktoren größer als zehn für Sätze mit untypischer Anzahl intersektiver Modifikatoren. Der Algorithmus ist sehr geeignet für die Verarbeitung lexikalisierter Grammatiken wie HPSG, und macht nur geringe Annahmen über die Grammatik, die weit weniger restriktiv sind als die Forderung nach einem semantischen Kopf in strukturgesteuerten Verfahren.

7 Linguistische Theorien zur Generierung

Einige vornehmlich funktional orientierte linguistische Theorien eignen sich besonders zur Generierung. Dazu gehören unter anderem Baumadjunktionsgrammatiken, die an anderer Stelle in diesem Band beschrieben werden. Sie werden wegen ihrer Flexibilität für oberflächennahe Aufgaben in der Generierung vielfach eingesetzt. Zwei speziell für die Generierung geeignete Theorien, Meaning Text Theorie und Systemisch Funktionale Grammatik, werden in der Folge kurz vorgestellt.

7.1 Meaning Text Theorie

Meaning-Text Theorie (MTT) ist ein extrem stratifiziertes Sprachmodell, wobei die semantische, die tiefe syntaktische, und die oberflächennahe syntaktische Repräsentationsebenen in Sprachen wie Englisch die wichtigsten Strata sind. Zentral für MTT ist die Rolle des Lexikons, mit einer speziellen Eigenschaft, den *lexikalischen Funktionen* [Mel82], das sind Pointer von einem Lexikoneintrag zu anderen Wörtern, die zu diesem Wort semantisch oder kollokational in Beziehung stehen.

Die theoretischen Konzepte von MTT sind in dem 'Explanatory Combinatorial Dictionary' [Mel87] implementiert, dessen Einträge in *semantische*, *syntaktische*, und *lexikalische* Zone untergliedert sind. Die semantische Zone eines Lexikoneintrags spezifiziert ein semantisches Netzwerk, das die mit diesem Eintrag assoziierte Bedeutung in Form von nächsteinfacheren Wortbedeutungselementen ausdrückt. Auf der Basis dieser Definitionen können große semantische Netzwerke inkrementell reduziert werden, indem einzelne Knoten, die komplexe Bedeutungen repräsentieren, für Teilnetze substituiert werden, die die Bedeutung eines solchen Knoten mit einfacheren Bedeutungselementen repräsentieren. Die syntaktische Zone spezifiziert die syntaktische Kategorie eines Lexikoneintrags, seine syntaktischen Eigenschaften, sowie die Art der Abbildung der semantischen Rollen des Eintrags auf tiefe und oberflächennahe syntaktische Repräsentationen. Die lexikalische Zone spezifiziert Bezüge zu semantisch verwandten Lexemen als Werte von etwa 60 lexikalischen Funktionen. Dazu gehören 'paradigmatische' Funktionen, die Verbindungen zu verwandten Wörtern derselben syntaktischen Kategorie ausdrücken, wie etwa Synonyme, Generalisierungen und Konverse, sowie 'syntagmatische' Funktionen, die den Wert von Kollokationen berechnen, für Wörter, die im selben Satz in einem spezifischen syntaktischen Bezug zu dem Lexikoneintrag stehen.

Die Generierung eines in Form eines semantischen Netzwerks spezifizierten Texts erfolgt dann in drei Phasen (das Netzwerk muß natürlich ontologisch zu den Lexikoneinträgen passen):

1. Netzwerkreduktionen und -vereinfachungen

2. Übergang vom semantischen zum tiefen syntaktischen Stratum

3. Übergang vom tiefen syntaktischen zum oberflächennahen syntaktischen Stratum

Durch Netzwerkreduktionen können umfangreiche Definitionen durch den damit assoziierten Term ausgedrückt werden, wie etwa 'more than one program of a type such that someone compiles something with these programs' durch 'compilers'. Damit ist die Behandlung einiger Fälle von Lexikalisierung im eigentlichen Sinne (gemäß Cahill) möglich, was eine besondere Fähigkeit von MTT im Vergleich zu anderen linguistischen Theorien darstellt. Bei Anwendbarkeit mehrerer Äquivalenzdefinitionen können unter Umständen ganz unterschiedliche Aufteilungen des Informationsgehalts erfolgen, so daß 'Fred überquerte den Fluß schwimmend' und 'Fred schwamm quer über den Fluß' (siehe das Kapitel über Lexikalisierung) aus dem gleichen zugrundeliegenden Netzwerk erzeugt werden können.

In der zweiten Phase bestehen oft Alternativen für das verbale Ausdrücken einzelner Knoten. So kann etwa ein mit 'name' bezeichneter Knoten entweder durch 'Name' oder 'heißen' ausgedrückt werden. Die Auswahl zwischen diesen Alternativen und unterschiedliche Zusammensetzungen werden gemäß der Thema/Rhema Struktur getroffen.

In der dritten Phase schließlich bestehen durch auf lexikalischen Funktionen definierte Regeln weitere Möglichkeiten zur Paraphrasierung, etwa durch Bildung von Funktionsverbkonstuktionen. Beispielsweise liefert eine bestimmte lexikalische Funktion zu 'entscheiden' die nominalisierte Form 'Entscheidung' und eine andere lexikalische Funktion das dazu passende Funktionsverb 'treffen'. Dabei sorgt die Definiertheit oder Nicht-Definiertheit einer lexikalischen Funktion für einzelne Lexeme für die Vermeidung der Bildung inkorrekter Phrasen.

Beim Einsatz von Meaning-Text Theorie zur Generierung muß die Eingabe zunächst in satzgrößenartige semantische Netzwerke zerlegt werden, die jeweils mit Markierungen zur Thema/Rhema Struktur angereichert sind, wie dies im System GOSSiP (Generation of Operating System Summaries in Prolog, [Ior91] gemacht wurde. Meaning-Text Theorie ist ein ehrgeiziger und wohlfundierter zur Repräsentation lexikalischen Wissens und geht dabei über die Fähigkeiten anderer Theorien hinaus, zum Beispiel durch die expliziten Netzwerkdefinitionen. Die Anwendung der Theorie setzt aber die Ausarbeitung solcher umfangreicher Lexika voraus, was einen enormen Entwicklungsaufwand erfordert. Darüberhinaus sind die auf diesen Definitionen arbeitenden Transformationsoperationen zur Exploration der alternativen Optionen mit hohem Rechenaufwand verbunden. Meaning-Text Theorie ist somit ein typisches Beispiel einer Theorie, die über eine erhebliche Ausdruckmächtigkeit und linguistische Abdeckung verfügt, diese Vorteile allerdings mit hohem

Effizienzverlust erkauft.

7.2 Systemisch Funktionale Grammatik

Systemisch-funktionale Grammatik ist eine Theorie natürlicher Sprache, für die der Begriff der Funktion zentral ist [Hal94]. Sie ist ähnlich wie Meaning-Text Theorie ein stratifiziertes Modell, und sie enthält neben den üblichen linguistischen Repräsentationsebenen auch außer-linguistische Aspekte, wie etwa soziale Kontexte. Neben der Organisation in Form der Strata, die auf *Abstraktion* basieren, gibt es eine inhaltliche Organisation in Form von *Metafunktionen*, die auf *Diversifikation* basieren, also unterschiedliche Perspektiven liefern. Es gibt drei Metafunktionen: *ideationale*, *interpersonale* und *textuelle*. Die ideationale Sicht bezieht sich auf inhaltliche, semantische Aspekte, also den reinen Informationsgehalt einer Äußerung. Die interpersonale Sicht betrifft den sozialen Bezug zum Hörer, besagt also wie der Informationsgehalt gesehen wird (hypothetisch, befragt, etc.). Die textuelle Sicht schließlich behandelt die kontextuelle Einbettung der Information, ist somit zum Beispiel für die Thema/Rhema Struktur zuständig. Die Erscheinungsform einer generierten Äußerung kann durch Spezifikationen beeinflußt werden, die jeweils unterschiedlichen Metafunktionen zuzuordnen sind.

Elemente des operationalen Teils einer Systemisch-funktionalen Grammatik sind *Systemnetzwerke*, das sind Vererbungsnetzwerke, die eine Menge zusammenhängender Auswahlmöglichkeiten darstellen. Jedes System in diesem Netzwerk verfügt über eine Eingangsbedingung und über zwei oder mehr paradigmatische Features, die grammatische Alternativen darstellen, von denen bei Ansteuern eines Systems eine auszuwählen ist. So besteht etwa bei dem System 'Satzmodus' die Wahl zwischen den Featuren 'imperativ' und 'indikativ'; letzteres entspricht wiederum der Eingangsbedingung für das System 'Indikativ Typ', mit den Wahlmöglichkeiten 'deklarativ' und 'interrogativ'. Beim Traversieren des Netzwerks werden die den gewählten Varianten entsprechenden paradigmatischen Features gesammelt. Das Traversieren des Netzwerks erfolgt durch iterative Ausführung eines der Systeme, deren Eingangsbedingung erfüllt ist (oder mehrerer parallel). Der Effekt dieser Ausführung kann zur Erfüllung der Eingangsbedingung weiterer Systeme beitragen.

Die Grammatik selbst ist mit der syntagmatischen und mit der semantischen Ebene verbunden. Die syntagmatische Organisation basiert auf mit paradigmatischen Features verbundenen Realisierungsfestlegungen, die etwa das Vorhandensein eines Features ausdrücken oder Reihenfolgebedingungen. So führt etwa das paradigmatische Feature 'wh-' zur Einführung einer wh-Funktion und legt fest, daß in der Ausgabe die wh- Konstituente vor dem finiten Verb kommt. Die Verbindung zum semantischen Stratum erfolgt mittels sogenannter 'Chooser', die mit Systemen verbunden sind um grammatische Entscheidungen semantisch zu motivieren. Die 'Chooser' sind intern als Entscheidungsbäume organisiert, bei deren Verzweigungen eine Reihe unterschiedlicher Faktoren zur Auswertung herangezogen werden können: textuelle Relationen, Domänenwissen, ein Benutzermodell und kommunikative Ziele.

Zwei der führenden auf Systemisch-funktionaler Grammatik basierende Generierungssysteme sind GENESYS [Faw90] und Penman [Man85]. GENESYS ist ein sehr reichhaltiges Satzgenerierungssystem, das von Benutzerspezifikationen gesteuert wird. Unter anderem kann es semantisch motivierte Spezifikationen für Intonation erzeugen. Penman ist in der Lage Sätze aus semantischer Eingabe zu erzeugen. Einige Weiterentwicklungen arbeiten an multi-lingualer Generierung mit mehreren, teilweise überlappenden Grammatiken für die Zielsprachen. Techniken zur Generierung mit Systemisch-funktionaler Grammatik werden detailliert in [Mat91] beschrieben.

8 Zusammenfassung

In diesem Beitrag wurden Techniken zur Generierung von natürlichsprachlichen Äußerungen — einer Folge von Sätzen — aus den Spezifikationen des Kontexts einer Situation vorgestellt und ihre Verwendung am Beispiel einer Wegbeschreibung erläutert. Dabei ist klar geworden, daß zur Bewältigung dieser Aufgabe eine Vielzahl heterogener Methoden erforderlich ist, die zur Grob- und Feinplanung und zur Realisierung eines Texts beitragen. Es bestehen in verschiedenen Aspekten Abhängigkeiten zwischen den mit diesen Aufgaben betrauten Prozessen. Sie werden trotzdem meist aus Gründen der Einfachheit sequentiell nacheinander ausgeführt, wobei die in einem System relevanten Abhängigkeiten oft durch spezielle Implementierungskonstrukte umgangen werden, oder schlicht ein gewisser Qualitätsverlust in Kauf genommen wird.

Für viele Teilaufgaben des Generierungsprozesses gibt es ein oder mehrere etablierte Verfahren zur Lösung der jeweiligen Teilaufgabe. Typischerweise sind diese Verfahren an einer lokal funktionalen Sicht orientiert, ohne Perspektive auf den gesamten Generierungsprozeß. Desweiteren sind die linguistischen, rhetorischen und auch psychologischen Kriterien zur Wahl zwischen alternativen Ausdrucksmöglichkeiten an verschiedenen Stellen der Verarbeitung oft zu allgemein um in einer konkreten Situation eine angemessene Entscheidung zu begründen. Sie müssen ergänzt werden durch spezifische, operationale Kriterien, gemäß der gewünschten Funktionalität des jeweils betrachteten Systems. Bekannte Verfahren leisten vorrangig die kohärente

Organisation von satzartigen Propositionen, sowie verschiedene Teilbereiche der Satzplanung, wie etwa Erzeugung referentieller Beschreibungen, das heißt identifizierender Referenzausdrücke. Am weitesten fortgeschritten sind Oberflächenrealisierungsverfahren, manche mit recht umfangreicher Abdeckung, jedoch meist nur für das Englische. Wenig methodische Unterstützung erhält man als Systemdesigner neben den ausreichend spezifischen Auswahlkriterien bei der Modellierung von Inferenzen und bei der Behandlung von Annahmen über den Hörer, speziell dann wenn diese Annahmen mit Unsicherheit behaftet sind und Entscheidungen über den zu generierenden Text dadurch zum Teil strategisch (vorsichtig oder spekulativ) motiviert werden.

Insgesamt erweist sich die Textgenerierung, speziell in den verschiedenen Planungsaspekten, als ziemlich schwierig, da man relativ viele zu Teil sehr unterschiedliche Verfahren kennen muß, um die Problematik des Gebiets zu verstehen und Anforderungen für die Funktionalität eines zu entwickelnden Systems einschätzen zu können.

9 Weiterführende Literatur

Eine ausführliche Diskussion zu Architekturen von Generierungssystemen findet man in [DS95]. Oberflächenrealisierungskomponenten mit umfangreicher Abdeckung sind SURGE [Elh99] und KPML [Bat97]. Einige der bekanntesten Generierungssysteme sind FOG [Gol94], das Wettervorhersagen in Englisch und Französisch produziert, IDAS [Rei95], das online technische Dokumentations aus einer Wissensbasis erzeugt, und MAGIC [McK97], das multimediale Kurzbeschreibungen zur Gesundheitsvorsorge generiert.

Literatur

[And94] Andre, E., Rist T.: *Generating Coherent Presentations Employing Textual and Visual Material.* Artificial Intelligence Review, 9:147–65, 1994.

[Bat97] Bateman, J.: *Enabling Technology for Multilingual Natural Language Generation: The KPML Development Environment.* Natural Language Engineering, 3:15–55, 1997.

[Cah98] Cahill, L.: *Lexicalization in Applied NLG systems.* RAGS deliverable, 1998.

[Car99] Carroll, J., Copestake A.-Flickinger D. Poznanski V.: *An Efficient Generator for (Semi-)Lexicalist Grammars.* In: *Proc.*

of the 7th European Workshop on Natural Language Generation (ENLGW-99), Seiten 86–95, 1999.

[Dal95] DALE, R., REITER E.: *Computational Interpretations of the Gricean Maxims in the Generation of Referring Expressions.* Cognitive Science, 19:233–263, 1995.

[Dal96] DALIANIS, H., HOVY E.: *Aggregation in Natural Language Generation.* In: ADORNI, G., ZOCK M. (Herausgeber): *Trends in Natural Language Generation: An Artificial Intelligence Perspective*, Seiten 88–105. Springer-Verlag, 1996.

[Dob93] DOBES, Z., NOVAK H.-J.: *From Constituent Planning to Text Planning.* In: HORACEK, H., ZOCK M. (Herausgeber): *New Concepts in Natural Language Generation.* Pinter Publishers, 1993.

[DS95] DE SMEDT, K., HORACEK-H. ZOCK M.: *Some Problems with Current Architectures in Natural Language Generation.* In: ADORNI, G., ZOCK M. (Herausgeber): *Trends in Natural Language Generation - An Artificial Intelligence Perspective.* Springer-Verlag, 1995.

[Elh97] ELHADAD, M., MCKEOWN K.-ROBIN J.: *Floating Constraints in Lexical Choice.* Computational Linguistics, 23:159–239, 1997.

[Elh99] ELHADAD, M., ROBIN J.: *SURGE: A Comprehensive Plug-in Syntactic Realization Component for Text Generation.* Computational Linguistics, 4, 1999.

[Faw90] FAWCETT, R., TUCKER G.: *Demonstration of GENESYS: A Very Large Semantically Based Systemic Functional Grammar.* In: *Proc. of COLING-90*, Seiten 47–49, 1990.

[Fei90] FEINER, S., MCKEOWN K.: *Coordinating Text and Graphics in Explanation Generation.* In: *Proc. of the 8th National Conference on Artificial Intelligence (AAAI-90)*, Band 1, Seiten 442–449, 1990.

[Gol94] GOLDBERG, E., DRIEDGER N.-KITTREDGE R.: *Using Natural Language Processing to Produce Weather Forecasts.* IEEE Expert, 9:45–53, 1994.

[Gro86] GROSZ, B., SIDNER C.: *Attention, Intention, and the Structure of Discourse.* Computational Linguistics, 12:175–204, 1986.

[Hal94] HALLIDAY, M.: *Introduction to Functional Grammar.* Edward Arnold, London, 1994.

[Hor96] HORACEK, H..: *On Expressing Metonymic Relations in Multiple Languages.* Machine Translation, 11:109–158, 1996.

[Hor98] HORACEK, H.: *Generating Inference-Rich Discourse Through Revisions of RST-Trees.* In: *Proc. of the 15th National Conference on Artificial Intelligence (AAAI-98)*, Seiten 814–820, 1998.

[Hov93] HOVY, E.: *Automated Discourse Generation Using Discourse Structure Relations.* Artificial Intelligence, 63:341–385, 1993.

[Ior91] IORDANSKAJA, L., KITTREDGE R.-POLGUERE A.: *Implementing the Meaning-Text Model for Language Generation.* In: PARIS, C., SWARTOUT W.-MANN W. (Herausgeber): *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, Seiten 293–312. Kluwer Academic, 1991.

[Kay96] KAY, M.: *Chart Generation.* In: *Proc. of ACL-96*, Seiten 200–204, 1996.

[Kit91] KITTREDGE, R., KORELSKY T.-RAMBOW O.: *On the Need for Domain Communication Knowledge.* Computational Intelligence, 7(4):305–314, 1991.

[Mai93] MAIER, E., HOVY E.: *Organizing Discourse Structure Relations Using Metafunctions.* In: HORACEK, H., ZOCK M. (Herausgeber): *New Concepts in Natural Language Generation*, Seiten 69–86. Pinter Publishers, 1993.

[Mai95] MAIER, E.: *Textual Relations as Parts of Multiple Links Between Text Segments.* In: ADORNI, G., ZOCK M. (Herausgeber): *Trends in Natural Language Generation - An Artificial Intelligence Perspective*, Seiten 68–87. Springer-Verlag, 1995.

[Man85] MANN, W., MATTHIESSEN C.: *Demonstration of the NIGEL Text Generation Computer Program.* In: *Systemic Perspectives on Discourse*, Seiten 50–83. Norwood Ablex, 1985.

[Man91] MANN, W., THOMPSON S.: *Rhetorical Structure Theory: Towards a Functional Theory of Text Organization.* Text, 3:243–281, 1991.

[Mar97] MARCU, D.: *From Local to Global Coherence: A Bottom-Up Approach to Text Planning.* In: *Proc. of the 14th National Conference on Artificial Intelligence (AAAI-97)*, Seiten 629–635, 1997.

[Mat91] MATTHIESSEN, C., BATEMAN J.: *Text Generation and Systemic-Functional Linguistics.* Pinter Publishers, London, 1991.

[McK85a] MCKEOWN, K.R.: *Discourse Strategies for Generating Natural-Language Text.* Artificial Intelligence, 27:1–41, 1985.

[McK85b] MCKEOWN, K.R.: *Text Generation: Using Discourse Strategies and Focus Constraints to Ge nerate Natural Language Text.* Cambridge University Press, Cambridge, 1985.

[McK97] MCKEOWN, K., PAN S.-SHAW J. JORDON D. ALLEN B.: *Language Generation for Multimedia Healthcare Briefings.* In: *Proc. of the 5th Conference on Applied Natural Language Processing (ANLP-1997)*, Seiten 277–282, 1997.

[Mel82] MELCUK, I.: *Lexical Functions in Lexicographic Descriptions.* In: *Proc. of the 8th Annual Meeting of the Berkeley Linguistic Society*, 1982.

[Mel87] MELCUK, I., POLGUERE A.: *A Formal Lexicon in the Meaning-Text Theory.* Computational Linguistics, 13(3–4):261–275, 1987.

[Met92] METEER, M.: *Expressibility and the Problem of Efficient Text Planning.* Pinter Publishers, 1992.

[Moo89] MOORE, J., SWARTOUT B.: *A Reactive Approach to Explanations.* In: *Proc. of the 11th International Joint Conference on Artificial Intelligence (IJCAI-89)*, Seiten 1504–1510, 1989.

[Moo92] MOORE, J., POLLACK M.: *A Problem for RST. The Need for Multi-Level Discourse Analysis.* Computational Linguistics, 18:537–544, 1992.

[Mos96] MOSER, M., MOORE J.: *Towards a Synthesis of two Accounts of Discourse Structure.* Computational Linguistics, 22:409–419, 1996.

[Nog91] NOGIER, J.-F., ZOCK-M.: *Lexical Choice as Pattern Matching.* In: NAGLE, T., NAGLE J.-GERHOLZ L. EKLUND P. (Herausgeber): *Current Directions in Conceptual Structures Research.* Springer, 1991.

[Par87] PARIS, C.: *The Use of Explicit User Models in Text Generation: Tailoring to a User's Level of Expertise.* PhD Thesis, Columbia University, 1987.

[Par88] PARIS, C.: *Tailoring Object Descriptions to a User's Level of Expertise.* Computational Linguistics, 14:64–78, 1988.

[Rei94] REITER, E.: *Has a Consensus NL Generation Architecture Appeared, and is it Psycholinguistically Plausible?* In: *Proc. of the 7th International Workshop on Natural Language Generation (INLGW-94)*, Seiten 163–170, 1994.

[Rei95] REITER, E., MELLISH C.-LEVINE J.: *Automatic Generation of Technical Documentation.* Applied Artificial Intelligence, 9:259–287, 1995.

[Rei00] REITER, E., DALE R.: *Building Natural Language Generation Systems.* Cambridge University Press, Cambridge, 2000.

[Sco90] SCOTT, D., DE SOUZA-C.: *Getting the Message Across in RST-based Text Generation.* In: DALE, R., MELLISH C.-ZOCK M. (Herausgeber): *Current Research in Natural Language Generation*, Seiten 47–74. Academic Press, 1990.

[Sha98a] SHAW, J.: *Clause Aggregation Using Linguistic Knowledge.* In: *Proc. of the 9th International Workshop on Natural Language Generation (INLGW-98)*, Seiten 138–147, 1998.

[Sha98b] SHAW, J.: *Segregatory Coordination and Ellipsis in Text Generation.* In: *Proc. of the 36th Annual Meeting of the Association for Computational Linguistics and the 17th International Conference on Computational linguistics (COLING-ACL'98)*, Seiten 1220–1126, 1998.

[Shi90] SHIEBER, S., PEREIRA F.-VAN NOORD G. MOORE R.: *Semantic-Head-Driven Generation.* Computational Linguistics, 16:30–42, 1990.

[Ste96] STEDE, M.: *Lexical Options in Multilingual Generation from a Knowledge Base.* In: ADORNI, G., ZOCK M. (Herausgeber): *Trends in Natural Language Generation: An Artificial Intelligence Perspective*, Seiten 222–237. Springer-Verlag, 1996.

[vO87] Oirsouw, R. van: *The Syntax of Coordination*. Croom Helm, Beckenham, 1987.

[Whi92] Whitelock, P.: *Shake-and-Bake Translation.* In: *Proc. of COLING-92*, Seiten 610–616, 1992.