

# Wissenschaftliche Textproduktion und Fachdomäne

## Sprachliche Realisierung wissenschaftlicher Inhalte in verschiedenen Fachdisziplinen und ihre computerlinguistische Modellierung

Wiebke Ramm und Claudia Villiger  
Saarbrücken

In this paper we investigate the communicative and textual properties of scientific articles in journals of two different scientific domains, medicine and law, which differ significantly in their responsibilities, methods and objectives. We characterize the nature of scientific activity in the two fields and describe the role scientific communication plays in each of them. Our analyses of weekly journals of the two disciplines examine how underlying states of affairs are dealt with linguistically and which communicative means are deployed by the authors. We then use these empirical results to assess the possibilities of computational linguistic support for the production of scientific texts in the two areas. Finally, we sketch a text generation architecture for the drafting of medical case reports.

### 1 Einleitung

In unserem Beitrag wollen wir der Frage nachgehen, wie sich wissenschaftliche Texte verschiedener Fachdisziplinen in ihren linguistischen und insbesondere in ihren kommunikativ-textuellen Eigenschaften unterscheiden und welche Rolle dabei die zugrundeliegenden Handlungsbereiche und Begriffssysteme des jeweiligen Fachgebiets spielen.

Unabhängig von der Fachdisziplin zeichnet sich wissenschaftliche Fachkommunikation durch eine Reihe gemeinsamer Merkmale und Anforderungen aus:

- Wissenschaftliche Texte dienen in erster Linie der *fachlichen Verständigung* zwischen Wissenschaftlern eines bestimmten Fachgebietes, beziehen sich also auf den fachinternen Kommunikationsbereich zwischen Experten. Einen weiteren Anteil der wissenschaftlichen Textproduktion nimmt der interdisziplinäre Austausch zwischen Experten verschiedener Fachdisziplinen sowie die Vermittlung wissenschaftlicher Zusammenhänge von Experten an Laien ein. Hier soll jedoch die fachinterne Kommunikation im Mittelpunkt stehen.
- Fachkommunikation muß *aktuell* sein. Typische Publikationsmedien in den meisten wissenschaftlichen Disziplinen sind daher Wochen- Monats- oder Vierteljahreszeitschriften. Solche Fachzeitschriften unterscheiden sich in ihrer Funktion im wissenschaftlichen Kommunikationsprozeß deutlich von

Büchern, die eine wesentlich längere Zeitspanne zwischen dem eigentlichen Schreibprozeß, der nachfolgenden Revision/Korrektur und dem tatsächlichen Erscheinungsdatum erfordern. Die Zeit zwischen dem Einreichen eines Fachartikels bei einer Zeitschrift und seiner Publikation (incl. Revision) beläuft sich dagegen durchschnittlich auf ca. 6–12 Monate.<sup>1</sup> Als noch aktuellere wissenschaftliche Kommunikationsplattform tritt außerdem gerade in letzter Zeit zunehmend das Internet mit seinem „*World Wide Web*“ in Erscheinung. Welche einschneidenden Veränderungen sich aus den Möglichkeiten dieses elektronischen Mediums für den gesamten Schreib-, Revisions- und Publikationsprozeß sowie für die traditionellen Rollen von Autor, Herausgeber und Leser ergeben, ist derzeit noch kaum abschätzbar.

- Wissenschaftlicher Austausch findet außerdem zunehmend *international* statt und macht damit entweder die multilinguale Publikation von wissenschaftlichen Ergebnissen erforderlich, oder aber das Schreiben in einer wissenschaftlichen „lingua franca“, wie sie beispielsweise das Englische für viele naturwissenschaftliche Forschungsdisziplinen darstellt, welche aber für einen Großteil der Autoren nicht die Muttersprache ist.
- Nicht zuletzt wird wissenschaftliche Information und Kommunikation immer *vernetzter*, wie man am aktuellen Internet-/WWW-Boom oder der massiven Werbung für die Nutzung von „Datenautobahnen“ ersehen kann, und greift immer mehr auf schnelle elektronische Medien zurück.

Was bedeutet dies nun für die Fachkommunikation zwischen Wissenschaftlern? Aus den Forderungen nach Aktualität und Internationalität in Verbindung mit der Nutzung neuer elektronischer Publikationsmedien ergeben sich neue Anforderungen an den wissenschaftlichen Textproduktionsprozeß. Uns interessiert hier vor allem die Frage, inwieweit sich dieser (Fach-)Kommunikationsprozeß mit computerlinguistischen Werkzeugen, insbesondere mit Hilfe der automatischen Textgenerierung, unterstützen läßt und ob dies für Texte verschiedener Fachgebiete gleichermaßen gilt. Gegenstand unseres Beitrages soll daher die Untersuchung der linguistischen Merkmale (textuell, grammatisch, lexikalisch) von Fachtexten aus zwei Fachgebieten sein, die bezüglich ihres Handlungs- und Gegenstandsbereiches sowie in ihrer wissenschaftlichen Methodologie sehr unterschiedlich sind: *Medizin* und *Jura*. Nach einer allgemeinen Charakterisierung von Fachkommunikation und ihren Funktionen in den beiden Wissenschaftsdisziplinen (Kapitel 2) wollen wir in Kapitel 3 anhand von Beispieltexten prüfen, welche Parameter den jeweiligen wissenschaftlichen Kommunikationsprozeß bestimmen (z. B. welche Rolle die Kommunikationsteilnehmer darin spielen) und welchen Einfluß das jeweilige Fachgebiet und die damit verbundenen wissenschaftlichen Ziele und Methoden auf die linguistische Erscheinungsform von Fachartikeln hat. In Kapitel 4 wollen wir auf

---

1 Skudlik (1990, 31) weist darauf hin, daß eine Publikation innerhalb von 6-9 Monaten in den Naturwissenschaften als schnell gilt.

der Basis dieser (text-)linguistischen Analysen der Frage möglicher computerlinguistischer Unterstützung bei der wissenschaftlichen Textproduktion in den beiden Fachgebieten nachgehen, wobei der Einsatz der Textgenerierung zur Erzeugung von Rohfassungen (*Drafting*) von Fachartikeln den Schwerpunkt bildet.

Als konkreten Vorschlag in diese Richtung skizzieren wir in Kapitel 5 ein System zur Generierung von Rohfassungen medizinischer Fachartikel. Das Schlußkapitel resümiert noch einmal unsere Ergebnisse.

## 2 Fachkommunikation und wissenschaftliche Domäne

Allgemeine Definitionen zu dem Begriff ‚Fachsprache‘ (Hoffmann 1985, 53; Busch-Lauer 1991, 11) heben als wesentliches kommunikatives Ziel von Fachsprachen das Ermöglichen der *Verständigung* zwischen Menschen, die in einem bestimmten Kommunikationsbereich arbeiten, hervor. Busch-Lauer (1991, 93) bezeichnet als kommunikatives Ziel eines wissenschaftlichen Zeitschriftenaufsatzes das „sachbetonte Informieren“ und das „forschend-analysierende Klären“. Deutliche Unterschiede gibt es in bezug auf die *Kommunikationsbereiche*, d. h. den Ausschnitt aus der gesellschaftlichen Wirklichkeit, in dem die jeweilige Fachsprache verwendet wird (vgl. Hoffmann 1985, 53). Darin gründet die z. B. von Hoffmann (1985, 62ff.) geäußerte Skepsis gegenüber dem Versuch der Formulierung einer „allgemeinen wissenschaftlichen Fachsprache“. In Typologien von Fachsprachen (z. B. bei Benes (1969, 227ff.) spielt der Kommunikationsbereich als ein Klassifikationsmerkmal eine wichtige Rolle. Wie können solche Kommunikationsbereiche aussehen? Als Beispieldomänen werden wir uns im folgenden auf die Fachgebiete Humanmedizin und Jurisprudenz konzentrieren.

In der *Humanmedizin* beschäftigt sich der Wissenschaftler vornehmlich mit den Ursachen und Erscheinungsformen von Krankheiten, deren Erkennung und Behandlung. Dabei kann auf exakte Grundlagen aus verschiedenen Naturwissenschaften wie z. B. Biologie und Chemie zurückgegriffen werden. Eine naturwissenschaftliche Betrachtungsweise des Menschen ist in diesem Zusammenhang vorrangig. Der Handlungsraum ist primär *nicht-sprachlich*, d. h. es müssen Handlungen in der physikalischen Welt vollzogen werden, die es erlauben, das erstrebte Ziel, die Wiederherstellung der Gesundheit, zu erreichen. Wie generell in den Naturwissenschaften ist der Geltungsbereich dieser Wissenschaft weitgehend universell, d. h. unabhängig von nationalen Grenzen und damit einem internationalen wissenschaftlichen Diskurs unterworfen, der sich zunehmend des Englischen als Vermittlungssprache bedient. In ihrem Wissenschaftssprachenprofil für die bundesdeutsche Wissenschaft hebt Skudlik (1990, 99) hervor, daß neben der Mathematik und den ‚reinen‘ Naturwissenschaften (z. B. Biologie oder Chemie) die Humanmedizin die stärkste anglophone Prägung aufweist. Dies zeigt sich beispielsweise darin, daß auch deutschsprachige Artikel in medizinischen Zeitschriften fast immer ein englisches Abstract enthalten.

Die *Rechtswissenschaft* steht in Beziehung zu anderen Geistes- und Sozialwissenschaften. Sie beschäftigt sich mit der Erkenntnis, der Auslegung und der Fortbildung des geltenden Rechts. Der Handlungsraum hat daher eine primär *sprachliche* Dimension, die sich z. B. in der Ausbildung von deklarativen Sprechakten („Im Namen des Volkes ergeht folgendes Urteil...“ etc.) niederschlägt. Außerdem ist hier der Geltungsbereich beschränkter, da eine Anbindung der Rechtswissenschaft an das (z. B. nationale) Rechtssystem, auf das sie sich bezieht, vorgegeben ist.<sup>2</sup> Große Bereiche der wissenschaftlichen Arbeit sind damit nationalsprachlich geprägt. Englisch als „lingua franca“ ist in der Rechtswissenschaft folglich weitaus weniger verbreitet (vgl. Skudlik 1990, 99).

Bezüglich ihrer Kommunikationsbereiche lassen sich also einige Unterschiede zwischen Humanmedizin und Rechtswissenschaft aufweisen. Während die Humanmedizin empirisch, daten- und handlungsorientiert ist und auf naturwissenschaftliche Grundlagen zurückgreift (Benes (1969) ordnet sie in seiner Typologie der Fachsprachen den Angewandten Naturwissenschaften zu), geht es in der Rechtswissenschaft um das Auslegen, Paraphrasieren und Fortbilden von sprachlich Gegebenem (z. B. den Gesetzestexten des jeweiligen Rechtssystems). Folglich unterscheidet sich das sprachliche Handeln des Humanmediziners von dem des Rechtswissenschaftlers: Während in der Humanmedizin das Austauschen von Fakten und Erfahrungen im Zentrum des Kommunikationsprozesses steht, ist es für den Rechtswissenschaftler von Bedeutung, seine Sicht auf einen Sachverhalt oder einen neuen Aspekt eines Sachverhalts möglichst vollständig und überzeugend darzustellen. Ein weiterer Unterschied läßt sich in der jeweiligen Fachterminologie ausmachen. Während sich die Medizin in präziser Weise mittels einer lateinisch orientierten Terminologie verständigt, greift die Rechtswissenschaft in hohem Maße auf lexikalisches Material zurück, das auch in der Alltagssprache Verwendung findet. Vor allem für Nicht-Juristen kommt es daher oft zu Problemen bei der Unterscheidung von wissenschaftlicher und alltagssprachlicher Verwendung von Wörtern und Sätzen. Eine Konsequenz der Sprachorientierung in der Rechtswissenschaft ist ihre Unterscheidung von drei Sprachebenen (vgl. Podlech 1975, 162): die Normsprache (z. B. Gesetzestexte), die Rechtfertigungssprache (z. B. Entscheidungsbegründungen) und die dogmatische Sprache, die im Zentrum der wissenschaftlichen Auseinandersetzung steht und daher hier Gegenstand der Untersuchung ist.

### 3 Analyse von Beispieltexten

Wie in der Einleitung erwähnt, spielt Aktualität wissenschaftlicher Kommunikation eine immer größere Rolle. Daher haben wir Texte aus einem Medium, einer *Wochenzeitschrift*, gewählt, das eine möglichst schnelle Publikation gewährleistet.

---

<sup>2</sup> zu den Problemen der Übertragung formaler Kategorien von verschiedenen Rechtssystemen vgl. z. B. Oksaar (1988).

In der Deutschen Medizinischen Wochenschrift (DMW) und der Neuen Juristischen Wochenschrift (NJW) werden aktuelle Forschungsergebnisse und Informationen aus den jeweiligen Fächern publiziert. In beiden Fällen handelt es sich um fachinterne Kommunikation, d. h. es liegt eine symmetrische Interaktion von Experte zu Experte vor. Dabei sind Textrezipienten und -produzenten von beiden Wochenzeitschriften sowohl Forscher als auch Praktiker der jeweiligen Fächer. Auch in bezug auf den Gegenstandsbereich der Artikel haben wir uns um Paralleltät bemüht, indem wir uns für beide Fächer auf Fallstudien, *Kasuistiken*, konzentriert haben, die in den beiden Disziplinen jedoch sehr unterschiedlichen Status haben, wie wir sehen werden.

In der folgenden Analyse möchten wir am Beispiel wissenschaftlicher Texte aus der Medizin und der Rechtswissenschaft zeigen, wie die verschiedenen Kommunikationsbereiche sich auf die sprachliche Realisierung der Texte und das kommunikative Ziel, das verfolgt wird, auswirken. Dabei werden wir *top-down* vorgehen, d. h. wir beginnen mit der Analyse der Makrostruktur und beobachten am Schluß die Phänomene auf Satzebene und die lexikalische Selektion.

Beim Vergleich der Makrostruktur medizinischer Kasuistiken fällt die *Homogenität der makrostrukturellen Gliederung* auf. Nach einer kurzen Einleitung, in der zumeist zusammenfassend das Krankheitsbild geschildert, mögliche Ursachen der Beschwerden angesprochen und die Relevanz der Untersuchung in der Praxis skizziert wird, wird zur eigentlichen Kasuistik übergeleitet. Die Teile der Kasuistiken, die mit Überschriften versehen werden, sind nahezu identisch: Sie beginnen mit der ‚Anamnese‘, es folgen die ‚Befunde‘ (z. B. in Text 1 ‚Klinischer Befund‘, ‚Klinisch-chemischer Befund‘, ‚Ergänzende Untersuchungen‘), dann der Teil ‚Therapie und Verlauf‘ und schließlich die ‚Diskussion‘. Die Literatur wird am Ende des Aufsatzes aufgeführt, manchmal wird zwischen Kasuistik und Literatur eine Dankagung eingefügt. Diese makrostrukturelle Gliederung der Kasuistiken ist nahezu *standardisiert*.<sup>3</sup> Die Abstracts, die in deutscher und englischer Sprache wiedergegeben werden, spiegeln diese Makrostruktur gewöhnlich wider.

Die Homogenität der Struktur, die wir anhand der makrostrukturellen Analyse medizinischer Texte beobachten konnten, läßt sich auch auf Absatzebene nachweisen. So beginnt die Anamnese mit der Nennung von Alter und Geschlecht des Patienten, es folgen Angaben zur Dauer und der Art der Beschwerden. Daran schließen sich weitere Beobachtungen an und die Anamnese endet mit der Beschreibung des Allgemeinbefindens und der Vorerkrankungen des Patienten. Die einzelnen Teile der Befunde weisen in regelmäßiger Abfolge entweder zuerst die Normalwerte und dann die von den Normalwerten abweichenden Befunde oder umgekehrt auf. Rein

---

3 Auf eine ähnliche makrostrukturelle Gliederung von Experimentalstudien in der Pädagogischen Psychologie weist Busch-Lauer (1991, 94) hin.

aufzählende Darstellung ist insbesondere bei der Wiedergabe der Normalwerte häufig:

- (1) Im Normbereich lagen die Werte für Hämoglobin, Thrombozyten, Gerinnungsparameter, Gesamteiweiß, Elektrolyte, Kreatinin, Harnstoff, g-GT, alkalische Phosphate, GOT, GPT, Lactatdehydrogenase, Kreatinkinase, Glucose, Amylase, Lipase sowie Angiotensin-Converting-Enzym, Lysozym und Komplementfaktoren. (Text 1, 201).

Abweichende Werte werden typischerweise als Merkmals-Wert-Paare aufgelistet, wobei das Satzthema das Merkmal und das Rhema dessen Wert bezeichnet:

- (2) Die Rachenschleimhaut war diskret gerötet. ... Der Blutdruck betrug 145/180 mm Hg. Der Klopfeschall über allen Lungenetagen war sonor; ... Das Abdomen war weich und nicht druckdolent; Leber, Milz und Lymphknoten waren nicht vergrößert tastbar. (Text 1, 201).

Auch in dem Kapitel ‚Therapie und Verlauf‘ wird versucht, so knapp wie möglich die Fakten darzustellen. Selbst der Schluß des Kapitels, der den Erfolg der Therapie beschreibt hat häufig aufzählenden Charakter:

- (3) Unter dieser Therapie normalisierte sich der Befund im Thorax-Röntgenbild bereits nach 2 Monaten (Abb. 3); lungenfunktionsanalytisch zeigte sich nach 6 Monaten ein Normalbefund mit normalem Transferfaktor und regelrechter Blutgasanalyse. (Text 1, 202).

Die Anordnung der Sachverhalte erfolgt in diesem Teil der Kasuistik chronologisch. Zeitangaben signalisieren die temporale Strukturierung des Kapitels:

- (4) Nach zehntägiger antibiotischer Behandlung ... In der daraufhin durchgeführten ... In der zugleich durchgeführten ... Daraufhin ... Innerhalb von wenigen Tagen ... nach einer Woche ... Nach 2 Wochen ... In sechswöchigen Intervallen ... (Text 1, 201-202).

Die Teile der medizinischen Texte, die die Diagnose des jeweiligen Falles beschreiben sind also streng *deskriptiv* gehalten – z. T. werden die Fakten sogar im Telegrammstil wiedergegeben – während im nachfolgenden Kapitel ‚Therapie und Verlauf‘ die globale Strukturierung durch den zeitlichen Ablauf gekennzeichnet ist. Zumindest für den letzten Teil der Kasuistiken, die Diskussion, könnte man erwarten, daß der vornehmlich deskriptive Stil zugunsten eines mehr argumentativen Stils aufgegeben wird. Doch auch für diesen Teil der Kasuistik gibt es eine mehr oder minder standardisierte Form. Begonnen wird zumeist mit einer Einordnung des Krankheitsbildes, woraufhin auf die Beschwerden eingegangen wird, die die meisten betroffenen Patienten erleiden. Hier werden häufig noch zusätzliche Fakten zum Krankheitsbild (z. B. Auftretenshäufigkeit von einzelnen Symptomen, die bei der Darstellung des Einzelfalles nicht berücksichtigt werden konnten) und weitere diagnostische Hinweise vermittelt. Danach wird nochmals zusammenfassend auf die Diagnose und die Therapie eingegangen, wobei von dem zuvor behandelten Einzelfall abstrahiert wird. Auch in diesem, wie in den anderen Teilen der Kasuistiken, ist auffallend, wie sehr der Autor/die Autoren um Zurückhaltung bemüht sind. Erst in diesem Teil der Kasuistik werden die Einschätzungen der Mediziner durch entsprechende Prädikationen (*machten unwahrscheinlich, war möglich etc.*), allerdings im Passiv oder durch Passivparaphrasen verbalisiert:

- (5) Im geschilderten Fall *machten* die nur wenig ausgeprägte Blutleukozytose und die Erhöhung der Konzentration des C-reaktiven Proteins eine ätiologische Relevanz der in der

ersten BAL nachgewiesenen Keime *unwahrscheinlich*. Da *Candida albicans* im Rachenabstrich nachgewiesen wurde, *war* eine endobronchiale Verschleppung von Pilzhyphen mit dem Bronchoskop *möglich*. Hinsichtlich des *Staphylococcus aureus* *ist* eine sekundäre Kontamination der Lavage *denkbar*. *Dafür spricht*, daß unter einer breit angelegten Antibiotikatherapie keine Besserung von klinischem und radiologischem Befund erzielt werden konnte. Uncharakteristische Ergebnisse einer erneuten BAL und einer transbronchialen Biopsie *ließen* die offene Lungenbiopsie *entscheidend* für die Diagnose *werden*. (Text 1, 203; Hervorhebungen der Autorinnen).

Wie schon bei der Charakterisierung der einzelnen Teile der medizinischen Kasuistiken deutlich wurde, werden vornehmlich Fakten vermittelt. Dies schlägt sich in der Satzstruktur nieder, für die kennzeichnend ist, daß häufig semantisch (nahezu) leere Verben gebraucht werden. Fachterminologie, v. a. nominale, wird rege gebraucht, wobei es kaum Unterscheidungsprobleme zwischen fach- und umgangssprachlichen Ausdrücken gibt. Bei der Betrachtung der Satzstruktur und des Lexikons wird deutlich, daß das Schreiben in einer Fremdsprache – in der Medizin fast immer das Englische – durch die Dominanz der Fachterminologie und die Konzentration auf die Fakten, die sich in dem entsprechenden Satzbau niederschlägt, erleichtert wird. Die Merkmale der sprachlichen Realisierung medizinischer Texte lassen sich aus ihrer kommunikativen Funktion erklären, die vornehmlich darin liegt, zu *informieren*. Als weitere Subfunktionen treten das Instruieren (z. B. in eher technisch orientierten zahnmedizinischen Aufsätzen) und das Empfehlen (manchmal in den Teilen ‚Therapie‘ und ‚Verlauf‘ und ‚Diskussion‘) auf. Will man aufgrund dieser kommunikativen Funktion medizinische Aufsätze einem bestimmten Texttyp (zur Texttypologie vgl. z. B. Werlich 1979 und Virtanen 1992) zuordnen, so bietet sich hierfür der *deskriptive Texttyp* an.<sup>4</sup>

Ganz anders fällt die Analyse der rechtswissenschaftlichen Fallstudien aus. Schon die Makrostruktur der Zeitschriftenaufsätze unterscheidet sich auffällig von der der medizinischen Kasuistiken. Die Kapitelzahl variiert und die Kapitelüberschriften sind von den jeweiligen inhaltlichen Aspekten geprägt. Da wir Texte ausgewählt haben, denen ein aktueller Fall und damit ein aktuelles Urteil zugrundeliegt, ist die Makrostruktur insofern homogen, als im ersten Teil des Textes auf den Ausgangsfall, die sogenannte Sachverhaltsbeschreibung (vgl. Podlech 1975, 163), und das darauf folgende Urteil eingegangen wird. Zum Teil wird dies in der jeweiligen Kapitelüberschrift ausgedrückt (z. B. „Der Fall und seine Probleme“ (Text 2), „Der dem Beschluß zugrundeliegende Fall“ (Text 3)). Eine weitere Parallelität der makrostrukturellen Gliederung liegt darin, daß am Ende des Textes zumeist eine Zusammenfassung des Dargestellten erfolgt, z. T. wird auch dies durch eine entsprechende Kapitelüberschrift (z. B. „Zusammenfassende Würdigung und Ausblick“ (Text 3)) angezeigt. Um den Überblick über die Globalstruktur des Aufsatz-

---

4 Die Zuordnung zu einem Texttyp soll hier bedeuten, daß der jeweilige Text vorwiegend, aber keinesfalls ausschließlich durch den gewählten Texttyp sprachlich realisiert wird (vgl. hierzu auch Lavid 1994 und Ramm/Rothkegel/Steiner/Villiger 1995).

zes zu gewährleisten, wird zumeist im Abstract oder im ersten Kapitel eine Übersicht (*Advanced Organizer*) aufgeführt.

Die Fallbeschreibung im ersten Kapitel erfolgt häufig in aller Kürze:

- (6) In dem Fall, der der Entscheidung zugrunde liegt, hatten zahlreiche Demonstranten die A8 in beiden Richtungen dadurch blockiert, daß sie sich in dichter Menge auf die Fahrbahnen stellten. Durch den so verursachten Stau wurde eine große Zahl von Kraftfahrern am Fortkommen gehindert. (Text 2, 160).

Ausführlicher dargestellt wird fast immer die zu diskutierende Entscheidung(en), sowie deren Begründung. In unserem Beispieltext wird gleich im Abstract die Stellung angezeigt, die der Autor einnimmt:

- (7) Der Beitrag kritisiert die Annahme des BGH, die Blockade einer Autobahn durch eine Barriere von Menschen sei eine Nötigung. (Text 2, 160).

Die Kapitelübersicht im ersten Kapitel dient dazu, den Leser über die Globalstruktur des Aufsatzes zu informieren und die argumentative Linie, die der Autor verfolgt, nachzuzeichnen.

- (8) Den beiden Entscheidungen ist zu widersprechen. Erstens haben die Demonstranten gegen die Kraftfahrer nicht „Gewalt“ i.S. des §240 StGB angewendet (dazu unten II). Zweitens wäre der Nötigungstatbestand selbst dann nicht erfüllt, wenn mit dem BGH die Anwendung von Gewalt anzunehmen wäre. Die Annahme, der Tatbestand sei erfüllt, widerspricht außerdem Art. 103 II GG (dazu unten III). Drittens ist statt des § 240 StGB § 315 bI Nr. 2 StGB anzuwenden (dazu unten IV). (Text 2, 160).

Das bedeutet, daß sobald der Fall und die entsprechenden Entscheidungen und Entscheidungsbegründungen skizziert wurden, der deskriptive Modus verlassen und eine mehr persuasive Haltung eingenommen wird. Während in der oben zitierten Passage aufzählende Adverbien in Satzthemaposition benutzt werden, um die Argumente des Autors zu unterstreichen, finden sich ansonsten in juristischen Aufsätzen zahlreiche argumentative Adverbien und Konjunktionen:

- (9) *Zwar* kann es sein, daß auch K nicht Gewalt anwendet, *nämlich dann, wenn* sie um die Lebensgefährlichkeit der Spritze nicht weiß. *Aber* selbst in diesem letzteren Fall ist es ausgeschlossen zu sagen, daß A gegen den P Gewalt anwendet, mag er auch mittelbarer Täter des Mordes sein. *Vielmehr*... (Text 2, 161; Hervorhebungen der Autorinnen).

Anhand der Textbeispiele dürfte schon deutlich geworden sein, daß auch die Rolle des Autors von juristischen Texten eine ganz andere ist, als in medizinischen Texten. Der Autor eines medizinischen Textes ist zurückhaltend – er läßt die Fakten für sich sprechen. Stammen die Zeitschriftenaufsätze von mehreren Autoren, ist häufig nicht klar erkennbar, wer die Veröffentlichung eigentlich produziert hat – im Vordergrund steht das jeweilige Institut und damit die entsprechende Forschergruppe. Der Autor eines juristischen Textes hingegen muß, um dem Gegenstand gerecht zu werden, Argumente gegeneinander abwägen, seine eigene Position begründen etc. Daher werden in fast allen Kapiteln Bewertungen durch den Autor vorgenommen:

- (10) durcheinandergebracht werden ... was nicht selten übersehen wird ... ist freilich wenig glücklich gewählt ... die nicht sonderlich gut gelungene Formulierung ... genau eine solch

Entleerung findet in den hier besprochenen Entscheidungen statt ... was man auch sofort sieht (Text 2, 160-162).

Dementsprechend häufig sind Verben, die Bewertungen ausdrücken oder Konnotationen implizieren. Ein weiteres Merkmal von juristischen Texten ist das Auftreten von Metaphern oder, wie das folgende Textbeispiel zeigt, von rhetorischen Figuren (häufig rhetorische Fragen):

- (11) Das heißt doch: Die Einschränkung des („schrankenlosen“!) Grundrechts gilt bei Nichtchristen als nicht vorhanden, wenn Christen (wer? mit welcher Legitimation?) der Meinung sind, Nichtchristen würden nur verhältnismäßig gering beeinträchtigt, zumindest, wenn sie ein normales Empfinden haben? *Und das alles nicht so wichtig nehmen? Und warum machen dann die Gegner des Kruzifix-Beschlusses so ein großes Geschrei? Und wo doch sogar – im Vergleich zu bis 80 cm großen Kruzifixen winzige und nur zeitweise sichtbare – Anti-AKW-Plaketten nicht tolerabel sind?* Demgegenüber bleibt festzuhalten:... (Text 3, 3352; Hervorhebungen der Autorinnen).

Die dominierende kommunikative Funktion von juristischen Zeitschriftenaufsätzen liegt im Beurteilen und Bewerten von Sachverhalten mit dem Ziel, daß sich der Leser der Argumentation anschließt. Sie ist damit *persuasiv*. Als weitere kommunikative Funktion muß wie für die medizinischen Texte das Informieren genannt werden. Diese Funktion ist allerdings nur in kürzeren Passagen und im Falle der Kasuistiken v. a. im ersten Kapitel (Fall- und Urteilsbeschreibung) vorherrschend. Die folgende Übersicht faßt die linguistischen Merkmale unserer Textanalysen, die charakteristisch für medizinische und juristische Zeitschriftenaufsätze sind, zusammen.

	Medizinische Texte	Juristische Texte
Kommunikative Funktion	Hauptfunktion: Informieren Weitere Funktionen: Empfehlen, Instruieren	Hauptfunktion: Überzeugen Weitere Funktionen: Informieren, Bewerten, Beurteilen
Texttyp	deskriptiv	argumentativ
Makrostruktur	stark standardisiert	kaum standardisiert
Satzbau	in vielen Teilen je nach Kommunikationsinhalt stark schematisch (z. B. Attribut-Wertzuordnung, temporale Elemente im Satzthema)	vielseitig, z. T. aufgrund der Argumentationen Parataxen häufig
Lexikalische Selektion	Fachterminologie häufige semantisch leere Verben	der Umgangssprache sehr ähnlich

Tabelle 1: Linguistische Merkmale medizinischer und juristischer Zeitschriftenaufsätze

## 4 Wissenschaftliche Textproduktion und automatische Textgenerierung

Was bedeuten diese unterschiedlichen linguistischen Merkmale juristischer und medizinischer Kasuistiken nun im Hinblick auf den potentiellen Einsatz computerlinguistischer Werkzeuge bei der wissenschaftlichen Textproduktion in den beiden Fachdisziplinen? Dieser Frage gehen wir im folgenden nach, wobei wir uns insbesondere auf die Möglichkeiten der automatischen Generierung von wissenschaftlichen Texten, genauer gesagt, von *Rohfassungen (Drafts)* von Fachartikeln konzentrieren. Ein solches „*Drafting*“ kann gerade vor dem Hintergrund des immer schnelleren wissenschaftlichen Fortschritts und den damit wachsenden Anforderungen an die Aktualität wissenschaftlicher Publikationen (auch beispielsweise via Internet) für viele Forschungsdisziplinen interessant sein. Wir gehen zunächst auf einige grundsätzliche Merkmale und Voraussetzungen der Textgenerierung sowie potentiell dafür in Frage kommende Textsorten ein und überprüfen dann Fachtexte aus unseren beiden Beispiel-Fachdisziplinen auf ihre Eignung für eine automatische Generierung.

Textgenerierungssysteme<sup>5</sup> finden heutzutage bereits eine Vielzahl von Anwendungen, wie etwa bei der natürlichsprachlichen Ausgabe von Datenbankinhalten z. B. als Auskunftssysteme oder Reportgeneratoren (Wetter-, Börsenberichte), als Schnittstelle zu Expertensystemen, als Generierungskomponenten von Dialogsystemen (z. B. zur Beratung) oder als Textgeneratoren (z. B. Bildbeschreibungen, Vertextung von Information aus Wissensbasen). Voraussetzung für jegliche Art der Generierung ist, daß die als Text zu generierende Information als formale, computerlinguistisch verarbeitbare Information vorliegt, wie z. B. Informationen aus Datenbanken, oder Wissensrepräsentationen. Der Generierungsprozeß selbst stellt eine komplexe Verarbeitung und Abbildung dieser Informationen in geeignete lexikalische, syntaktische, semantische und textuelle Strukturen dar. Traditionell wird der Generierungsprozeß in zwei Stufen unterteilt, die strategische Generierung, bei der der zu generierende Textinhalt festgelegt (geplant) wird, und die taktische Generierung, die die sprachliche (lexikalische, syntaktische, textuelle) Form der einzelnen Sätze bestimmt (vgl. z. B. Busemann/Novak 1995).

Die Möglichkeiten der heute verfügbaren Systeme beschränken sich in erster Linie auf die Generierung technischer bzw. sachbezogener Texte, da hier eine Modellierung der im Text kommunizierten Informationen und ihre Abbildung in sprachliche Strukturen eher möglich ist als in Texten mit abstraktem Textgegenstand oder komplexerer kommunikativer Struktur. Die Verarbeitung solcher vorwiegend propositionalen Information wird typischerweise durch Wissensrepräsentationssprachen, wie sie in der Künstlichen-Intelligenz-Forschung (KI) entwickelt wurden, unterstützt. Die Einbeziehung interpersonaler Information spielt in Anwendungskontexten eine Rolle, in denen Modellierung von Benutzerprofilen und -intentionen rele-

<sup>5</sup> Für einen Überblick vgl. z. B. Kempen (1989).

vant ist, wie beispielsweise in Erklärungskomponenten für Expertensysteme. Ein Schwerpunkt liegt jedoch generell auf der Generierung propositionaler Information.

Hieraus ergeben sich einige Konsequenzen und Beschränkungen für die Sorten von Texten, die potentiell für eine automatische Generierung in Frage kommen. Als am besten geeignet haben sich sachbezogene Texte mit klarer, möglichst standardisierter Makrostruktur und klarer rhetorischer Organisation erwiesen. Typische Textsorten, die bereits automatisch produziert werden können, sind Wetter-, oder Börsenberichte und verschiedene Arten von technischer Dokumentation wie Handbücher oder Bedienungsanleitungen. Obwohl die Qualität der maschinell generierten Texte zunehmend besser wird, ist häufig noch immer eine Nachedition durch den Menschen erforderlich, da insbesondere die Modellierung von Textphänomenen (wie z. B. Anaphern, die Thema-Rhema-Struktur oder die Informationsverteilung im Satz) den meisten Generierungssystemen noch erhebliche Schwierigkeiten bereitet. Die maschinell erzeugten Texte stellen also meist Rohfassungen dar, die dann noch einmal überarbeitet werden müssen.

Wie ist nun unsere wissenschaftliche Fachtextsorte „Zeitschriftenartikel in einer Wochen- oder Monatsschrift“, die wir oben in bezug auf ihr unterschiedliches linguistisches Erscheinungsbild in den Fachdisziplinen Medizin und Jura untersucht haben, im Hinblick auf die Einsatzmöglichkeiten der automatischen Textgenerierung einzuschätzen? Die Analyse medizinischer Fachartikel in Kapitel 3 hat eine starke Sachorientierung mit signifikanter Häufung definierter Fachterminologie, eine relativ distanzierte Haltung des Autors bzw. der Autoren zu dem im Text dargelegten Sachverhalt, d. h. eine vorwiegend propositionale Ausrichtung ergeben. Ferner wiesen die Texte eine standardisierte globale Textorganisation auf, die zudem mit charakteristischen rhetorischen Mustern auf Satzebene korrelierten. Damit sind die wichtigsten Bedingungen für eine automatische Textgenerierung erfüllt. Wie ein solches Generierungssystem aussehen könnte, wird im folgenden Kapitel skizziert.

Anders sieht es jedoch in bezug auf die juristischen Fachartikel aus. Obwohl es auch hier um Fälle und Sachverhalte geht, steht nicht in erster Linie ihre (neutrale) Vermittlung im Mittelpunkt, sondern die Sicht, die Perspektive, die der Autor dazu einnimmt und mit argumentativen und evaluativen Mitteln darlegt. Die juristischen Texte wiesen eine viel stärker metakommunikative Orientierung auf, die nicht ohne weiteres mit den üblichen Instrumentarien der Wissensmodellierung im KI-Sinne zu erfassen wäre. Dies spiegelte sich in den wesentlich komplexeren und variablen stilistischen Mustern auf Satzebene wieder. Es wäre daher wesentlich schwieriger, die linguistischen Eigenschaften der Texte in einem Textgenerator zu simulieren. Das heißt juristische Fachartikel kommen wegen ihrer komplexen rhetorischen Organisation und ihrer weniger starken Sachorientierung für eine automatische Generierung kaum in Frage. Interessant für das Fachgebiet Jura kann der Einsatz von Computertechnologie jedoch in solchen Zusammenhängen sein, in denen es

eher um die juristische Normsprache und nicht die dogmatische Sprache, wie in den untersuchten Fachartikeln, geht. Hier sei insbesondere auf die Nutzung von Datenbanken zur Definition juristischer Begriffe oder zu Gesetzestexten verwiesen (so z. B. JURIS, vgl. Cyrus/Wild 1990).

## 5 Skizze eines Systems zur Generierung medizinischer Fachtexte

Wir wollen nun am Beispiel eines existierenden Textgenerators skizzieren, wie eine Architektur zur Generierung medizinischer Fachtexte aussehen könnte, die sowohl die Eigenschaften der Makrostruktur dieser Fachtextsorten als auch die typischen lexikalischen und syntaktischen Muster auf Satzebene reflektiert. Als Szenario dient uns der Entwurf einer Rohfassung (*Drafting*) medizinischer Fallbeschreibungen. Die Generierungsplattform, an der der Ansatz illustriert werden soll, ist das multilinguale Textgenerierungssystem *KPML*, bzw. dessen deutsche Komponente *Komet-Penman*, das am Institut für Integrierte Publikations- und Informationssysteme (IPSI) der GMD Darmstadt entwickelt wurde (vgl. Teich 1992 und Bateman 1995). Der Generator wird in anwendungsorientierten Forschungskontexten bereits zur Generierung von Texten genutzt, die ähnliche Textsorteneigenschaften haben, wie unsere medizinischen Beispieltexte, so z. B. für die Generierung von Lexikoneinträgen für eine Kunst- und Künstlerenzyklopädie (vgl. Bateman/Teich 1995).

Die Systemarchitektur orientiert sich in wichtigen Bereichen an Ideen der systemisch-funktionalen Linguistik (SFL) (vgl. Halliday 1994). Wichtig für die Erfassung sowohl der satzgrammatischen als auch der textuellen Merkmale einer Textsorte ist die Repräsentation der stratifikationellen Organisation des Sprachsystems, wie sie in der systemisch-funktionalen Linguistik postuliert wird: Eine sprachliche Äußerung wird als das Ergebnis eines komplexen Auswahlprozesses aufgefaßt, der rekursiv unter den Optionen der linguistischen Abstraktionsebenen (*Strata*) von Lexikogrammatik, (Satz-)Semantik, Register und Genre wählt.<sup>6</sup> Textsorteneigenschaften werden auf der Ebene von *Genre* und *Register* (als Enkodierungen der Kontexte von Kultur und Situation) repräsentiert. Die typische Makrostruktur von Texten eines Genres, d. h. ihre typische globale syntagmatische Organisation, wird als sog. *Generic Structure Potential (GSP)* erfaßt (vgl. Hasan 1984). Eine GSP-Struktur besteht aus den Textsegmenten (*stages*), die für die Texte eines Genres charakteristisch bzw. obligatorisch sind. Linguistische Ressourcen auf allen *Strata* werden als *systemische Netzwerke* dargestellt, d. h. als paradigmatische Auswahl-systeme, die die Wahlmöglichkeiten unter verschiedenen Typen linguistischer Information repräsentieren.

---

6 Für einen Überblick über das systemisch-funktionale Sprachmodell vgl. Matthiessen/Bateman (1991).

Unsere Textsorte „medizinische Kasuistik“ ist in dieser Architektur in einer Taxonomie von Textsorten, einem systemischen Netzwerk von Genres repräsentiert.<sup>7</sup> Die mit dieser paradigmatischen Einordnung der Textsorte assoziierte globale (syntagmatische) Textorganisation entspricht in etwa der folgenden GSP-Struktur:

GSP	
1	Zusammenfassung
2	Kasuistik
2.1	Anamnese
2.2	Befundung
2.3	Therapie und Verlauf
3	Diskussion

Tabelle 2: Typische GSP-Struktur medizinischer Kasuistiken

Die Satzgenerierung erfolgt auf der Basis von Konzepten einer taxonomischen Wissensrepräsentation, die in einem sog. Domänenmodell die domänenspezifischen lexikalischen Konzepte enthält (vgl. Bateman/Kasper/Moore/Whitney 1990). Die Satzgrammatik, die als komplexes Auswahlnetzwerk zwischen einer Vielzahl grammatischer Optionen auszuwählen hat, verarbeitet als Input Satzrepräsentationen in Form semantischer Prädikat-Argument-Strukturen, die die sprachlich zu realisierenden Konzepte in formal-logischer Notation enthalten.

Ein altbekanntes Problem in der Textgenerierung ist die Interaktion und Kontrolle zwischen Makro- und Mikroplanung, d. h. die Interaktion zwischen der globalen Textplanung und der lexikogrammatistischen Realisierung der einzelnen Sätze, das auch als sog. „*generation gap*“ bezeichnet wird (vgl. Meteor 1992). Gemeint sind hiermit u. a. die Probleme bei einer Kontrolle der textuellen Organisation auf Satzebene (z. B. in bezug auf die Thema-Rhema-Struktur oder die Informationsverteilung im Satz). Hier können wir uns die Ergebnisse unserer Textanalysen zunutze machen: Die Untersuchung der Makro- und Mikrostruktur der medizinischen Beispieltex-te hat klare Korrelationen zwischen der kommunikativen Funktion eines Textsegmentes/GSP-Stages bzw. des Kommunikationsinhaltes und seiner satzgrammatischen Realisierung ergeben. Diese Information kann dazu dienen, die für ein Textsegment charakteristische lokale Textorganisation auf Satzebene zu kontrollieren, wenn sie entsprechend mit den GSP-Stages assoziiert wird.

<sup>7</sup> Den theoretischen Hintergrund für das Genre-Netzwerk, das dem KPML-System zugrunde liegt, bildet die Arbeit von Martin (1992).

Wir wollen dies kurz anhand unserer Beispieltex-te illustrieren (vgl. Textanalysen in Kap. 3):

- Die Generierung des Textsegmentes „Anamnese“ kann sich an der sequentiellen Verbalisierung der für dieses GSP-Stage obligatorischen Domänenkonzepte „Alter“ und „Geschlecht“ des Patienten, „Krankheitssymptome“, „Entwicklung der Symptome“, „Vorerkrankungen“ und „Begleiterkrankungen“ orientieren.
- Für die nachfolgende Befundung ist die Auflistung von Befundungskriterien und ihren Werten kennzeichnend, und zwar als zwei Typen von Thema-Rhema-Struktur:
  - a) die Auflistung der Normwerte als Rhema des ersten Satzes (vgl. Beispiel (1)) und
  - b) die Nennung der von der Norm abweichenden Befundungswerte, wobei das Befundungskriterium jeweils in thematischer und der Wert in rhematischer Position erscheint (vgl. Beispiel (2)). Die Generierungsgrammatik kann die Wortstellung in diesem Textsegment entsprechend kontrollieren.
- Im GSP-Stage „Therapie und Verlauf“ sind temporale Konstruktionen (vor allem Präpositionalphrasen und Adverbien, vgl. Beispiel (4)) in thematischer Position charakteristisch, wie unsere Textanalysen zeigen haben. Die Wortstellungskomponente der Satzgrammatik erhält also eine entsprechende Anweisung, in diesem Teil des Textes temporale Elemente an den Satzanfang zu stellen.
- Das Kapitel „Diskussion“ schließlich zeichnet sich durch die Verwendung von Prädikationen aus, die Einschätzungen auf einen Sachverhalt zum Ausdruck bringen (wie „möglich sein“, „denkbar sein“, „wahrscheinlich sein“; vgl. auch Beispiel (5)) und hierfür auf eine relativ begrenzte Menge von Lexemen (Adjektive) zurückgreifen. Dies kann bei der lexikalischen Selektion der Verben bzw. der Prädikate, die in diesem Segment generiert werden, berücksichtigt werden.

Die zu generierende Information könnte dem System beispielsweise mit Hilfe einer Eingabemaske verfügbar gemacht werden, die die für die einzelnen GSP-Stage die jeweils relevante Information in strukturierter Form erfaßt und auf die entsprechenden Konzepte der Wissensrepräsentation abbildet. Eine vom Verarbeitungsaufwand her komplexere Alternative, für die es jedoch bereits für andere medizinische Textsorten Prototypen gibt, stellt die automatische Wissensakquisition aus gesprochensprachlichem Input dar (vgl. z. B. Schröder 1994).

Zum Abschluß wollen wir noch auf eine Erweiterungsmöglichkeit verweisen, die vor allem im Zusammenhang mit der oben erwähnten Internationalisierung der wissenschaftlichen Fachkommunikation interessant ist, nämlich die multilinguale Generierung von Texten. Multilinguale Textgenerierungssysteme wie beispielsweise das oben erwähnte *KPML*-System haben sich in letzter Zeit als interessante

Alternative zu mehrsprachigen maschinellen Übersetzungssystemen erwiesen, da die Wissensquellen für beide Arten von computerlinguistischen Anwendungen zum Teil sehr ähnlich sind (vgl. Hartley/Paris 1995; Teich 1995). Da die Medizin als naturwissenschaftlich orientierte Fachdisziplin, wie oben erwähnt, deutlich anglophon geprägt ist, kann ein simultanes Drafting einer englischen und deutschen Version eines Fachartikels, insbesondere für Wissenschaftler, deren Muttersprache nicht Englisch ist, einen Beitrag zur Verbesserung des internationalen wissenschaftlichen Austauschs liefern.

## 6 Resümee

In unserem Beitrag haben wir kommunikative und textuelle Eigenschaften medizinischer und juristischer Fachaufsätze untersucht. Die beiden Disziplinen zeichnen sich durch deutliche Unterschiede sowohl in ihrem Handlungs- als auch Kommunikationsbereich aus, was zu charakteristischen sprachlichen Reflexen in den jeweiligen Fachtexten führt. Die primär informierende kommunikative Funktion der medizinischen Texte korreliert mit einer faktenzentrierten Darstellung der Sachverhalte und wird entsprechend mit Hilfe des deskriptiven Texttyps realisiert. Die dem medizinischen Handlungsbereich zugrundeliegende Strukturierung (,Anamnese‘, ,Befundung‘, ,Therapie und Verlauf‘ etc.) dient gleichzeitig als Vorlage für die Makrostruktur der Texte. Ganz anders bei den rechtswissenschaftlichen Zeitschriftenaufsätzen, deren kommunikative Funktion primär evaluativer und persuativer Natur ist und daher bei der sprachlichen Realisierung der argumentative Texttyp bevorzugt wird. Für die sprachliche Umsetzung dieser unterschiedlichen kommunikativen Funktionen haben wir charakteristische lexikalische und syntaktische Muster identifiziert.

Ein Ziel unserer Textanalysen war es abzuschätzen, inwiefern computerlinguistische Werkzeuge, wie z. B. Textgenerierungssysteme, den wissenschaftlichen Textproduktionsprozeß unterstützen können und welche Rolle dabei die zugrundeliegende Domäne (Fachdisziplin) spielt. Wegen ihrer Sachbezogenheit und standardisierten Makrostruktur sind medizinische Fachartikel eher für eine automatische Generierung geeignet als juristische Aufsätze, die erhebliche Probleme sowohl bei der Wissensmodellierung als auch bei der Kontrolle der Textstruktur bereiten. Um dies beispielhaft zu illustrieren, skizzierten wir im letzten Kapitel ein Textgenerierungssystem zur Erzeugung von Rohfassungen medizinischer Fachartikel.

## Literatur

- Bateman, John A./ Kasper, Robert T./ Moore, Johanna D./ Whitney, Richard A. (1990): A general organization of knowledge for natural language processing: the PENMAN upper model. Marina del Rey CA: USC/Information Sciences Institute [Technical report]
- Bateman, John A. (1995): KPML: The KOMET-Penman multilingual resource development environment. In: Proceedings of the Fifth European Workshop Natural Language Generation. Leiden: University of Leiden, Faculty of Social and Behavioural Sciences, 219-222
- Bateman, John A./ Teich, Elke (1995): Selective Information Presentation in an Integrated Publication System: An Application of Genre-Driven Text Generation. In: Information Processing and Management 5 (31), 753-767
- Benes, Eduard (1969): Zur Typologie der Stilgattungen der wissenschaftlichen Prosa. In: Deutsch als Fremdsprache (6), 225-233
- Busch-Lauer, Ines-Andrea (1991): Englische Fachtexte in der Pädagogischen Psychologie. Frankfurt/Main u. a.: Lang
- Busemann, Stephan/ Novak, Hans-Joachim (1995): Generierung natürlicher Sprache. In: Görz, Günther (Hrsg.): Einführung in die Künstliche Intelligenz. 2. Aufl. Bonn: Addison-Wesley, 492-557
- Cyrus, Christoph/ Wild, Yves (1990): JURIS leichtgemacht. Der Weg zur erfolgreichen Recherche. 2. Aufl. Düsseldorf: Werner-Verlag
- Halliday, Michael A. K. (1994): An Introduction to Functional Grammar. 2. Aufl. London: Arnold
- Hartley, Anthony F./ Paris, Cecile L. (1995): Supporting Multilingual Document Production: Machine Translation or Multilingual Generation? In: Proceedings of the IJCAI'95 Workshop on Multilingual Generation, Montreal, Canada, August 1995
- Hasan, Ruqaiya (1984): The nursery tale as a genre. In: Nottingham Linguistic Circular (13), 71-192
- Hoffmann, Lothar (1985): Kommunikationsmittel Fachsprache. Eine Einführung. Tübingen: Narr
- Kempen, Gerard (1989): Language generation systems. In: Bátori, Istvan S./ Lenders, Winfried/ Putschke, Wolfgang (Hrsg.): Computational Linguistics – An International Handbook on Computer Oriented Language Research and Applications. Berlin, New York: de Gruyter, 471-480
- Lavid, Julia (1994): Thematic Development in Texts. Madrid: Universidad Complutense de Madrid [Technical report, Dandelion, Esprit Basic Research Project 6665, Deliverable R1.2.1.]
- Matthiessen, Christian M.I.M./ Bateman, John A. (1991): Text Generation and Systemic-Functional Linguistics: Experiences from English and Japanese. London: Frances Pinter
- Martin, James R. (1992): English Texts. System and Structure. Philadelphia, Amsterdam: Benjamins
- Meteer, Marie W. (1992): Expressibility and the problem of efficient text planning. London: Pinter Publishers
- Oksaar, Els (1988): Fachsprachliche Dimensionen. Tübingen: Narr
- Podlech, Adalbert (1975): Die juristische Fachsprache und die Umgangssprache. In: Petöfi, János/ Podlech, Adalbert/ Savigny, Eike von (Hrsg.): Fachsprache – Umgangssprache. Wissenschaftstheoretische und linguistische Aspekte der Problematik, sprachliche Aspekte der Jurisprudenz und der Theologie, maschinelle Textverarbeitung. Kronberg/Ts.: Scriptor Verlag, 161-189
- Ramm, Wiebke/ Rothkegel, Annely/ Steiner, Erich/ Villiger, Claudia (1995): Discourse grammar for German. Saarbrücken: Universität des Saarlandes [Technical report, Dandelion, Esprit Basic Research Project 6665, Deliverable R2.3.2.]
- Schröder, Martin (1994): Interpreting Clinical Questions – Medical Text Analysis Supports Image Presentation. In: Barth, G./ Günter, A./ Neumann, B. (Hrsg.): KI-94, Anwendungen der Künstlichen Intelligenz. Berlin u. a.: Springer-Verlag, 53-68
- Skudlik, Sabine (1990). Sprachen in den Wissenschaften. Deutsch und Englisch in der internationalen Kommunikation. Tübingen: Narr

- Teich, Elke (1992): KOMET: Grammar Documentation. Darmstadt: GMD/IPSI [Technische Dokumentation]
- Teich, Elke (1995): Towards a Methodology for the Construction of Multilingual Resources for Multilingual Generation In: Proceedings of the IJCAI'95 Workshop on Multilingual Generation, Montreal
- Virtanen, Tuija (1992): Discourse Functions of Adverbial Placement in English. Clause-Initial Adverbials of Time and Place in Narratives and Procedural Place Descriptions. Abo: Abo Akademis Förlag
- Werlich, Egon (1979): Typologie der Texte: Entwurf eines textlinguistischen Modells zur Grundlegung einer Textgrammatik. Heidelberg: Quelle und Meyer

## Quellen

- Text 1: M. Thomas et al. (1996): Bronchiolitis obliterans mit organisierender Pneumonie. Differentialdiagnose der antibiotikarefraktären Pneumonien. In: Deutsche Medizinische Wochenzeitschrift (121), 200-204
- Text 2: Joachim Hruschka (1996): Die Blockade einer Autobahn durch Demonstranten – eine Nötigung? In: Neue Juristische Wochenschrift 3 (49), 160-164
- Text 3: Gerhard Czermak (1995): Der Kreuzifix-Beschluß des Bundesverfassungsgerichts, seine Ursachen und seine Bedeutung. In: Neue Juristische Wochenschrift 51 (48), 3348-3353