

Rezensionen

Stefan Schierholz:

Lexikologische Analysen zur Abstraktheit, Häufigkeit und Polysemie deutscher Substantive, Niemeyer Verlag 1991, DM 108,

Was ist ein abstraktes Substantiv? Wann tritt ein Nomen häufig auf, warum ist es polysem? Vorweg: Wer Antworten auf solche Fragen sucht, wird das vorliegende Buch (eine Göttinger Dissertation) rasch wieder weglegen. Dies ist keine lexikologische Studie in dem Sinn, daß Reflexionen über die semantischen Eigenschaften einzelner Wörter angestellt würden, die Auslöser für Häufigkeit oder Polysemie wären. Ziel dieser Arbeit ist es vielmehr, mit Hilfe quantitativer Methoden Zusammenhänge zwischen der (operational definierten) Abstraktheit, Häufigkeit und Polysemie von Nomina zu untersuchen. Als Basis hierfür dienen Nomina, die entweder als Stichwörter oder als sog. Kernkonstituenten der Bedeutungsparaphrasen in den Einträgen eines einsprachigen deutschen Wörterbuchs auftreten (im wesentlichen dem Duden-Deutschen Universalwörterbuch, 1. Auflage 1982).

"Lexikologisch" sind die vorgenommenen Untersuchungen also insofern, als das Material einem Lexikon entnommen ist und die Untersuchungsergebnisse Rückschlüsse auf den Aufbau dieses Wörterbuchs zulassen.

Zu Beginn der Arbeit wird festgelegt, welche Teile des Universalwörterbuchs als statistische Datengrundlage dient, wie Abstraktheit, Häufigkeit und Polysemie gemessen werden und welche Zusammenhänge zwischen diesen Größen untersucht werden. Es folgen eine Dar-

stellung der Untersuchungsergebnisse sowie eine Diskussion, welchen Einfluß Versuchsordnung und Datenbasis auf die Ergebnisse hatten.

Zunächst aber die Hypothesen, die zu Beginn der Arbeit als Gegenstand der empirischen Untersuchung postuliert werden:

1. je häufiger ein Wort, desto abstrakter ist es
2. je häufiger ein Wort, desto mehrdeutiger ist es
3. je mehrdeutiger ein Wort, desto abstrakter ist es

Anstelle einer Herleitung oder Begründung dieser Hypothesen wird nur kurz auf Meier 1967 und Zipf 1949 verwiesen. Die ersten beiden Hypothesen leuchten jedoch sofort ein: Bestimmte Wörter sind deshalb häufig, weil sie ein breites Spektrum von Gegenständen oder Sachverhalten denotieren; ein breites Spektrum entspricht einem abstrakten (im Sinne von allgemeinen) Begriff. Zieht man in diesem Spektrum Grenzen, ergibt sich ein hohes Maß an Polysemie: für jeden Abschnitt des Spektrums eine Lesart. Die dritte Hypothese schließlich wirkt wie ein Umkehrschluß aus 1 und 2: wenn mehrdeutig, dann auch häufig und somit abstrakt.

Umso erstaunlicher ist, daß von den drei Hypothesen zwei falsifiziert werden, wie sich unten ergeben wird. Es lohnt sich also, genau hinzusehen, wie Abstraktheit, Häufigkeit und Polysemie denn hier definiert und gemessen werden.

Beginnen wir mit der Abstraktheit. Zunächst gilt es, ein Mißverständnis zu vermeiden: Abstraktheit wird im statistischen

Teil der Arbeit nicht, wie in der Semantik üblich, als Gegensatz zu Konkretheit verstanden, sondern als Allgemeinheit: ein Nomen gilt dann als abstrakt, wenn es in der Hierarchie der Definitionen per genus proximum et differentiam specificam, die vielen Bedeutungsexplikationen zugrunde liegen, möglichst weit oben steht. Besonders abstrakte Nomina sind daher beispielsweise *Wesen*) *Mensch*) *Ding*) *Stoff*, *Weise*; wenig abstrakte *Klagelaut*) *Küche* oder *Lebensalter*. Abstraktheit als Gegensatz zu Spezialisierung (*Tier* - *Yorkshireterrier*) korreliert nicht mit Abstraktheit als Gegensatz zu Konkretheit (*Geist* - *Körper*- wie auch immer man hier die Grenze zieht). - Bei der Diskussion der Resultate scheint Schierholz dann allerdings die Interpretation von Abstraktheit als Gegensatz zu Konkretheit zu bevorzugen.

Wie mißt man Abstraktheit? Schierholz unterstellt, daß ausreichend vielen Paraphrasen im DDUW eine klassische Definitionsstruktur zugrunde liegt, und bestimmt zunächst die sog. "Kernkonstituenten" einer Explikation. Darunter versteht er diejenigen Nomina, die sich in den Substitutionsrahmen *Ein* <Lemma> *ist ein*< Erklärungswort> einsetzen lassen, unterscheidet also nicht zwischen Hyperonymen, Hyponymen und Synonymen. Da sich diese Prozedur nicht fehlerfrei maschinell durchführen läßt (unterstellt, man hätte eine maschinenlesbare Version des Lexikons), zieht Schierholz eine exakt festgelegte Stichprobe, die er manuell untersucht; grob gesagt, das erste Nomen auf jeder Seite des Lexikons. Für die 1482 untersuchten Nomina erhält er 1110 verschiedene Kernkonstituenten, von denen die meisten wiederum als Stichwort im Lexikon belegt sind. Durch rekursive Anwendung der Prozedur auf die Definitionen der Kernkonstituenten ergeben sich Definitionsketten; eine der längsten lautet:

Tischfernsprecher - *Tischtelefon* -
Telefon - *Apparat* - *Gerät* - *Gegenstand*
 - *Ding*.

Ding wird wieder durch *Gegenstand* erklärt, so daß die Kette wegen Zirkularität abbricht.

An dieser Stelle erhält man die ersten Resultate zur Struktur des Lexikons: im Deutschen sind die Ketten länger als etwa im Französischen, weil dort die typisch deutschen Bandwurmkomposita wegfallen, die zu Beginn der Ketten stehen können. In nicht zu vielen Fällen werden die Ketten zirkulär, in viel zu vielen Fällen münden sie in Nominalisierungen, die von Schierholz grundsätzlich nicht als Oberbegriffe akzeptiert werden: *Beschaffenheit*: *das Beschaffensein* einer *Sache*. Das naheliegendste Maß für die Abstraktheit eines Nomens *x* wäre nun die Zahl der Elemente, die in den Ketten vor *x* stehen; *x* kann jedoch in verschiedenen Ketten verschieden viele Elemente vor sich haben. Es wird daher gezählt, wieviele Wörter an *n*-ter Stelle irgendeiner Definitionskette stehen; als Abstraktheitswert von *x* wird dann der Durchschnitt über die relative Häufigkeit von *x* auf den *n* verschiedenen Ebenen berechnet.

Polysemie und Häufigkeit lassen sich schneller erklären: Der Grad der Polysemie eines Wortes wird festgelegt durch die Zahl der unterschiedlichen Kernkonstituenten in den Explikationen seiner Lesarten; am mehrdeutigsten ist *Zug* mit 29 Kernkonstituenten. Als Häufigkeit eines Nomens wird bezeichnet, wie oft es in den untersuchten Explikationen auftritt; das häufigste Nomen ist *Teil*.

Heraus kommt, daß die Hypothese 1 von den statistischen Daten gestützt wird: wenn häufig, dann auch abstrakt. Die Hypothesen 2 und 3 werden dagegen falsifiziert.

Nun kann man diese Falsifikation in zwei Richtungen deuten: entweder die Hypothesen stimmen tatsächlich nicht, oder es liegt an der Untersuchungsmethode. Fangen wir mit letzterem an.

An dieser Stelle ist hervorzuheben, daß die vorliegende Arbeit sich in hohem Maße durch Präzision und Übersichtlichkeit aus

zeichnet. Der Untersuchungsgegenstand wird exakt definiert, das Verfahren erscheint absolut nachvollziehbar. Dieses Ideal empirischer Untersuchungen erleichtert dem Rezensenten seine Kritik, scheint mir aber ausgerechnet der Pferdefuß zu sein, der die Untersuchungen zu Fall bringt: weil nämlich operationalisiert wurde, was nicht so einfach zu operationalisieren ist. Die statistische Analyse gerät zu abstrakt, geht nicht auf die inhaltlichen Hintergründe ein.

Es beginnt mit den Kernkonstituenten: Nicht umsonst gibt es einen Unterschied zwischen Hyperonymen und Synonymen, und wenn Abstraktheit über die Länge der Definitionsketten definiert wird, dann dürfen diese Ketten nur Hyperonyme enthalten und keine Synonyme. Den größten Ausreißer stellt das häufigste aller Nomina dar, nämlich *Teil*, das meist als Indikator für eine partonymische Explikation fungiert und nicht ohne weiteres mit zu den Kernkonstituenten gezählt werden dürfte: eine *Abschlagszahlung* ist kein *Teil*, auch wenn sie *Teil einer zu leistenden Zahlung* ist. Zumindest die Unterscheidung von Hyperonymen und Synonymen hätte durch leichte Verfeinerung des Substitutionstests (*Jedes <Lemma> ist...*) berücksichtigt werden können; bei *Teil*, *Art* usw. wird es schon schwieriger.

Bei der Bestimmung der Polysemie fällt auf, daß Schierholz nicht etwa die im Lexikon durch Nummern unterschiedenen Lesarten zählt, sondern die unterschiedlichen Kernkonstituenten. In manchen Fällen fallen dadurch Lesarten weg, z.B. bei

- Notation:*
1. *Aufzeichnung; System von Zeichen oder Symbolen einer Metasprache*
a) *das Aufzeichnen von Musik in*
 2. *Notenschrift;*
b) *Notenschrift*
 3. *Aufzeichnung der einzelnen Züge einer Schachpartie*

Gewertet werden *Aufzeichnung*, *System*

und *Notenschrift*, Lesart 3 fällt flach, weil sie dieselbe Kernkonstituente aufweist wie Lesart 1, und Lesart 2a enthält kein "ordentliches" Hyperonym. Lesart 1 wird dafür gleich doppelt gezählt.

Als Begründung dafür, daß er nicht die vorgegebene Einteilung übernimmt, nennt Schierholz die vermeintliche Unsitte der Lexikographen, nicht-synonyme Paraphrasen wie *Aufzeichnung* und *System von Zeichen* unter eine Lesart zu subsumieren. Was dem Statistiker zuviel an Variation ist, ist dem Lexikologen allerdings noch zuwenig: Schließlich stellen Lesarten keine objektiven Entitäten dar, sondern Abstraktionen aus Verwendungskontexten, die keineswegs unbedingt durch eine einzelne Paraphrase beschrieben werden können, eigentlich eher noch ausführlicher dargestellt werden müßten.

Hinzu kommt, daß lexikalische Explikationen zwar sorgfältig gemacht sind, aber ihre eigenen Gesetze haben. Beispiele: Zwar wurde die Hypothese 1 (wenn häufig, dann abstrakt) im ganzen nicht widerlegt, einige zentrale Lexeme fallen aber aus der Reihe: *Gesamtheit* steht an 4. Stelle der Häufigkeitsliste, aber an 175. der Abstraktheitskala. *Gruppe* hat den Häufigkeitsrang 10, den Abstraktheitsrang 195. Bei beiden hat man das Gefühl, der Häufigkeitsrang sei - bedingt durch das Aufbauschema von Explikationen - unangemessen hoch, der Abstraktheitsrang dagegen zu niedrig, vielleicht weil die Hyponyme von *Gruppe* oder *Gesamtheit* weniger durch lexikalisierte Komposita ausdifferenziert sind als etwa die von *Tier* (Rang 13), *Einrichtungsgegenstand* (Rang 32) oder *Amt* (Rang 77). (Selbst *Schreibpapier* und *Foto* haben einen doppelt so hohen Abstraktheitsgrad wie *Gesamtheit!*) - Im übrigen konstituieren auch *Gesamtheit* und *Gruppe* in den meisten Fällen Partonymie- und keine Hyponymierelationen und müßten daher differenzierter betrachtet werden. Wo viele Abstraktheitswerte intuitiv unangemessen erscheinen, ist auch das Scheitern der dritten

Hypothese nicht erstaunlich: *Element, Modell, Grund* zählen als stark polysem, aber nicht als abstrakt. Allein *Größe* weist beide Merkmale auf. Bei den häufigen Substantiven schließlich ist es allein *Linie*, das eine hohe Mehrdeutigkeit aufweist - in einem Lexikon sind die häufigsten Wörter eben *Person, Mensch, Vorrichtung, Pflanze*, und die sind nicht mehrdeutig.

Schierholz selbst verteidigt seine Hypothesen, indem er auf die Mängel der Datengrundlage verweist, insbesondere auf den hohen Anteil von Nominalisierungen des Typs *Beschaffensein* - eine sicher berechtigte Kritik.

Wären die Hypothesen zu retten, wenn "subjektivere" Bewertungen bei der Berechnung von Abstraktheit und Polysemie einflößen? Zumindest bei der Korrelation dieser Größen mit der Häufigkeit ist zu beachten, daß ein Substantiv nicht einfach "häufig" ist, sondern häufig relativ zu einem bestimmten Korpus; und die Schierholz'schen Kernkonstituenten stellen eben ein stark fachsprachliches Korpus dar, das andere Häufigkeitsverteilungen aufweist als etwa die deutsche Zeitungssprache. Und zwischen Abstraktheit und Polysemie besteht vielleicht tatsächlich kein zu enger Zusammenhang.

Was bleibt? Für quantitativ orientierte Linguisten eine Arbeit, die sich durch ausführlich begründete, detaillierte statistische Analysen auszeichnet; die von ihm verwendeten, hier nicht dargestellten statistischen Methoden zur Berechnung und Korrektur der Korrelationskoeffizienten will Schierholz als exemplarisch für quantitativ-linguistische Untersuchungen verstanden wissen.

Für den nicht quantitativ arbeitenden Computerlinguisten ist bestenfalls Schierholz' Vorschlag interessant, ein "Polyfunktionalitätswörterbuch" aufzustellen, das angibt, wie oft ein Lemma als Erklärungswort gebraucht wird. Mit anderen Worten, die einem Lexikon zugrundeliegende Genushierarchie müßte einmal vollständig rekonstruiert werden (unter getrennter Berücksichtigung der Synonyme und Partonyme, natürlich). Die kritische Durchsicht dieser Hierarchie wäre nicht nur für die Bewertung und Weiterentwicklung eines Wörterbuchs interessant, sondern ließe vielleicht auch eine bessere Beurteilung der hier untersuchten Hypothesen zu. Und was vor allem bleibt, ist ein Appell an die deutschen Wörterbuchverlage, ihre Daten in maschinenlesbarer Form endlich freizugeben, damit Untersuchungen wie die vorliegende und die skizzierte in größerem Maßstab durchgeführt werden könnten als nur mit dem ersten Nomen auf jeder Seite.

Literatur:

Meier, Helmut 1967: *Deutsche Sprachstatistik*. Hildesheim, 2. Aufl.. 1967.

Sambor, Jadwiga / Hammer!, Rolf (Hrsg.) 1991: *Definitionsfolgen und Lexemnetze*. Band I. Lüdenscheid: Richter Altmann Medienverlag.

Zipf, George K. 1949: *Human behavior and the principle of least effort*. Cambridge; Mass.

Stephan Mehl, Universität Duisburg

Den Hintergrund für Schierholz' Arbeit den er allerdings mehr andeutet als ausführlich darstellt - bildet ein Zweig der quantitativen Linguistik (vgl. Sambor/Hammerl 1991), der sich mit stochastischen Eigenschaften der Lexik befaßt und für den die Falsifikation der Schierholz'schen Hypothesen von Bedeutung wäre.

Rezensionsangebote

Exemplare der folgenden Bücher liegen der Redaktion zur Besprechung vor. Kompetente LeserInnen, die gerne eines oder mehrere Bücher rezensieren wollen, setzen sich bitte mit der Redaktion in Verbindung.

Leitner, Gerhard (ed): *New Directions in English Language Corpora: Methodology, results, software Developments.* (Topics in English Linguistics; Bd. 9). 368 S., Berlin, New York: Mouton de Gruyter, 1992

Part I: Corpus design and text encoding:

The optimum corpus sample size? (Pieter de Hahn), *Corpus sampling* (Jeremy Clear), *International Corpus of English: Corpus design - problems and suggested solutions* (Gerhard Leitner), *The Northern Ireland transcribed corpus of speech* (John Kirk), *Problems in the compilation of a corpus of standard Caribbean English: A pilot study* (Christian Mair), *The Text Encoding Initiative: A progress report* (Lou Burnard),

Part II: Automated syntactic and semantic text analysis:

Pinpointing problematic tagging decisions (Nancy Belmore), *Inferences and lexical relations* (Willem Meijs), *Tracing cohesive relations in corpora samples using a machine readable dictionary* (Sylvia Janssen), *Developing a scheme for annotating text to show anaphoric relations* (Steve Fligelstone), *SUSANNE - a deeply analysed corpus of American English* (Geoffrey Sampson), *Information retrieval and corpora* (Stephan Sutton and Anthony McEnery),

Part III: Corpora in language description:

Relative infinitives in spoken and written English (Christer Geisler), *Who(m)? Case marking of wh-pronouns in written British and American English* (Edgar W. Schneider), *Discourse category and text type classification: Procedural discourse in the Brown and LOB corpora* (Kay Wikberg), *Opaque and transparent features of Indian English* (S. V. Shastri), *Computer analysis of spelling variants in Chaucer's (Canterbury Tales,*

Geoff Barnbrook), Pitch contours and tones in the Lancaster IIBM spoken English corpus (Gerry Knowles), *What do you think of that: A pilot study of phraseology of the core words in English?* (Antoinette Renouf), *Sequences of spatial and temporal adverbials in spoken and written English* (Hilde Hasselgard), *Grammatical or native like?* (Göran Kjellmer), *Collocation and bilingual text* (Jacques Noel),

Menzel, Wolfgang: *Modellbasierte Fehlerdiagnose in Sprachlehrensystemen.* (Sprache und Information; Bd. 24). 230 S., Tübingen: Niemeyer, 1992

Der Entwurf fehlersensitiver Analysesysteme für größere Bereiche der natürlichen Sprache ist eine aufwendige und in zahlreichen Einzelfragen noch keinesfalls überschaubare Aufgabe. Dies gilt gleichermaßen für die antizipationsbasierten wie für die modellbasierten Ansätze. Die vorliegende Arbeit versucht, aus dieser Not eine Tugend zu machen: Ausgehend von einer tragfähigen Prinziplösung für sehr kleine, aber dennoch praktisch relevante Beispielprobleme wird durch schrittweise Hinzunahme zusätzlicher Phänomene eine wachsende Leistungsfähigkeit der Diagnoseverfahren angestrebt. Nicht für alle dieser Erweiterungsschritte liegen dabei schon praktische Erfahrungswerte aus 00-plementationsexperimenten vor. Dennoch ist es bereits möglich, eine vorläufige Bilanz der bisherigen Resultate zu ziehen. Wichtigstes Ergebnis der Untersuchungen ist sicherlich die Tatsache, daß eine modellbasierte Diagnose für Regularitäten der natürlichen Sprache überhaupt realisierbar ist. Dieses Ergebnis ist keineswegs trivial, da modellbasierte Verfahren im allgemeinen mit dem Ruf eines außerordentlich hohen rechentechnischen Aufwands verbunden sind. Hinzu kommt, daß die modellbasierte Diagnose nach einer hochgenauen Modellierung der korrekt / inkorrekt-Unterscheidung verlangt und erfahrungsgemäß jeder Versuch zur Durchsetzung dieser Forderung im Bereich der Syntax natürlicher Sprachen auf erhebliche Schwierigkeiten stößt. Erzielt

werden kann dieses Ergebnis dank der Tatsache, daß - für die Zwecke des Sprachunterrichts - eine Zerlegung der Grammatik in sehr kleine isolierte Teilbereiche möglich ist. Eine solche Beschränkung auf stark eingegrenzte Aufgaben erlaubt immer noch die Realisierung recht sinnvoller Anwendungsbeispiele und bringt darüber hinaus sogar einige nicht unerhebliche didaktische Vorteile mit sich: In derartig eingegrenzten Gebieten der Grammatik lassen sich Diagnoseergebnisse erzielen, die hinsichtlich ihrer Komplexität, Präzision, Vollständigkeit und Plausibilität den Ansprüchen, die sich bei einem Einsatz für die Zwecke des Fremdsprachenerwerbs ergeben, vollauf gerecht werden. Es konnte gezeigt werden, daß sich für einfache Modelle eine Diagnoseprozedur entwerfen läßt, die wegen ihrer verhältnismäßig geringen Aufwandsparameter auch auf relativ leistungsschwacher Hardware lauffähig ist. Die Untersuchungen konzentrierten sich vorrangig auf Beispiele mit einer einzigen Dominanzregel, bei denen die Vorteile der Arbeit in eingeschränkten Sprachbereichen besonders deutlich hervortreten. Für den Spezialfall der Übungen mit deutschen Präpositionalgruppen ergeben sich auch bei Einbeziehung von Anordnungsrestriktionen und Forderungen zur Dominanzverträglichkeit maximale Diagnosezeiten von ein bis zwei Sekunden auf einfachen Standard-PC's mit einer Taktfrequenz von unter zehn Megahertz. Damit stellt die verfügbare Hardwarekonfiguration kein entscheidendes Kriterium für die Anwendung modellbasierter Techniken dar. Bei dem Einsatz der beschriebenen Verfahren für vergleichbare einfache Übungen dürfte weder im Schul- noch im Heimbereich mit ernsthaften hardwareseitigen Schwierigkeiten zu rechnen sein.

Rothkegel, Annely: Text Knowledge and Object Knowledge. (Communication in artificial intelligence series). 194 S., London, New York: Pinter Publishers. April 1993
1. *Knowledge representation (Information Theory)* 2. *Computational Linguistics*

[Library of Congress Catalogin] Text production is taken here to be a process in which text knowledge and object knowledge interact. While the latter concerns the conceptual organization of world knowledge, the former refers to conventionally determined processes for the ordering and presentation of knowledge for communicative purposes. The model presented here introduces TEXT ACTIONS as instruments of representation. They link action-orientated and textual aspects of representation with lexical and syntactic realization. Their theoretical base is in speech act theory, and their function is to describe a text grammar with two levels, for the processes of selection and composition in the text itself. Selection is mainly determined by the knowledge frames that structure the various object fields. This is exemplified from the domain of book announcement texts, in which books are the objects. But selection also takes place in relation to the presentation of the objects. For example, this may take a descriptive, evaluative or argumentative form, and this in turn has consequences for the lexical and syntactic selections. Composition concerns the entire organization of the text, which is reflected ultimately in linearization. Various types of sequencing operation are introduced to specify the connectivity in the text (FRAMING; CHAINING and PAIRING). Finally all the elements are linked together in a TEXT SPACE model.

Contents: 1. Introduction. 2. What is meant by knowledge? 3. Text generation. 4. Text actions. 5. Text grammar I: Selections. 6. Text grammar II.1: Principles of composition. 7. Text grammar II.2: Sequencing operations (SOP). 8. Integrated view of the text grammar.

Rothkegel, Annely: Textualisieren. Theorie und Computermodell der Textproduktion. (Sprachwelten; Bd. 8), 262 S., Frankfurt/M., Berlin, New York, Paris, Wien: Verlag Peter Lang, 1993

Wenn wir Texte schreiben, z.B. um über andere Gegenstände oder Sachverhalte zu informieren, wenden wir komplexe Strate-

gien an, die eine sprachliche Spur im Text hinterlassen. Die Hypothese ist, daß die Textualisierung von Inhalten zu bestimmten kommunikativen Zwecken sprachliche Implikationen für die Sequenzierung und die damit korrespondierende Konnexionsstruktur des Textes hat. Textualisieren wird so als ein Prozeß aufgefaßt, in dem funktionale, thematische und sprachliche Aspekte aufeinander bezogen werden. Er ist linguistisch beschreibbar in Form von Texthandlungen, die sprechakttheoretisch begründet und im Rahmen eines Computermodells der Textproduktion operationalisiert werden. Das Modell wird exemplarisch anhand von Buchankündigungstexten entwickelt.

Aus dem Inhalt: Handlungsorientierte Texttheorie - Wissen im Text - Sequenzierungsstrategien und Textkonnexion - Textgrammatik und computerunterstützte Textproduktion.

Schweizer, Harald: THLI (Sprachkritik als Ideologiekritik. Zur Grammatikversion am Beispiel von QRB; Bd. 1). 173 S., Tübingen: Francke Verlag, 1991

Bader, Winfried: THLI (Simson bei Delila. Computerlinguistische Interpretation des Textes Ri 13-16, Bd. 3.). 468 S., Tübingen: Francke Verlag, 1991

Die Reihe THLI ("Textwissenschaft - Theologie - Hermeneutik - Linguistik - Literaturanalyse - Informatik") will vor allem der Methodenreflexion neue Impulse geben. Bei dem Versuch, eine Verstehenstheorie zu konzipieren, die Intersubjektivität im Interpretationsakt ermöglicht, erweisen sich die Instrumente der Informatik zunehmend als nützliches Hilfsmittel. Die Bände entwickeln innovative methodische Orientierungen, die - auch wenn sie hier im Bereich der Theologie angewandt werden - von allgemeinem Interesse sind.

Folgende Punkte sind hervorzuheben:

=> der von der Zeichendefinition her neu definierte SYNTAX-Begriff und die ergiebigen methodischen Folgerungen (incl. der computerlinguistischen Umsetzbarkeit): inhaltsfreie Morphologie,

erschöpfende Ausdrucksanalyse eines Textes

=> die damit zusammenhängende klare Trennbarkeit von SYNTAX und SEMANTIK, die eine entsprechend klar unterschiedene algorithmische Behandlung nach sich zieht (Distribution vs. Datenbank)

=> im Fall von THLI 1 (Schweizer: Sprachkritik als Ideologiekritik) außerdem die praktischen Erprobungen in einem vorgegebenen Rahmen zum Verhältnis SEMANTIK und PRAGMATIK. Die Vorteile dieses Konzepts: Der zeichentheoretischen Klarheit entspricht eine solche Programmentwicklung; beides bewährt sich in der praktischen Analyse und führt zu Ergebnissen, die neu und höchst anregend sind.

Immer wieder stößt man auf Ratlosigkeit bezüglich des Dreischritts SYNTAX-SEMANTIK- PRAGMATIK (Syntax nur partiell ausgenommen). Demgegenüber sind die beiden Bücher - nach Meinung des Verlags - einige Schritte weiter.

Außerdem haben Verlage mehrere Bücher vorgestellt, die der Redaktion umgehend zu Rezensionszwecken zur Verfügung gestellt werden können:

Dik, Simon C.: Functional Grammar in Prolog. An Integrated implementation for English, French and Dutch. (Natural Language Processing 2). 264 S., Berlin/ New York: Walter de Gruyter, 1992

Heintzeler, Mirjam: Raumausdrücke im Konzeptlexikon - Die Darstellung der Komposition lokaler Verben und Präpositionen in einem konzeptuellen Lexikon. 272 S., Konstanz: Hartung-Goue Verlag, 1992

Das Thema der Arbeit liegt zwischen Künstlicher Intelligenz und Sprachwissenschaft. Drei Themenbereiche - einer aus der KI, zwei aus der Sprachwissenschaft - bilden den Hintergrund der Arbeit: Maschinelle Übersetzung und Wissensrepräsentation; sprachlicher Ausdruck von Raum, insbes. sogenannte lokale Verben und Präpositionen; und Verbsemantik unter dem Blickwinkel der Zeitkonstitution. Aus der kritischen Diskus

sion der Forschungen zu diesen drei Bereichen werden folgende Grundlagen für weitere Arbeiten entwickelt: Das Konzeptlexikon als Teil eines Textverstehens- und Übersetzungsmodells mit der Framesprache K 0 L3 für die Formalisierung des Konzeptlexikons und seiner Konzepte; die Klassifikation lokaler Präpositionen des Deutschen, insbesondere im Hinblick auf die in der bisherigen Forschung eher vernachlässigten direktionalen und exogenen Präpositionen, die als sogenannte transformative Präpositionen zusammengefaßt werden; und eine für die Zuordnung einfacher Verben bzw. auch komplexer Ausdrücke anwendbare Sachverhaltsontologie.

Im Rahmen dieser Ergebnisse wird ein Ansatz für die Komposition lokaler Verben und Präpositionen vorgestellt. Hierbei spielen bei der Erfassung der Komposition prozeßhafter Verben und transformativer Präpositionen die Präpositionen gegenüber den Verben eine zentrale Rolle: Die Problematik der Umkategorisierung eines prozeßhaften Verbs zu einem ereignishaften Verb als Teil eines komplexen Ausdrucks mit transformativer Präposition wird dadurch gelöst, daß die transformative Präposition – und nicht das Verb – den Sachverhaltstyp des gesamten Ausdrucks bestimmt. Motiviert ist der Ansatz durch die Betrachtung deutschfranzösischer lexikalischer Kontraste, die darin bestehen, daß französische transformative Verben wie *entrer* und *sortir* in gewisser Weise deutschen transformativen Präpositionen wie *in* (mit Akkusativ) und *aus* entsprechen. Eine mögliche Verallgemeinerung der Arbeit besteht in der Übertragung des vorgeschlagenen Ansatzes auf andere Phänomene der Zeitkonstitution.

Bei Durchsicht aller im weiteren Sinne zum LDV-Forum gehörenden Unterlagen sind eine Reihe – durchaus nicht uninteressanter – Rezensionsexemplare zutage gefördert worden, die beim besten Willen nicht mehr als Neuerscheinungen gelten können. Sollte sich eine InteressentIn finden, die willens und in der Lage ist, eine Rezension für das LDV-Forum zu schreiben, so werde ich die entsprechenden Exemplare gerne weiterreichen. Ansonsten werde ich diese Bücher im Laufe des Septembers wieder an die Verlage zurückschicken.

Kucera, Antonin; Rey, Alain; Wiegand, Herbert Ernst; Zgusta, Ladislav (ed): Lexicographica. International Annual for Lexicography (Bd. 3); 293 S.; Tübingen: Niemeyer, 1987

Kucera, Antonin; Rey, Alain; Wiegand, Herbert Ernst; Zgusta, Ladislav (ed): Lexicographica. International Annual for Lexicography (Bd. 4); 278 S.; Tübingen: Niemeyer, 1987

Luckhardt, Heinz-Dirk: Sprache und Information. (Der Transfer in der maschinellen Sprachübersetzung; Bd. 18), 221 S., Tübingen: Niemeyer, 1987

Wilss, Wolfram; Schmitz, Klaus-Dirk (ed): Sprache und Information. (Maschinelle Übersetzung: Methoden und Werkzeuge. Akten des 3. Internationalen Kolloquiums des Sonderforschungsbereichs 100 Elektronische Sprachforschung; Bd. 16), 324 S., Tübingen: Niemeyer, 1987

Naumann, Sven: Generalisierte Phasenstrukturgrammatik: Parsingstrategien, Regelorganisation und Unifikation. 180 S., Tübingen: Niemeyer, 1988

Steiner, Erich; Schmidt, Paul; Zelinsky Wibbelt, Cornelia (ed): Prom Syntax to Semantics. Insights from Machine Translation. 262 S., London: Pinter Publishers, 1988

Zock, Michael; Sabah, Gerard (ed): Advances in Natural Language Generation. An Interdisciplinary Perspective. Vol. 1; 200 S.; London: Pinter Publishers, 1988.

Zock, Michael; Sabah, Gerard (ed): Advances in Natural Language Generation. An Interdisciplinary Perspective. Vol. 2; 176 S.; London: Pinter Publishers, 1988

Laffling, John: Towards High-Precision Machine Translation based on Contrastive Textology (Distributed Language Translation 7). 178 S., Berlin; New York: Walter de Gruyter, 1992