

## LDV-FORUM

Forum der Gesellschaft für Linguistische Datenverarbeitung GLDV

LDV-Forum 8.1-2 (1991) Forum der Gesellschaft für Linguistische Datenverarbeitung e.V.

### Herausgeber

Gesellschaft für Linguistische Datenverarbeitung e.V. (GLDV)

*Anschrift:* Prof. Dr. Burghard Rieger, Universität Trier, FB II: LDV /CL, D-5500 Trier, Postfach 3825, Tel.: (0651)201-2272/2270; FAX (0651)2013928; Email: rieger%utruert@unido.uucp.de oder unido!utruert!rieger

### Redaktion

Burghard Rieger, Roland Fraese, Amancio Kolompar

### Wissenschaftlicher Beirat

Karin Haenelt, Hans Haugeneder, Peter Hellwig, Gerhard Knorz, Jürgen Krause, Winfried Lenders, Dietmar Rösner

Erscheinungsweise Doppelnummer im Jahr 1991, ansonsten halbjährlich zum 31. März und 30. September

### Bezugsbedingungen

Für Mitglieder der GLDV ist der Bezugspreis des LDV-Forum im Jahresbeitrag mit eingeschlossen. Jahresabonnements können zum Preis von DM 50,- (incl. Versand), Einzellexemplare zum Preis von DM 20,- (zuzügl. Versandkosten) bei



## Editorial

Mit diesem nun endlich vorliegenden *LDV-Forum*, das auch in 1991 wiederum nur als Doppelnummer erscheinen konnte, ist hoffentlich die lang währende Durststrecke beendet, welche leider für die Herausgabe, Redaktion und Produktion der letzten Hefte des *LD V-Forum* kennzeichnend war. Nach der von Beirat und Vorstand vorgeschlagenen und auf der letzten Mitgliederversammlung beschlossenen Neuordnung sind zumindest die Rahmenbedingungen dafür geschaffen, daß nicht nur ein regelmäßiges Erscheinen sondern auch ein kontinuierlicher Ausbau des wissenschaftlichen Profils wie des informatorischen Niveaus des *LD V-Forums* in Zukunft gesichert werden kann.

1. Schon vor einem Jahr hatte Gerhard Knorz an gleicher Stelle beklagt, daß in der Vergangenheit - verbunden mit seiner *Ein-Personen Verantwortungs-Organisations-Herstellungs-Arbeit* am *LDV-Forum* ein regelmäßiges Erscheinen zweier Hefte im Jahr nur durch eine Art selbst-ausbeuterischem Arbeitsaufwand zu sichern waren, daß seine veränderten (äußeren und inneren) Lebensumstände aber einen solchen Einsatz nicht mehr erlaubten. Mit unserem allfälligen, hier wiederholten Dank an Gerhard Knorz, der das *LDV-Forum* seit 1985 nicht nur redaktionell betreut sondern ganz entscheidend auch in Form und heutigem Erscheinungsbild mitgeprägt hat, verbindet sich inzwischen wieder die (begründete) Hoffnung, ihn und seine langjährige Erfahrung auch weiterhin dem *LD V-Forum* erhalten zu können.

2. Daß die vorliegende Ausgabe des *LD V-Forum* wiederum noch eher den alten Gegebenheiten als den neuen Möglichkeiten entspricht, ist unter anderem in den veränderten politischen Bedingungen begründet, die einer Umsetzung unserer für 1991 anvisierten "Trierer Lösung" entgegenstanden: unerwarteter Ausfall erwarteter Ressourcen. Die vereinte Suche nach einem praktikablen Ausweg lief nur auf eine Zwischenlösung hinaus, die - mehr schlecht als recht - das alte Problem der *Ein-Personen-Verantwortungs-Organisations-Herstellungs-Arbeit* nunmehr von Darmstadt (Knorz) nach Trier (Rieger) verlagerte, anstatt es zu lösen. Dazu hätte es schon Ende 1990 etwa der Beschlüsse zur Neuordnung der Mitgliedsbeiträge bedurft, wie sie durch die Mitgliederversammlung der *GLD V* am 26. September 1991 in Trier gefaßt wurden, die nun erstmals auch für das *LD V-Forum* eine entsprechende Basis einer hoffentlich langfristigen Konsolidierung schaffen. So ist denn die verspätet erscheinende Doppelnummer 8.1-2 (1991) in manchem bestenfalls ein Abbild ihrer Vorgängerin: wiederum weitgehend von einem gemacht, dessen begrenzte Arbeitskraft und übrige Verpflichtungen bei zunehmender Abstinenz der freiwilligen Mitarbeit anderer am *LD V-Forum* ein früheres Erscheinen und ein neu es Erscheinungsbild bisher verhinderten.

3. Die von Vorstand und Beirat angestellten Überlegungen zur wissenschaftlichen, redaktionellen und ökonomischen Sicherung des Periodikums der *GLDV* als *wissenschaftliche Zeitschrift* gehen davon aus, daß das Spektrum des *LDV-Forums* inhaltlich nicht nur nicht eingengt, sondern - der zunehmenden Breite der Linguistischen Datenverarbeitung entsprechend - diese noch deutlicher als bisher in den Bereichen ihrer akademischen Ausbildung, theoretischen Forschung und Entwicklung, praktischen Erprobung und Anwendung, sowie ihren beruflichen Perspektiven durch thematische Profilierung einzelner Hefte akzentuiert werden sollte. Die entwickelten organisatorischen Vorstellungen sehen hierzu vor

- => einen Herausgeber (Editor) zu gewinnen, der seinen Namen mit dem wissenschaftlichen Ansehen und Niveau der Zeitschrift verbinden kann, weil er langfristig für Konzeption, Planung und redaktionelle Arbeit verantwortlich zeichnet,
- => ein Herausbergremium (Editorial Board) zu etablieren, dem u. a. der Vorsitzende oder Mitglieder des *Vorstands* der *GLDV* und die Mitglieder des *Wissenschaftlichen Beirats* angehören,
- => eine Redaktion einzurichten, die längerfristige Mitarbeit z.B. durch bezahlte Teilzeitkräfte sichert,
- => die Attraktivität des Publikationsortes *LDV-Forum* dadurch zu erhöhen, daß Fachbeiträge vor ihrer Veröffentlichung einer rigoroseren Begutachtung und Auswahl unterworfen werden,
- => die ökonomische Basis der Zeitschrift zu verbessern, was über Werbung von Vollmitgliedern, von Abonnenten und Inserenten geschehen kann.

4. Freilich sind derartige Vorschläge nur so gut wie die Personen, die sie erfolgreich umzusetzen vermögen. Davon gibt es in jedem Fall zu wenige, unter unseren Leserinnen und Lesern aber hoffentlich doch noch einige, die hierzu gewonnen werden können. Ihnen allen darf ich dabei zu bedenken geben, daß die Mitgliedschaft in einer wissenschaftlichen Gesellschaft wie der *GLDV* nicht einen durch Beitragszahlung abgegoltenen Kauf von Dienstleistungen darstellt sondern eine durch eben diese Mitgliedschaft erklärte Bereitschaft zur aktiven Mitarbeit. Sie betrifft sowohl die Unterstützung der satzungsgemäßen Ziele einer Gesellschaft, die durch Bündelung der Aktivitäten von Einzelnen der Förderung ihrer gemeinsamen (gleichen oder ähnlichen) Interessen dienen, als auch die Teilnahme am korporativ(n) Leben einer solchen Gesellschaft, das in den (aktiven und passiven) Wahlen, den Veranstaltungen und Publikationsaktivitäten besteht. Die Mitarbeit am und im *LDV-Forum* gibt hierzu beispielhaft umfassende Gelegenheit.

5. Das nächste Heft 9.1 (1992) des *LD V-Forum* wird pünktlich erscheinen. An Fachbeiträgen wird dieses Heft einige der (überarbeiteten) Vorträge enthalten, die für die beiden *GLDV-Sektionen "Unification Based Grammars and Models"* und "*Natural Language Processing and Tools*" im Rahmen der *QUAL/CO-91* angenommen worden waren. Weiterhin wird das Heft einen Tagungsbericht über die *First Quantitative Linguistics Conference (QUAL/CO-91)* vom 23.-27. September 1991 in Trier enthalten und - neben den bekannten Rubriken--einen ersten Beitrag einer neuen Folge, die über die Situation von *LDV & NL?* in anderen Ländern aufgrund von Informationsreisen aktuell berichten wird: zunächst aus Rußland. Schwerpunktthema .für Heft 9.2 (1992): Empirische Sprachdaten und *LDV /CL*.

Trier, im November 1991

der Redaktion bestellt werden. Eine Ausnahmeregelung gilt für diese Doppelnummer, die zum Preis von DM 30,- abgegeben wird.

#### **Titelgestaltung**

Werbestudio Zimmermann, D-6083 Biebesheim

#### **Fachbeiträge**

Fachbeiträge werden vor Veröffentlichung von mindestens zwei ReferentInnen begutachtet. Manuskripte sind daher möglichst frühzeitig an die Herausgeber zu schicken, neben Typoskript oder Ausdruck (dreifach) möglichst auch auf Diskette (5 1/4" bzw. 3 1/2") oder elektronisch (Email: `ldvforum%utru@unido.uucp.de` oder `unido!utru!ldvforum`) als ASCII oder LATEX-Datei (*LDVforum.sty* wird zugesandt).

#### **Rubriken**

Die namentlich gezeichneten Beiträge geben ausschließlich die Meinung der Autoren wider. Einreichungen sind -wie bei Fachbeiträgen - an die Redaktion zu übermitteln.

#### **Redaktionsschluß**

Für alle Rubriken mit Ausnahme der als Fachbeiträge eingereichten Manuskripte:  
für Heft 9.1/92: 28.02.1992; für Heft 9.2/92: 31.08.1992

#### **Herstellung**

Druckerei Print-In, Schönbornstr. 11, D-5500 Trier

#### **Auflage**

550 Exemplare

#### **Anzeigen**

Preisliste und Informationen: Prof. Dr. Johann Haller, Institut für Angewandte Informationsforschung (IAI), Martin-Luther-Str. 14, D6600 Saarbrücken 3; Tel.: (0681) 39313; FAX (0681) 397482; Email: `haller@iai.sbsvax.uucp.de`

#### **Bankverbindung GLDVforum**

(Prof. Rieger): Stadtparkasse Trier (BLZ 585 500 80) Kto. Nr. 680.280

B.R.

## Überprüfung der Lauttreue der deutschen Orthographie

Carl Ludwig Naumann (Aachen), Gerd Willee

### 1 Überblick

Mehrfach hat Naumann die Lauttreue der deutschen Orthographie qualitativ und quantitativ untersucht, zuletzt und am breitesten in Naumann1989a. Die empirische Basis war die von W. Ortmann vorgelegte Aufbereitung des Kaeding-Corpus (Kaeding1898) (v.a. Ortmann1976). Das Kaeding-Corpus ist zwar bislang an Umfang unübertroffen. (Vgl. Tab. 1, Spalte AnzWö = Anzahl der Wörter) Aber wegen des hohen Alters der Materialgrundlage (erschienen 1898) und seiner fragwürdigen Repräsentativität war eine Überprüfung an neueren Texten wünschenswert. Diese Überprüfung wurde an Material aus dem LIMAS-Corpus (Glas1974) vorgenommen. Hier sollen Grundlagen und Verfahren der Überprüfung und die Ergebnisse - höchste Übereinstimmung der interessierenden Zahlen mit dem Kaeding-Corpus vorgestellt werden. Vorberichte über Ziele und Verfahren haben wir gegeben in NaumannjWillee1988 und Willee1987.

Nur soweit erforderlich, unterscheiden wir im Text zwischen "Laut" und "Phonem" (vgl. Naumann1989a:238f.), und vermeiden bei hinreichender Klarheit auch eine Kennzeichnung der Beispiele mit Klammern verschiedener Art.

### 2 Methodologische Probleme I: Statistik

Das Ausgangsinteresse dieser Lauttreue-Untersuchungen ist ein didaktisches, denn die Frage der Lauttreue stellt schon immer einen zentralen Aspekt des Erst-Schreib- und -Lese-Unterrichts dar (Streit um die "Ganzheitsmethode" bis in die 70er Jahre) sowie des

Rechtschreibunterrichts (Grundwortschatz-Diskussion in den 80ern), wobei Reform-Interessen hier hintangestellt werden sollen.

In der Geschichte der Auseinandersetzungen fällt auf, daß grundsätzliche statistisch-methodologische Fragen ungeklärt blieben, was angesichts der Hitzigkeit dieser Kämpfe in Schule und Öffentlichkeit nicht bloß von akademischem Interesse ist:

#### . Corpus oder Lexikon als Zählbasis?

Daß hier ein Unterschied bestehen kann, ist z. T. von den Beteiligten an den Auseinandersetzungen gar nicht bemerkt worden. So beabsichtigt Karstädt1930, Mohr1891 zu widerlegen durch eine verbesserte Definition der Zählseinheiten und schreibt, er habe" ... ebenfalls ein Wörterbuch ausgezählt ... "(S. 236), während Mohr Texthäufigkeiten ermittelt hatte. In der vorliegenden Untersuchung war die Entscheidung gegen das Lexikon und für ein Corpus vorgeprägt von dem spezifischen didaktischen Interesse: Für den beginnenden Schreibern ist der Lernnutzen solcher Einheiten wichtig, die in seinen Texten vermutlich oft auftreten werden.

#### • Repräsentativität?

Sie ist ebenfalls bei fast allen Corpora kaum diskutiert; am klarsten ist sie gegeben beim Limas-Corpus. Allerdings ist man trotzdem in einer Klemme: Man braucht prinzipiell große Corpora; die gibt es aber (vgl. Tabelle 1) nur für geschriebene, genauer: gedruckte Texte. Nun ist eigentlich Lauttreue nicht zu definieren ohne Beachtung der Perspektive des anfänglichen Lernalters, sie wäre also eigentlich auch in Corpora mit gesprochener Kindersprache zu prüfen. Diese aber

sind, vgl. nochmals Tabelle 1, unerfreulich klein, was beim großen Aufwand der Verschriftung verständlich ist und sich vermutlich nicht allzu bald ändern wird.

. Phonem, Wortform, Lemma oder Morphem als Zählseinheiten?

. Was heißt genau 'Lauttreue' ?

Beide Fragen sind nicht zu klären ohne rechtsschreibtheoretische Überlegungen. Die Entwicklungen der letzten 11 Jahre - als deren Initiatoren hier die Nerius-Gruppe in Rostock/Berlin-DDR und G. Augst erwähnt seien - haben die Pulverschwaden aus alten Kampfzeiten etwas abziehen lassen; daher ist jetzt der Blick frei dafür, daß der Kampf um "ganzheitlich oder synthetisch" aus orthographietheoretischer Sicht eigentlich überflüssig war. Der folgende Abschnitt 3 - als Exkurs zu sehen - soll dies zeigen.

### 3 Methodologische Probleme 11: Orthographietheorie

Zunächst schien sich in der beginnenden linguistischen Durchdringung der Orthographie das Lauttreue-Problem sogar gänzlich überflüssig zu machen: Die Begeisterung über die Entdeckung, daß die alte Rede von "Prinzipien der Schrift" linguistisch neu gefaßt und somit befriedigender fundiert werden kann, daß also neben den Phonemen auch Morpheme und Syntagmen, in geringerem Grade (und partiell umstritten, vgl. zuletzt Naumann1990a und Naumann1990b) Lexeme, Silben und prosodische sowie textuelle Aspekte in die Schreibregeln eingehen, schien die Frage nach der Lauttreue obsolet zu machen. Wenn nämlich so viele Prinzipien wirken und man sich ihr Verhältnis untereinander als ungeordnet glaubt vorstellen zu sollen, ist Lauttreue ein nutzloses Konstrukt.

Man muß aber ein regelhaftes Ineinandergreifen der Prinzipien annehmen, insbesondere eine enge Verbindung zwischen Phonemen und Morphemen als Aspekten von Regeln:

. Die sg. Auslautverhärtung ist z.B. eine strenge Regel in der deutschen Orthographie. Eine bestimmte morphologisch definierte Form muß lautlich analysiert werden: Infinitiv bzw. 1. Person Plural Praesens bei den Verbformen, obliquer

Kasus oder Plural bei den Nomina und Adjektiven - eigentlich aber jede Form mit Vokal nach den kritischen Stammfinalen Konsonanten [p, t, k, f, s; auch l.:]. Also *Bund* unter Bezug auf *Bundes*, *Gleis* <- *Gleise*, *halb* <- *halbe*.

. Ähnlich, wenn auch nur im Flexionsbereich mit gleicher Strenge, nicht bei Wortbildung oder Etymologie, funktioniert die Regel der <ä>-Schreibung: Lauttreues [a] und in einem weiteren Sinne lauttreues [au/ao] in der bestimmten morphologisch definierten Form (bei Nomina hier jedoch: Nominativ Singular) "begründen" eine <ä>- bzw. <äu>-Schreibung. Also *Wände* <- *Wand*, *läuft* <- *laufen*.

Einen Terminus von Maas benutzend, kann man in beiden Fällen von lauttreuen "Stützformen" sprechen, also von ausgezeichneten Allomorphen, die die lautliche Bezugsbasis für die Schreibungen darstellen.

. Lautnachbarschaftsregeln sind z. T. auf die Morphemgrenze zu beziehen. Z.B. gelten die Regeln: "[s] vor [t, p] nicht als <sch>, sondern als <s>; [n] vor [k] nicht als <ng>, sondern als <n>; [kv] als <qu>" nur innerhalb von Morphemgrenzen. Man kann das leicht an folgenden schönen Fehlern verdeutlichen: *Fisterrine*, *Löspapier*; *Lankornreis*; *Trinquasser*. Wenn also das phonematische und das morphologische Prinzip so eng ineinandergreifen, ist "Lauttreue" keineswegs ein überflüssiger Begriff, sondern kann sinnvoll gefaßt und fruchtbar gemacht werden. Denn sie kann in die Weiterentwicklung der Prinzipien-Idee eingebracht werden: Eine geordnete und quantifizierte Prinzipien-Integration erlaubt es, sich einem Orthographie-System im umfassenden Sinne zu nähern.

(Vgl. Naumann1989a, bes. Kap. 2.6.)

## 4 Lösungsansatz

Die in Abschnitt 2. gestellte Frage nach den Zählseinheiten ist damit soweit zu beantworten: Phoneme und Morpheme müssen als Zählbasis berücksichtigt werden. Zu prüfen sind allerdings noch Wortformen und Lemmata.

Die Idee, Wortformen zu zählen, wäre aus computerlinguistischer Sicht attraktiv, weil Wortformen am bequemsten

verarbeitet werden können. Ihre sachliche Angemessenheit steht aber gar nicht erst zur Debatte. Orthographiesystematisch sind sie nämlich nur relevant in einem schwachen Punkt, der Reduktion von einigen konsonantischen Geminaten und Trigeminaten in Flexion bzw. Komposition (*du lachst*, aber *du bremsst* [mit nur einem <s>, obwohl der Stamm mit s endet und die Endung mit s beginnt]; *Sauerstoffflasche*, aber *Schiffahrt*.).

Mit der Entscheidung für Morpheme scheint zugleich die Wahl auf Lemmata mit einem Lemmanamen gefallen zu sein. Dies ist jedoch in einem wichtigen Bereich problematisch, nämlich bei den orthographisch allomorphen Verben. Die Häufigkeit der Fehler *\*bekamm*, *\*kamm* statt *bekam*, *kam* und *\*erschrack* statt *erschrak* deutet auf ein besonderes Zusammenwirken von phonematischen und morphologischen Regelaspekten: Die zweifelsfreie semantische Ähnlichkeit bzw. Identität der Teillemmata *komm* und *kam*, *schreck* und *schrak* ist orthographisch unerheblich gegenüber ihrer lautlichen Verschiedenheit. Dies ist um so bedeutender, wenn man die sonst so starke graphische Morphemkonstanz im Deutschen beachtet, die das <h> sogar durch solche Stammformenreihen hindurch erhält, wo die Längenschreibung <ie> für langes [i] auftritt; vgl. *sehen*, *sieht*. - Zumindest diese "orthographisch verkrachten Verwandten", also die orthographisch allomorphen Verben, können nicht als Lemmata unter nur einem Lemmanamen geführt werden.

Bei der hier entwickelten Auffassung verschiebt sich die Frage nach der Lauttreue gegenüber früher in einer doppelten Weise. Erstens müßte man genauer von einer Laut-Orientierung sprechen; dann kann man daneben eine Morphem- und eine Syntaxorientierung stellen (um die drei wichtigsten und didaktisch relevanten linguistischen Bezugsebenen zu nennen). Und zweitens ist es so nicht nötig, die Frage als absolute Entscheidungsfrage zu stellen: "Ist die deutsche Orthographie laut treu oder nicht?". Aber als Frage nach einer Skalierung stellt sie sich sinnvoll neu: "Wie lautreu ist die deutsche Orthographie?", zumal dann das grundsätzlich festgestellte Ineinandergreifen von Phonematik und Morphologie präziser erfaßt wer-

den kann.

Eine dritte, unterrichtspraktisch auch interessante Auffassung von Lauttreue wird hier nicht verfolgt. (Vgl. aber Meyer-Schepers 1990; Dummer-Smoch/Hackethal 1988) Sie versuchen, Lernwörter nach abnehmender Lauttreue anzuordnen, um so Schriftanfängern und Legasthenikern einen optimalen (Neu)Einstieg zu ermöglichen.

Lauttreue ist im Rahmen dieser Untersuchung also aufgefaßt als Grad der Einfachheit, mit dem gesprochene Sprache in geschriebene übertragen wird. Aber reine Index-Werte sind offensichtlich zu grob, sie müssen bei bestimmten Abweichungen durch Bedingungen, z.B. morphologische, ergänzt werden.

Der Grad der Lauttreue wird in dieser Untersuchung bestimmt durch Auszählung der Verschriftungsmöglichkeiten von (Phonemen und einigen Phonemverbindungen), so wie sie sich in einem Corpus darstellen, und zwar unter definierten morphologischen Bedingungen.

## 5 Verfahren

Wir haben die Verschriftungsregeln für deutsche Phoneme in den strukturell deutschen Morphemen des LIMAS-Corpus untersucht. Die Ergebnisse stellen wir den erwähnten Kaeding-Ortmann-Daten gegenüber. Dafür gingen wir in folgenden Untersuchungsschritten vor:

Wie schon bei der Arbeit am Kaeding-Material wurden die Korrespondenzen zwischen Lauten und Schriftzeichen auf der Morphemebene untersucht. Die aus beiden Corpora extrahierten Listen von Wortformen sind vergleichbar: Ortmann bezieht sich bei dem Kaeding-Corpus (Umfang ca. 10 Mio. lfd. Wortformen) auf alle Wortformen-Types, die häufiger als 100 mal im Corpus repräsentiert sind, und erhält dadurch eine Liste von 7995 Types, die insgesamt ca. 87% aller Wortformen bei Kaeding abdecken. Das LIMAS-Corpus enthält ca. 1,2 Mio. laufende Wortformen, alle Wortformen mit einer Frequenz  $\geq 10$  wurden herangezogen. Diese 7668 Wortformen-Types decken ca. 83% aller Textwortformen des LIMAS-Corpus ab (vgl. auch Tab. 1.). Beide Listen stimmen zu

etwa 65% in ihrem Wort bestand überein; der Wandel im Lexem-Bestand im Laufe von über 70 Jahren ist also offensichtlich - unabhängig von der bei Kaeding nicht sehr systematischen Vorgehensweise beim Sammeln der Texte für sein Corpus.

## Unterschiede in der Art der Bearbeitung:

Das Kaeding-Material wurde weitgehend manuell bearbeitet (ausgehend von der maschinellen Aufbereitung durch Ortmann), das LIMAS-Material soweit wie möglich maschinell, wobei der Computer grundsätzlich als Datenspeicher eingesetzt wurde, vorhandene aufbereitete Sprachdaten herangezogen wurden und z. T. bereits bestehende Sprachverarbeitungsprogramme Anwendung fanden. Wo eine maschinelle Bearbeitung zu aufwendig oder ganz unmöglich war, wurden intellektuelle Bearbeitungsschritte durchgeführt (Vgl. auch Willee 1987).

Das LIMAS-Material wurde in zwei Schritten phonemisiert, und zwar zunächst mittels der Liste der phonemisierten Kaeding- Wortformen von Ortmann, soweit es sich um Wortformen handelte, die in beiden Listen enthalten waren. Die dadurch noch nicht erfaßten Wortformen wurden durch das Programm PHONOL bearbeitet (vgl. Breuer /Brustkern/Willee 1979). Schließlich wurde alles noch einmal manuell korrigiert.

Die Zergliederung in Morphe bzw. morphartige Gebilde erfolgte auf der Phonem-Ebene manuell, auf der Graphemebene mithilfe des Mackensen-Lexikons (vgl. Bünting 1969) maschinell mit manueller Ergänzung bzw. Korrektur. Letztere war unumgänglich, da bei dem Mackensen-Lexikon eine zum Teil von unserem Ansatz abweichende Definition von Morph/Morphem zugrundelag.

(Eine Zergliederung in morphartige Gebilde, wie etwa Bei/spiel, ist in unserem Zusammenhang durchaus sinnvoll, weil zwar nicht eine strikte morphologische Segmentierung, dafür aber eine orthographische Behandlung wie bei "normalen" Morphemen möglich ist. Für die Orthographie, zumindest unter didaktischem Gesichtspunkt,

ist keine historisch-etymologische oder strikt morphologische Kompetenz zugrunde zu legen; man kann weithin mit der Unterstellung einer synchron-etymologischen, morphologisch teils unterdeteminierten Kompetenz arbeiten.)

Es wurden ca. 18000 Wortsegmente (Tokens) isoliert, die ca. 2000 Types entsprechen.

Diese verschiedenen Paare wurden halb-automatisch weiter segmentiert in Paare von Phonemen und zugehörigen Graphemen (resp. jeweiligen Clustern); ca. 8840 dieser Paare (Tokens) konnten zu 94 Types reduziert werden.

## 6 Ergebnisse

Die statistische Auswertung der Phoneme, zuzüglich der Cluster /ks, kv, ts/, und ihrer Schreibungen im Deutschen ergibt eine weitgehende Übereinstimmung der Häufigkeit von Regeln nach LIMAS- und Kaeding-Corpus; vgl. Tabelle 2 sowie 3. (Auch die Daten von Thome 1989 stimmen, wenngleich schwächer, mit beiden Untersuchungen überein. Allerdings ist Thomes Corpus kleiner: 10.000 Phoneme, so daß wir seine Abweichungen für weniger bedeutsam halten.)

Die Übereinstimmung zwischen Verschriftungsregeln im Kaeding- und im LIMAS-Material bestätigt die Vermutung, daß auch offenkundige Verschiebungen im Lexikonbestand (der gleichgebliebene Wortbestand macht ca. 65% beider Listen aus) nicht den Phonem- und insoweit den orthographischen Bestand verändert haben.

## 7 Divergenzen

Einige Zahlenverschiebungen - vom Gesamtcorpus her gesehen marginal - sind bei den Fremdwörtern sowie den Schreibungen des /i:/ zu beobachten (Vgl. Tab. 3). Man könnte prüfen, ob sie faktisch zu erklären sind (im Kaeding-Material fallen veraltete französische Fremdwörter auf), ob sie aus Spezifika der Corpora herrühren (Protokolle von Reichtagsdebatten, daher vielleicht im Kaeding-Material öfter <ph> aus dem bekanntlich dort häufigen "Paragraph"; mehr pronominale

Verkettungen, daher häufigere <ih>-Morpheme?) oder ob sie methodologische Artefakte darstellen (unterschiedliche Handhabung der Fremdwörter in Zweifelsfällen). Wenn die Auftretenshäufigkeit von /ks/ nicht sowieso gering wäre, wäre eine Aufklärung der Divergenz zwischen den beiden Hauptfällen in den bei den Corpora interessant.

Einige Buchungsfehler in der Kaeding-Ortmann-Auswertung wurden beim Zahlenvergleich bemerkt, die aber zu eher marginalen Verschiebungen führen: <jeg> hat auslautend [k], <ver> und <zer> sind bei <t> zu buchen, <zur> hat bei der hier zu unterstellenden sorgfältigen Sprechweise (vgl. Naumann 1989a, Kap.2 und 3) langes [u], nicht kurzes.

Selbst die genannten Zahlenunterschiede sind aber ohne grundsätzliche Bedeutung; es kommt hier auf markante Unterschiede an wie die geringe Häufigkeit des <h> als Vokallängezeichen oder die Größenordnung der Fremdwort- und Sonderschreibungen.<sup>1</sup> Didaktisch erheblich ist dann auch, daß wir für die Verschriftungsregeln Corpus-Häufigkeiten gezählt haben; zusätzlich sollten Aspekte der Lexikon-Häufigkeit beachtet werden. Denn wenn neben einem Verschriftungshauptfall hohe Abweichungen auftreten (v statt f, s statt sch.; auch i oder ih statt ie), ist die lexikalische Belegung zu prüfen. beginitemize

- . <s> statt <sch> unterliegt einer strengen Kontextregel;
- . <ih> statt <ie> kann als sehr kurze Liste von Morphemen formuliert werden, ähnlich <v> statt <f> als eine immerhin überschaubare Liste.

Im Gegensatz zu solchen didaktisch einfach zu nutzenden Fällen ist allerdings <i> statt <ie> nur als lange Liste darstellbar.

<sup>1</sup> Vgl. die Darstellungen der Verteilungen für Verschriftungshäufigkeiten im LIMAS-Corpus als Torten-Diagramme (Abb. 1 - 6, entsprechend den Tabellen 2 und 3) sowie die tabellarische Gesamtdarstellung der Werte beim LIMAS-Corpus in Tab. 4 Wld5.

## 8 Didaktische Folgerungen

Die so demonstrierte relativ starke Lautorientierung der deutschen Orthographie macht ihre Nutzung beim Schriffterwerb sinnvoll, wenn sie auch kein hegemonialer Ansatz sein darf.

Auch echte Ausnahmen lassen sich hier gut markieren: Listen mit wenigen Morphemen bieten sich als wortweise zu lernende Abweichungen vom Normalfall an (Der Normalfall muß aber als solcher bekannt gemacht werden!). Das gilt für Doppelvokalschreibungen, <ieh>, vielleicht auch <dt> und einige Fremdwortbesonderheiten. Das gilt ebenso für <v> statt <f> und <ih> statt <ie>2).

Die bislang auch in didaktischen Werken oft irreführend dargestellte Mengenverteilung bei den Langvokalschreibungen mit <h> erzeugt vermutlich eine Häufung entsprechender Fehler. (Menze11985: 15% mehr fälschlich gesetzte als fälschlich fehlende <h>!). Hier ist nicht der Platz, die interessanten orthographietheoretischen und didaktischen Ursachen dieses Irrtums aufzurollen. Hinzuweisen ist aber darauf, daß in diesem Phänomen der rationale linguistische Kern der Ganzheitsmethodik zu sehen ist: Wenn man eine solche Mengenverteilung schon als Regel auffaßt, so wäre es doch eine für den Lerner nutzlose Regel, weil ohne brauchbare Vorhersagesicherheit. (Der sonst in der Rechtschreibdidaktik so viel mißbrauchte alberne Spruch, daß die Ausnahme die Regel bestätige, hat hier einen Platz.) 'Der Fehlschluß der Ganzheitler, man dürfe nur ganze Wörter lernen, folgt also aus der Übertreibung einer richtigen Beobachtung.

U. E. ist die relativ große Menge von Morphemen, die in diesem Bereich zu lernen sind (die mit <h>!), eines der drei orthographischen Gebiete, die wegen Schwierigkeit der "Regel" ein Auslernen-Wollen der Orthographie in der Grundschulzeit als unsinnig erweisen. (Daneben die andersartig schwierigen Bereiche Großschreibung und Regeln für <s 58 - ß>.)

Die Ergebnisse unserer Untersuchung - sowohl was den diachronen Vergleich angeht als auch was die quantitative und qualitative Analyse der Verschriftungsphänomene betrifft, sollten ihren Niederschlag finden in der Orthographie-Didaktik. Schließlich läßt sich an den Beispielen ablesen, daß offenbar nur zu oft 'Schwierigkeiten' in der Orthographie geübt werden, nur weil sie so schön systematisch

übbar sind - unabhängig von ihrer tatsächlichen Relevanz im Alltagsgeschäft des Schreibens. Unsere Untersuchung bestätigt an aktuellem Material, daß didaktische Relevanz auch auf statistische Daten gestützt sein muß. Der Rechtschreibstoff - und zwar nicht nur die in der anfänglichen Grundwortschatz-Diskussion favorisierten "Wörter" oder die inzwischen deutlicher hervortretenden Morpheme (Augst1989), sondern eben auch die Phonem-Verschriftungen (Naumann 1989a und Naumann1989b) - zeigt so klare und damit lern-nützliche Trends, daß es eine üble Verschwendung wäre, sie zu mißachten.

## Literatur

- Augst1984a** Augst, Gerhard, Hrsg. (1984a) Kinderwort. Der aktive Kinderwortschatz (kurz vor der Einschulung) Frankfurt/M., Bern, New York: Lang = Theorie und Vermittlung der Sprache 1.
- Augst1989** Augst, Gerhard (1989) Schriftwortschatz. Untersuchungen und Wortlisten zum orthographischen Lexikon bei Schülern und Erwachsenen. = Theorie und Vermittlung der Sprache Bd. 10. Frankfurt/Bern/New York/Paris: Lang
- Augst1985** Augst, Gerhard (Hrsg.) (1985) Graphematik und Orthographie. Frankfurt
- BalhornF** = Fehlercorpus aus Balhorn/Vieluf/Vieluf1984
- Balhorn/Vieluf/Vieluf1984** Balhorn, Heiko, Jochen Vieluf, Ulrich Vieluf (1984) Bericht: Erhebung des Grundwortschatzes in Schulaufsätzen. Hamburg: (masch. verf. ).
- Breuer/Brustkern/Willee1979** Breuer, Brustkern, Willee (1979) PHON-TEXT - eine PHONOL-Anwendung zur Erzeugung synthetischer deutscher Sprache. In: SDv 1,2/1979
- Bünting1969** Bünting, Karl Dieter (1969) Morphologische Strukturen deutscher Wörter, Bonn
- Dummer-Smoch/Hackethal1988** Lisa Dummer-Smoch, Renate Hackethal (1988) Handbuch zum Kieler Leseaufbau. Kiel: Veris 2. Aufl.
- Glas1974** Glas, Reinhold (1974) Ein Textkorpus für die deutsche Gegenwartssprache, in: G.Nickel/H.Raasch (Hrsg.), IRAL-Sonderband, Heidelberg
- Hesse/Wagner1985** Hesse, Harlinde, Klaus R. Wagner (1985) Der Grundwortschatz der Primarstufe. Wortlisten und Anleitungen für den Aufbau eines klassenbezogenen Grundwortschatzes. Dorsten: Spectra - Multi Media.
- Kaeding1898** Kaeding, Friedrich Wilhelm, Hrsg. (1898) Häufigkeitswörterbuch der deutschen Sprache - festgestellt durch einen Arbeitsausschuß der deutschen Stenographiesysteme. Steglitz bei Berlin.
- Karstädt1930** Karstädt, Otto (1930) Das Diktat als schriftliche Darstellung. In: Methodische Strömungen der Gegenwart, hrsg. von O. K., 18. Auflage. Berlin und Leipzig: Julius Beltz Verlag, S. 235 - 242.
- LIMAS** = LIMAS-Corpus und abgeleitetes Material (Vgl. Glas1974.)
- Maas1989** Utz Maas (1989) Grundzüge der deutschen Orthographie. Osnabrück: Universität/FB Sprache - Literatur 2. Aufl.
- Mahlstedt1985a** Mahlstedt, Dagmar (1985 a) Grundwortschatz und kindliche Schriftsprache. In: Diskussion Deutsch 81, S. 89 - 105.
- Mahlstedt1985b** Mahlstedt, Dagmar (1985 b) Listen zum Grundwortschatz (masch. verf. )
- Mannheim I** = Mannheimer Corpus I und daraus abgeleitetes Material.
- MenzelF** = Fehlercorpus zu Menze11985, S. 14 - 19.
- Menze11985** Menzel, Wolfgang (1985) Rechtschreibunterricht. Praxis und Theorie. = Beiheft Praxis Deutsch 69.
- Meyer-Schepers1990** Ursula Meyer-Schepers (1990) Linguistik und Problematik des Schriftspracherwerbs. Bochum (masch. verf.)
- Mohr1891** Mohr, J.(1891) Unsere Methode der Rechtschreibung. Kritik derselben und Vorschläge zu ihrer Umgestaltung. Flensburg: Aug. Westphalen.
- Naumann1985** Naumann, Carl Ludwig (1985) Grundwortschatz-orientiertes Rechtschreiblernen. In: Augst1985, S. 260 - 270.

- Naumann1987** Naumann, Carl Ludwig (1987) Ordnung oder Willkür? Beiträge zu einem systematischen Rechtschreibunterricht. In: Die Grundschule 11, S. 14 - 18.
- Naumann1989a** Naumann, Carl Ludwig (1989a) Gesprochenes Deutsch und Orthographie. Linguistische und didaktische Studien zur Rolle der gesprochenen Sprache in System und Erwerb der Rechtschreibung. Frankfurt, Bern, etc.: Peter Lang. = Theorie und Vermittlung der Sprache, hrsg. von Gerhard Augst und Rudolf Beier, Band 8.
- Naumann1989b** Naumann, Carl Ludwig (1989b) Plädoyer für die Arbeit mit Grundwortschätzen. In: Ingrid M. Naegele/Renate Valtin: LRS in den Klassen 1 - 10. Handbuch der Lese - und Rechtschreibschwierigkeiten. Weinheim und Basel: Beltz, S. 181 - 185.
- Naumann1990a** Naumann, Carl Ludwig (1990a) 'Rechtschreibwörter und Rechtschreibregelungen', Handreichung zum Lehrplan Grundschule Sprache. Mit einem Beitrag von Ingrid Niedersteberg. Redaktion: Frank Schindler, Elke Diekmann, unter Verwendung von Material von Siegfried Asmuth. Soest 1986, 2. Aufl. 1987, 3., überarbeitete Aufl.
- Naumann1990b** Naumann, Carl Ludwig (1990b) Nochmals zu den Prinzipien der Orthographie. In: Christian Stetter (Hrsg.): Zu einer Theorie der Orthographie. Interdisziplinäre Aspekte gegenwärtiger Schrift- und Orthographieforschung. Tübingen: Niemeyer, S. 145 - 162.
- Naumann/Schindler1988** Naumann, Carl Ludwig, Schindler, Frank (1988) Wörter für einen Rechtschreibgrundwortschatz. In: H.H. Munske u. a. (Hrsg.) Deutscher Wortschatz. Lexikologische Studien. FS L.E. Schmitt, Berlin. S. 827 - 847.
- Naumann/Willée1988** Naumann, Carl Ludwig, Willée, Gerd (1988) Regelmäßigkeiten in der Beziehung zwischen gesprochener und geschriebener deutscher Sprache. In P. G. Roper (Hrsg.): Computers in Literary and Linguistic Research. Proceedings of the Thirteenth International Conference, Norwich 1986. Paris/Genf. S. 147 - 151.
- Nerius1987** Nerius und Autorenkollektiv (1987) Deutsche Orthographie. Leipzig: VEB Bibliographisches Institut.
- Oberender1951** Oberender, W. (1951) Der Häufigkeitswortschatz einer 3.Klasse. In: Die Schul warte 4, 349 - 357.
- Ortmann1976** Ortmann, Wolf Dieter (1976) Beispielwörter für deutsche Rechtschreibübungen. 7995 hochfrequente Wortformen der Kaeding-Zählung, rechner-sortiert nach Phonem-Graphem-Beziehungen, hrsg. vom Goethe-Institut München, Projekt Phonotheek. - Mit Ergänzungen aus Computer-Listen zu nicht abgedruckten Beispielen (Frequenz über 100).
- Pfeffer1964** Pfeffer, J. Alan (1964) Grunddeutsch. Basic (Spoken) German Word List. Grundstufe. Englewood Cliffs, N. J.
- Pregel/Rickheit** Pregel, Dietrich, Gert Rickheit (- im Erscheinen -) Wörterlisten aus dem Braunschweiger Corpus.
- Rosengren1972** Rosengren, Inger (1972) Ein Frequenzwörterbuch der deutschen Zeitungssprache. Die Welt, Süddeutsche Zeitung 1, CWK Glerup Lund Schweden.
- Ruoff1981** Ruoff, Arno (1981) Häufigkeitswörterbuch gesprochener Sprache. Tübingen.
- Schaffert** = mitabgedrucktes Material einer Untersuchung aus dem Jahre 1934, bei Oberender1951
- Schlag1921** Schlag, Johannes (1921) Häufigkeitsproben aus dem Sprachschätze von sechs- und achtjährigen Kindern. In: pädagogisch-psychologische Arbeiten aus dem Institut des Leipziger Lehrervereins, Bd. XI. Leipzig.
- Thomé1989** Thome, Günther (1989) Rechtschreibfehler und Orthographie. In: Der Deutschunterricht 6/1989, S. 29 - 38
- Wängler1963** Wängler, Hans Heinrich (1963) Rangwörterbuch hochdeutscher Umgangssprache, Marburg.
- Wagner1984** Wagner, Klaus R. (1984) Computerlisten zu 3 Kindern aus dem Dortmunder Corpus. V gl. Hesse/Wagner1985.

**Willée1987** Willée, Gerd (1987) Maschinengestützte Ermittlung und Wichtung von Verschriftungsregeln im Deutschen. Vortrag gehalten auf dem XIV. Internationalen Linguistik-Kongreß in Berlin-DDR 1987. Erscheint in den Proceedings

**Tabelle 1: Orthographierelevante empirische Wortschatzzählungen**

nach: Naumann (1989a) S. 286f.

NAME <sup>1)</sup>	JAHR	MODUS	ALTER	A <sub>INF</sub>	A <sub>WÖ</sub>	EINH.	GRUPPE
Schlag <sub>1</sub>	1921	mü	6 J.	2	27	Wo	1. Kinder, mündlich
Wagner	1975 (1985)	mü	6,8,9 J.	3	45	Wo	
Augst <sub>mü</sub>	1984	mü	6 J.	10	24(Lex.)	Gr	
Pre/Ri <sub>mü</sub>	1987	mü	1.- 4. Kl.	800	190	Gr <sup>2)</sup>	
Schlag <sub>2</sub>	1921	mü + schr	2.Kl.	1 Kl.	18	Wo	2. Kinder, schriftlich
Schaffert	(1934)	schr	3.Kl.	1 Kl.	17	Gr	
Oberender	1951	schr	3.Kl.	1 Kl.	22	Gr	
Pre/Ri <sub>schr</sub>	1987	schr	3.+ 4.Kl.	200	68	Gr <sup>2)</sup>	
Balhorn u.a	1984	schr	4.Kl.	190	114	Wo+Gr	
Mahlstedt	1985	schr	1.- 4.Kl.	33 Kl.	54	Wo+Gr	
Augst <sub>schr</sub>	1989	schr	4., 10.Kl.	20	284	Gr <sup>3)</sup>	3. Fehler
Balhorn <sub>F</sub>	1984	Fehler	4.Kl.	190	9	Wo+Gr	
Menzel <sub>F</sub>	1985	Fehler	2.-10.Kl.	950	20	Wo+Gr	4. Erwachsene, mündlich
Wängler <sub>mü</sub>	1963	mü	Jgd.+Erw.	-	80	Wo	
Pfeffer	1964	mü	Jgd.+Erw.	450	595	Gr	
Ruoff	1981	mü	Erw.	343	500	Gr	
Freiburg	(1974) <sup>4)</sup>	mü	Jgd.+Erw.	-	477	Wo	5. Erwachsene, Druck
Kaeding	1898	schr	Erw.	-	10908	Wo	
Wängler <sub>schr</sub>	1963	schr	Erw.	-	80	Wo	
Mannheim I	(1970)	schr	Erw.	-	2145	Wo	
Mannheim II	(1973)	schr	Erw.	-	331	Wo	
Limas	(1971)	schr	Erw.	-	1000	Wo	
Řosengren	1972	schr	Erw.	594	2977	Wo+Gr	
Bonner Ztg.1	(1974)	schr	Erw.	-	3040	Wo	

Abkürzungen:

A<sub>INF</sub> = Anzahl der Informanten

A<sub>WÖ</sub> = Textlänge in tausend Wörtern,  
Lex = Lexikonumfang

EINH. = Gezählte Einheiten :

Wo = Wortformen, Gr = Lemmata

JAHR in ( ) = nicht erschienen

Fußnoten:

1) Indices können aus der Spalte  
MODUS erschlossen werden

2) Bei Pregel/Rickheit sind nur Nomina,  
Verben und Adjektive gezählt

3) Auch morphemisiert dargeboten

4) Später teilpubliziert

**Tabelle 2: Ausgewählte Verschriftungshäufigkeiten  
Kaeding - Limas im Vergleich: Konsonanten**

Die Zahlenpaare für Kaeding

und Limas sind übereinander geschrieben.

Für Vergleichszwecke sind die Spalten wie in Naumann (1989a) Seite 88f. gezählt.

1. Häufigk. im Corpus	2. Haupt- schreibung	3. Dopp- lung	4. Auslaut- verhärtung	5. Auslaut- verweichung	6. andere deutsche Schreibung	7. Umgeb.- bedingung	8. Fremde Schreibungen
2,37 2,73	f	59	ff 4 3		v 36 37		ph 1,4 0
7,63 - 7,11 7,11	t	77 79	tt 4 2	d 19 19	dt 0,3 0,3		th 0,2 0
4,58 3,43	s /s/	76 79	ss 7 6		B 16 15		
1,09 0,96	sch/ş/	55 52			s 44 48	zw. Wortan- fang und p + t	ch 0,2 0,1
1,30 1,37	k	69 69	ck 11 11	g 18 20			ch 2 0,1
1,97 1,93	ch /ç/	92 91		g 8 9		nach i	
0,02 0,03	chs/ks/	63 73					x 37 27

### Tabelle 3: Ausgewählte Verschriftungshäufigkeiten Kaeding - Limas im Vergleich: Langvokale und Diphthonge

Die Zahlenpaare für Kaeding

und Limas sind übereinander geschrieben

Für Vergleichszwecke sind die Spalten wie in Naumann (1989a) Seite 92ff. gezählt.

1. Alle Vokale	4. Langvokale			
	a. Sonderfälle	b. ohne Kennzeichnung	c. mit <h>	d. Dopplung
	ieh	i	ih	ie
2,48	1	9	12	78
2,64	1	6	8	85
(4.) alle anderen Langvokale zusammen				
7,04	vor Mehrfach- konsonanz	vor einfachem Konsonanten	mit <h>	Dopplung
7,30	Beisp.: Erde, Mond	Wesen, Name	nehmen, Wahl	See, Moor
	0,02 (fehlt)	86 88	11 12	0,006 0,6
	ai	ei	eih	
2,38	0	99	1	
2,67	3	99	0,02	

Tabelle 4: Liste der verschiedenen Phonem-Graphem-Paare  
(sortiert nach Frequenzen) \*)

absol. Frequenz	relat. Frequenz	Beispiel- Segment	Phonem (-Cluster)	Graphem (-Cluster)
354690	10.2791	SCHEIN	N	N
325945	9.4460	SCHEITER	ð	E
273998	7.9406	SCHARF	R	R
212384	6.1550	SCHEID	D	D
193410	5.6051	SCHAFT	T	T
117318	3.3904	SCHICHT	I	I
106112	3.0666	SCHAFF	A	A
96578	2.7911	SCHER	E1	E
96143	2.7862	SCHALT	L	L
93701	2.7155	STETS	S	S
92599	2.6835	SCHIMPANS	M	M
91757	2.6591	SCHEIB	AE	EI
77281	2.2396	SCHIEB	I1	IE
75786	2.1963	SCHIMPANS	Z	S
74007	2.1447	SCHLAG	G	G
69821	2.0234	SCHWANK	V	W
68594	1.9878	SCHEIB	B	B
67741	1.9577	SCHLUCK	U	U
63628	1.8388	SCHENK	E	E
60687	1.7587	SCHICHT	Ç	CH
55784	1.6166	SCHAFT	F	F
45300	1.3128	SCHEID	T	D
44507	1.2898	STAB	A1	A
42946	1.2411	SCHLOSS	O	O
40407	1.1710	SCHMELZ	T S	Z
35289	1.0226	VATER	F	V
34433	0.9978	SCHAU	AO	AU
32383	0.9384	SCHENK	K	K
32031	0.9282	GEHEIM	H	H
30666	0.8862	BIOGRAPH	I3	I
29923	0.8671	SCHLANG	η	NG
28166	0.8162	SCHLUG	U1	U
25193	0.7301	SCHLACHT	X	CH
25100	0.7253	SCHOB	O1	O
24432	0.7080	SCHAFF	f	SCH
22746	0.6591	SPALT	f	S
22294	0.6442	AKTUELL	U3	U
19064	0.5509	SPEZIF	E3	E
17535	0.5081	SCHIEsz	S	sz
16298	0.4723	BRÄNN	N	NN
16092	0.4663	SPÜR	Y1	Ü
13931	0.4037	SCHNELL	L	LL
12720	0.3686	SCHIMPANS	P	P
11511	0.3335	SCHÈH	E1	EH
10632	0.3072	SCHÄFT	E	A
10010	0.2900	SCHOB	P	B
9406	0.2725	SCHLAG	K	G
8393	0.2432	JANUAR	J	J
7609	0.2205	SCHLÜSS	Y	Ü

\*) Die Kodierung der Phoneme folgt im Wesentlichen derjenigen von Ortmann1976, S. 26-28, V1 f.; der Murrel-vokal wird als ð dargestellt, der velare Nasal als η, /sch/ als f und ß als sz.

Tabelle 4: Liste der verschiedenen Phonem-Cluster (Sortiert nach Frequenz)

Phonem-Cluster (-Cluster)	Frequenz	Beispiel	Segment	Phonem	Cluster
7360	0.2132	STAHL	A1	AH	
7206	0.2088	SCHLOSS	S	SS	
7134	0.2067	STAMM	M	MM	
7128	0.2065	IHM	I1	IH	
7048	0.2042	SCHÄTZ	T S	TZ	
6958	0.2016	SCHLEUNIG	OQ	EU	
5924	0.1716	SCHATT	T	TT	
5859	0.1697	SCHLEUNIG	C	G	
5619	0.1629	MIR	I1	I	
5618	0.1628	SCHÖN	Q1	Ö	
5365	0.1554	SCHÄD	E2	Ä	
5048	0.1462	SCHICK	K	CK	
5042	0.1461	SCHÖPF	Q	O	
4165	0.1203	SCHOKOLADE	O3	O	
3148	0.0912	SCHAFF	F	FF	
3108	0.0900	STOHL	O1	OH	
2957	0.0856	BLÜH	Y1	UH	
2940	0.0852	SCHENK	η	N	
2778	0.0805	STRÄHN	E2	ÄH	
2417	0.0699	APPARAT	A3	A	
1303	0.0377	STRAUCH	OQ	AU	
1189	0.0344	SCHIEH	I1	IEH	
980	0.0284	FROH	Q1	OH	
880	0.0255	APPARAT	P1	PP	
824	0.0238	STAAT	A1	AA	
781	0.0226	STADT	T	DT	
708	0.0205	SPERR	R	RR	
704	0.0204	ACHS	K S	CHS	
619	0.0179	STUHL	U1	UH	
597	0.0173	SCHNEE	E1	EE	
347	0.0100	VIELLEICHT	I3	IE	
325	0.0094	CHRIST	K	CH	
271	0.0078	LEIH	AE	EIH	
254	0.0073	MARXISMUS	K S	X	
234	0.0067	INDIVIDU	V	V	
198	0.0057	KAISER	AE	AI	
165	0.0047	BEQUEM	K V	QU	
104	0.0030	INSTITUTION	T S	T	
57	0.0016	ANALYS	Y1	Y	
40	0.0012	PRÄSIDENT	E3	A	
36	0.0010	CHEF	f	CH	
34	0.0009	BIOGRAPH	F	PH	
34	0.0009	BIOGRAPH	F	PH	
32	0.0008	ANALYS	Y	Y	
31	0.0008	BURO	Y3	U	
26	0.0007	ARITHMET	T	TH	
23	0.0006	FLAGG	G	GG	
19	0.0005	BOOT	O1	OO	
	Y	SCHLUSS	0.2502	5009	

**Tabelle 5: Liste der verschiedenen Phonem-Graphem-Paare  
(sortiert nach Phonemen und Frequenz) \*)**

absol. Frequenz	relat. Frequenz	Beispiel- Segment	Phonem (-Cluster)	Graphem (-Cluster)
29923	0.8671	SCHLANG	ŋ	NG
2940	0.0852	SCHENK	ŋ	N
24432	0.7080	SCHAFF	f	SCH
22746	0.6591	SPALT	f	S
36	0.0010	CHEF	f	CH
325945	9.4460	SCHAITER	ð	E
106112	3.0666	SCHAFF	A	A
91757	2.6591	SCHIB	AE	EI
198	0.0057	KAISER	AE	AI
271	0.0078	LEIH	AE	EIH
34433	0.9978	SCHAU	AO	AU
44507	1.2898	STAB	A1	A
7360	0.2132	STAHL	A1	AH
824	0.0238	STAAT	A1	AA
2417	0.0699	APPARAT	A3	A
68594	1.9878	SCHIB	B	B
60687	1.7587	SCHICHT	Ç	CH
5859	0.1697	SCHLEUNIG	Ç	G
212384	6.1550	SCHID	D	D
63628	1.8388	SCHENK	E	E
10632	0.3072	SCHÄFT	E	Ä
96578	2.7911	SCHER	E1	E
11511	0.3335	SCHIH	E1	EH
597	0.0173	SCHNEE	E1	EE
5365	0.1554	SCHÄD	E2	Ä
2778	0.0805	STRÄHN	E2	ÄH
40	0.0012	PRÄSIDENT	E3	Ä
19064	0.5509	SPEZIF	E3	E

\*) Die Kodierung der Phoneme folgt im Wesentlichen derjenigen von Ortmann 1976, S. 26-28, V1 f.; der Murmelvokal wird als ð dargestellt, der velare Nasal als ŋ, /sch/ als f und ß als sz.

55784	1.6166	SCHAFT	F	F
35289	1.0226	VATER	F	V
3148	0.0912	SCHAFF	F	FF
34	0.0009	BIOGRAPH	F	PH
23	0.0006	FLAGG	G	GG
32031	0.9282	GEHEIM	H	H
117318	3.3904	SCHICHT	I	I
77281	2.2396	SCHIEB	I1	IE
7128	0.2065	IHM	I1	IH
5619	0.1629	MIR	I1	I
1189	0.0344	SCHIEH	I1	IEH
30666	0.8862	BIOGRAPH	I3	I
347	0.0100	VIELLEICHT	I3	IE
8393	0.2432	JANUAR	J	J
32383	0.9384	SCHENK	K	K
9406	0.2725	SCHLAG	K	G
5048	0.1462	SCHICK	K	CK
325	0.0094	CHRIST	K	CH
704	0.0204	ACHS	K S	CHS
254	0.0073	MARXISMUS	K S	X
165	0.0047	BEQUEM	K V	QU
96143	2.7862	SCHALT	L	L
13931	0.4037	SCHNELL	L	LL
92599	2.6835	SCHIMPANS	M	M
7134	0.2067	STAMM	M	MM
354690	10.2791	SCHEIN	N	N
16298	0.4723	BRANN	N	NN
42946	1.2411	SCHLOSS	O	O
6958	0.2016	SCHLEUNIG	OQ	EU
1303	0.0377	STRÄUCH	OQ	ÄU
25100	0.7253	SCHOB	O1	O
3108	0.0900	STOHL	O1	OH
19	0.0005	BOOT	O1	OO
4165	0.1203	SCHOKOLADE	O3	O

34	0.0009	BIOGRAPH	F	PH
12720	0.3686	SCHIMPANS	P	P
10010	0.2900	SCHOB	P	B
880	0.0255	APPARAT	P	PP
5042	0.1461	SCHÖPF	Q	Ö
5618	0.1628	SCHÖN	Q1	Ö
980	0.0284	FRÖH	Q1	ÖH
273998	7.9406	SCHARF	R	R
708	0.0205	SPERR	R	RR
93701	2.7155	STETS	S	S
17535	0.5081	SCHIEsz	S	sz
7206	0.2088	SCHLOSS	S	SS
193410	5.6051	SCHAFT	T	T
45300	1.3128	SCHEID	T	D
5924	0.1716	SCHATT	T	TT
781	0.0226	STADT	T	DT
26	0.0007	ARITHMET	T	TH
40407	1.1710	SCHMELZ	T S	Z
7048	0.2042	SCHATZ	T S	TZ
104	0.0030	INSTITUTION	T S	T
67741	1.9577	SCHLUCK	U	U
28166	0.8162	SCHLUG	U1	U
619	0.0179	STUHL	U1	UH
22294	0.6442	AKTUELL	U3	U
69821	2.0234	SCHWANK	V	W
234	0.0067	INDIVIDU	V	V
25193	0.7301	SCHLACHT	X	CH
7609	0.2205	SCHLÜSS	Y	Ü
32	0.0008	ANALYS	Y	Y
16092	0.4663	SPÜR	Y1	Ü
2957	0.0856	BLÜH	Y1	ÜH
57	0.0016	ANALYS	Y1	Y
31	0.0008	BÜRO	Y3	Ü
75786	2.1963	SCHIMPANS	Z	S

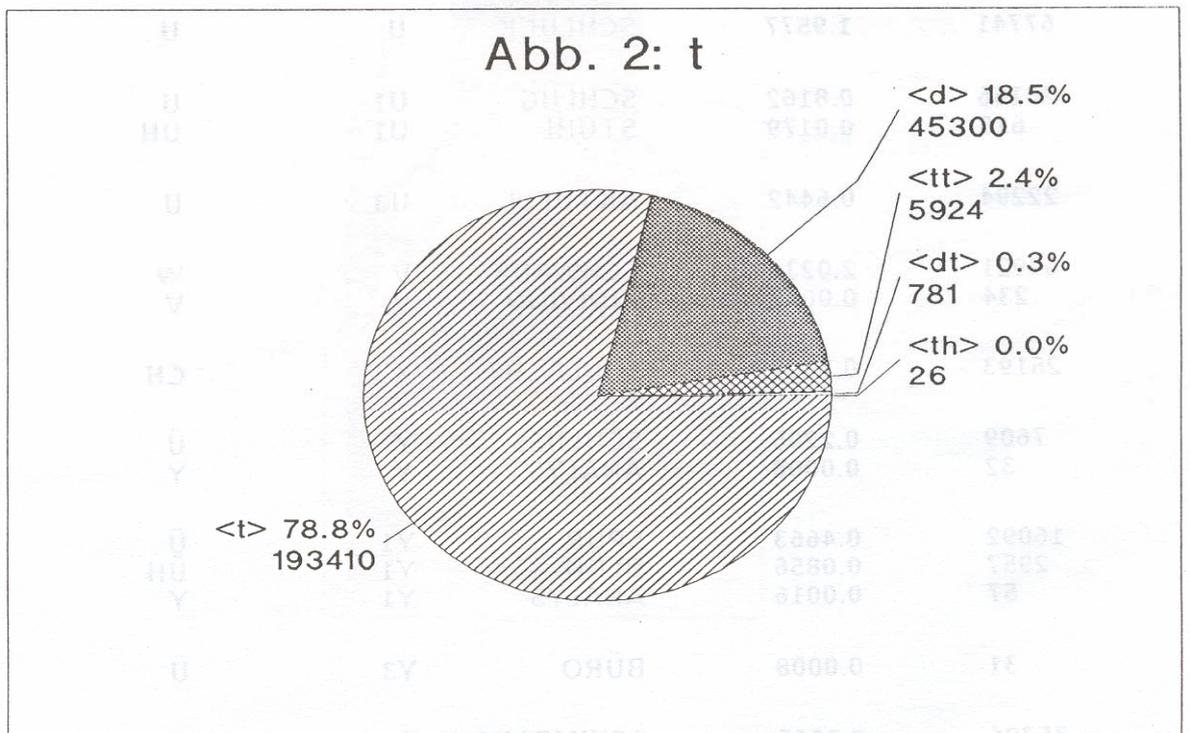
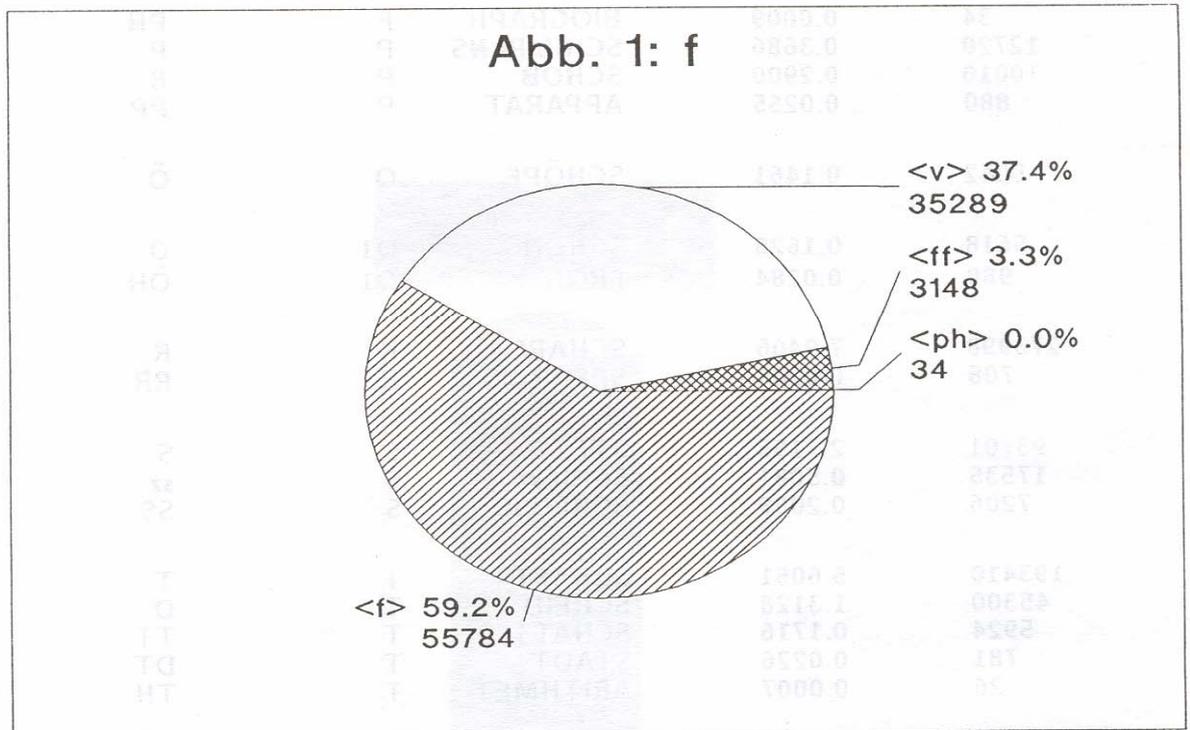


Abb. 3: s

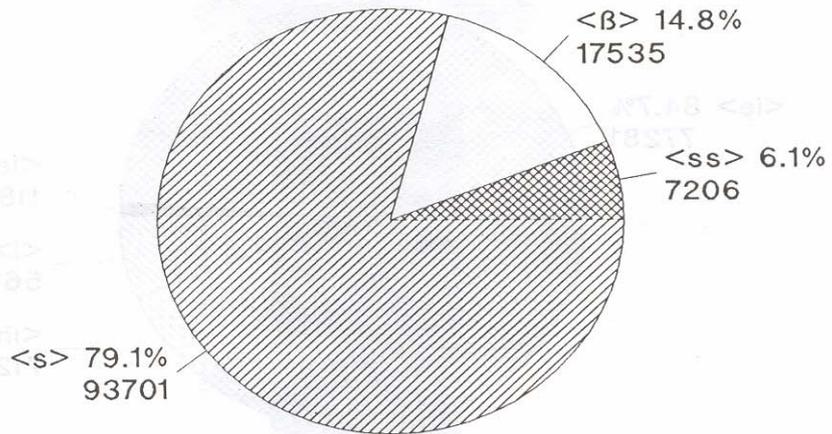


Abb. 4: sch

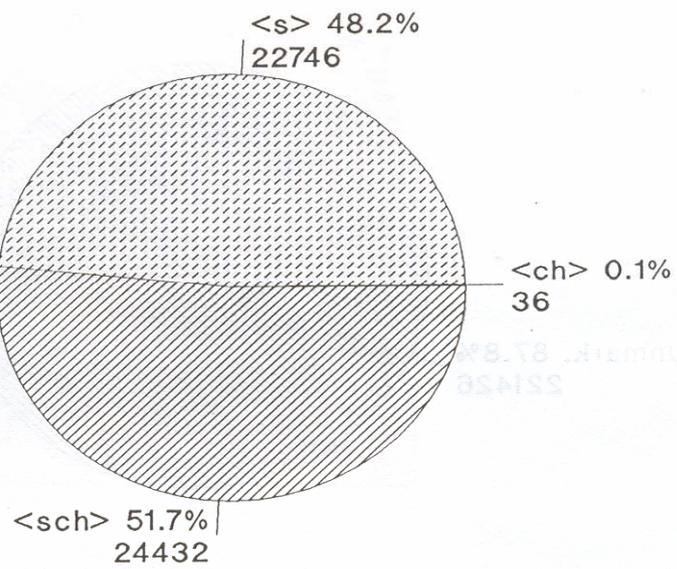


Abb. 5: ie

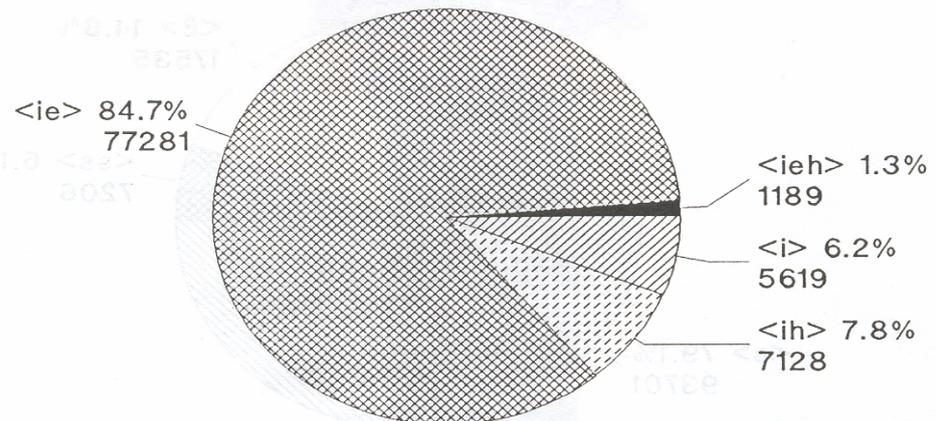
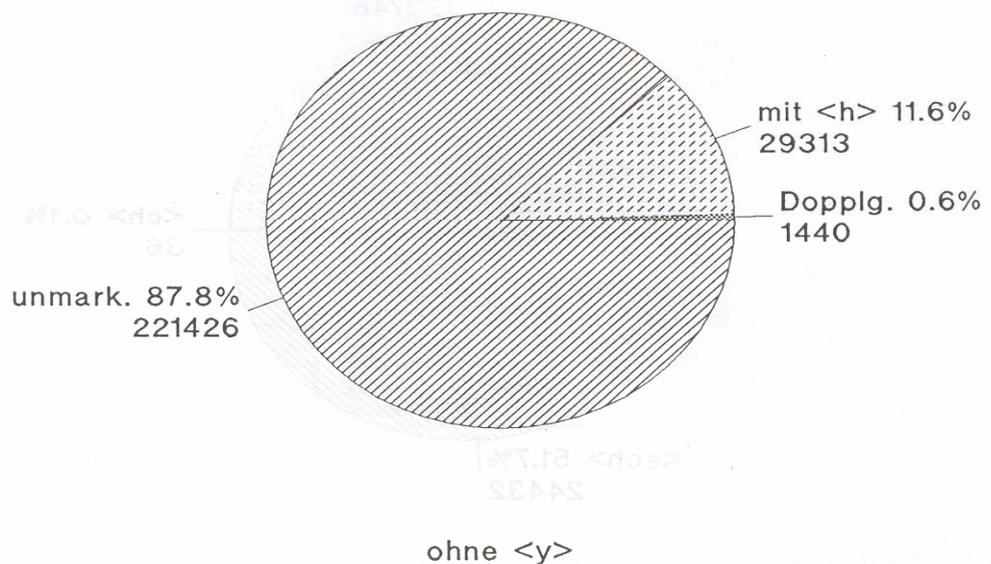


Abb. 6: andere Langvokale



## Ein topikbasiertes Retrievalsystem in Hypertext

Heinz J. Weber, Constantin Thiopoulos  
e-mail: weber@utrurt.uucp.de

### 1 Einleitung

Retrievalsysteme, die auf booleschem Retrieval basieren, sind - im Hinblick auf die exponentiell wachsenden Mengen von gespeicherten Informationen - nicht effizient, da sie nur in der Lage sind, durch eine boolesche Kombination von keywords eine Menge von gespeicherten Dokumenten zu bestimmen. Der Benutzer muß dann aus diesen Dokumenten die Stellen, die ihn interessieren, manuell herausfinden. Außerdem muß er die "richtige" Abfrage stellen.

Hier wird dagegen ein Retrievalsystem vorgestellt, das:

- Die gespeicherten Dokumente automatisch in Topiks aufteilt, eine entsprechende Topikhierarchie erstellt und diese als Hypertextstruktur generiert. Dadurch ist es möglich, daß der Benutzer nur einen Topik eines Dokumentes zu lesen hat, in dem ein keyword vorkommt. Falls er daraus schließt, daß auch andere Teile des Dokumentes für ihn von Interesse sind, kann er durch Aktivieren der entsprechenden Hypertextverbindungen zu diesen gelangen, sonst geht er zu dem entsprechenden Topik des nächsten Dokuments.
- Das Retrieval genauer macht durch das Vorschalten eines Kollokationfilters von Lexemen, die für den Benutzer interessant sind.
- Mit Hilfe eines Hypertextsystems das Retrieval auf das Aktivieren von graphischen Repräsentationen und Navigieren zurückführt, so daß der Benutzer die für ihn relevante Information bestimmen kann, ohne eine Abfrage explizit formulieren zu müssen.

### 2 Topikhierarchien

Texte werden oft nicht als isolierte Informationseinheiten verarbeitet, sondern können eine große Vielfalt zwischentextlicher Beziehungen aufweisen, die den Rezipienten in die Lage versetzen, durch Überschreiten der

Einzeltext-Grenzen Informationen zu sammeln und zu integrieren. Dies trifft u. a. auf wissenschaftliche Artikel zu, insbesondere jedoch auf Nachrichten. Hier besteht bei der Sammlung und Integration von Informationen die Möglichkeit, sowohl Texte der selben Tagesausgabe als auch Texte von zeitlich verschiedenen Ausgaben miteinander zu verbinden. Generell hängt es vom jeweiligen Leser ab, welche Texte in welcher Weise miteinander verbunden werden. Aber auch in den Texten selbst können Markierungen und Wegweiser vorkommen, die zur Orientierung dienen. Häufig verbinden diese Verweise Texte miteinander, in denen das selbe oder ein verwandtes Thema (im folgenden als "Topik" bezeichnet) behandelt wird. In Nachrichten dürfte dieser Typ von Text-Text-Relationen der wichtigste sein. Verbindungen zwischen Informationseinheiten führen zu einer Hypertext-Struktur, einer nicht-linearen Anordnung von Texten, die von der sequentiellen Präsentation (wie sie z.B. bei Nachrichten üblich ist) abweicht. [6]

Texte in nicht-linearer Anordnung werden von Hypertext-Systemen verwaltet. Die zwischen den Texten bestehenden logischen Verbindungen müssen entweder intellektuell/manuell hergestellt oder maschinell ermittelt werden. Bei maschineller Überführung von vorhandenen Texten in ein Hypertext-System muß eine Sprachverarbeitungs-komponente vorausgesetzt werden, d.h. ein Text-Parser und ein Link-Generator.

Der Beitrag stellt ein experimentelles Hypertext-System vor, in dem Einzeltexte des Typs Nachrichten nach thematischen Gesichtspunkten maschinell vernetzt werden sollen. Im folgenden soll eine Topik-Analyse skizziert werden, der jeder Einzeltext unterzogen wird. Ergebnis der Topik-Analyse ist die Zerlegung eines Textes in etikettierte Textsegmente und deren Anordnung in einer Topik-Hierarchie.

Für den Aufbau der Topik-Hierarchie soll das Vorgehen an einem kleinen Textsample demonstriert werden. Die drei Texte sind semi-authentisch: um Ablenkung zu vermeiden, werden fiktive Eigennamen verwendet.

*Beispiel 1*

T1, S1

*Unter der Leitung von H. Schulze traf eine Delegation der FCS-Fraktion in Gladz ein.*

T1, S2

*Dort begannen Gespräche mit Delegierten des Ostsylianischen Parlaments...*

T2, S1

*Eine Delegation der FCS-Fraktion besuchte das Ostsylianische Parlament in Gladz.*

T2, S2

*In einer Rede gab H. Schulze, Vorsitzender der FCS-Fraktion, der Hoffnung Ausdruck, daß die Grenze zwischen beiden sylianischen Staaten eine Grenze ohne Mauern und Stacheldraht werden wird.*

T3, S1

*Mitglieder der FCS-Fraktion besuchten das Ostsylianische Parlament in Gladz.*

T3, S2

*Es ist das erste offizielle Treffen von Parlamentariern beider sylianischer Staaten.*

T3, S3

*In der Zwischenzeit äußerte M. Meyer, Geschäftsführer der CFS, Zweifel bezüglich Kontakten dieser Art.*

T3, S4

*Der Besuch einer westsylianischen Delegation könnte Statusfragen berühren, die nicht zu leicht genommen werden sollten.*

Es wird nun versucht, jeden Text in kohäsive Segmente zu zerlegen; das geschieht entweder durch Zusammenfassung von Sätzen zu größeren Informationseinheiten oder durch Markierung von "Schnitten" zwischen Sätzen, die dann als Segmentgrenzen betrachtet werden. Grundlage der Zusammenfassung sind das Vorkommen von Kohäsionsmitteln wie Konnektive oder Anaphora am Satzanfang (z.B. "dort" in T1: S2), die Rekurrenz von Lexemen und Eigennamen, sowie Indikatoren für einbettende und eingebettete Strukturen (z.B. Sprechhandlungsverben in Verbindung mit dem Konjunktiv bei Redewiedergabe; vgl. T3: S3 & S4). "Rekurrenz" ist vorerst in einem engen Sinn definiert als Wiederauftreten des selben Lexems oder Namens

(z.B. "Fraktion" in T2: S1 & S2) oder als Vorkommen von Mitgliedern der selben Wortfamilie in aufeinanderfolgenden Sätzen (z.B. "Delegation" - "Delegierte" in T1: S1 & S2). Abwesenheit dieser Kohäsionsmittel führt zu einem "Schnitt" zwischen Sätzen. Für das Textsample ergibt sich daraus folgende Gliederung.

*Beispiel 2*

T1 Segment 1

S1: ... eine Delegation ... S2:

Dort... Delegierten...

T2 Segment 1

S1: ... der FCS-Fraktion ... S2:

... der FCS-Fraktion ...

T3 Segment 1

S1: ... das ... Parlament ... S2:

... Parlamentariern ... Segment

2

S3: H' äußerte ... Zweifel ... S4:

,.. könnte ... sollten.

Es ist auch möglich, daß ein Text nach diesem Analyseschritt aus einem einzigen Segment besteht oder daß er so viele Segmente hat wie Sätze. Im Durchschnitt weisen die Texte unseres Testcorpus (300 Nachrichtentexte) vier Segmente auf, ein Segment besteht aus durchschnittlich drei Sätzen. Hinsichtlich der Ermittlung von Textsegmenten ist klar, daß es eine Reihe weiterer Indikatoren gibt, sowohl für Kohäsion als auch für Diskontinuität (z.B. Vorkommen von Synonymen wie 'besuchen' - 'Treffen' - 'Kontakte' in T3 oder die Abfolge von unbestimmtem und bestimmtem Artikel).[9] Jedoch beim derzeitigen Stand des Experiments sollen nur eine geringe Anzahl robuster und verlässlicher Informationen herangezogen werden.

Die skizzierte Textzerlegung liefert eine Text-Interpunktion, die als Grundlage für die weiteren Schritte dient, die Binnen-Analyse der Textsegmente und die Konstruktion einer Topik-Hierarchie.

Die Binnen--Analyse resultiert in einer Repräsentation hypothetischer Topiks und einer Etikettierung der Segmente, die zum Aufbau einer Topik-Hierarchie führt. Folgende Segment-Etiketten werden unterschieden: Haupttopik, Subtopik, Redewiedergabe, Annotation. Obwohl sie aufgrund verhältnismäßig einfacher struktureller

Kriterien ermittelt werden können (was nicht heißt, daß diese Aufgabe trivial ist, denn die entsprechenden Segmente können sich über mehrere Sätze erstrecken), lassen sich die Segment-Etiketten auf einer abstrakteren Beschreibungsebene interpretieren, z.B. als Indikatoren für spezifische Textualitätsmodi (wie Narration, Argumentation oder Deskription) oder für bestimmte Text-Instanzen (wie Autor oder Figur). Haupt- und Subtopik können als Vertextungen im Modus "Narration" angesehen und der primären Text-Instanz "Autor" zugeordnet werden. Narrationsindikatoren für Nachrichten sind Präteritum und dritte Person bei den Verben. Die Unterscheidung zwischen Haupt- und Subtopik beruht auf mehreren Faktoren. Ein Subtopik

- ist durch Adverbiale oder Konjunktionen eingeleitet (vgl. "in der Zwischenzeit" in T3),
- kann im Tempus wechseln (z.B. von Imperfekt zu Plusquamperfekt),
- führt neue Referenten ein (vgl. 'Meyer', 'CFS' in T3).
- steht meist in nicht-initialer Textposition,

Für eine Entscheidung zugunsten der Etikettierung als Subtopik genügen zur Zeit zwei der vier Faktoren. Die Bezeichnungen Haupt- und Subtopik sind nicht mißzuverstehen als Abstufungen hinsichtlich der inhaltlichen Relevanz eines Textsegments, sie bezeichnen nur die Position in einem syntaktischen Gefüge - analog zur Satz-Hierarchie "Hauptsatz" und "Nebensatz", die ja auch nicht als Gewichtung von Information anzusehen ist.

Für T3 aus dem Textsample ergibt sich folgende Darstellung:

### Beispiel 3

T1 Segment 1  
Haupttopik I: Mitglieder der FCS-Fraktion besuchten das Ostsylvanische Parlament in Gladz. ...

#### Segment 2

Subtopik I: In der Zwischenzeit äußerte M. Meyer, Geschäftsführer der CFS, Zweifel bezüglich Kontakten dieser Art. ...

Annotationen können als Vertextungen im Modus "Kommentierung" interpretiert werden. Indikator ist der Tempuswechsel vom Präteritum ins Präsens. Die entsprechende Text-Instanz für Kommentierung ist der "Autor".

### Beispiel 4

#### T3 Segment 1

Haupttopik I: Mitglieder der FCS-Fraktion besuchten das Ostsylvanische Parlament in Gladz.

Annotation I: Es ist das erste offizielle Treffen von Parlamentariern beider sylvanischer Staaten.

Redewiedergabe ist einer untergeordneten, nicht-autonomen Textinstanz zuzuordnen (z.B. Sachverhaltsbeteiligte) und ist im Modus "Zitierung" vertextet. Zitierungsindikatoren sind Ausdrücke für Sprechhandlungen (z.B. "...äußerte... Zweifel") in Verbindung mit Konjunktionen, Konjunktiv oder Anführungszeichen. Segmente, die als Redewiedergabe etikettiert werden, sind wiederum als "Text" zu behandeln und in "Haupttopik", "Subtopik" etc. zu unterteilen, wobei die Möglichkeit von Rekursion via "Redewiedergabe" besteht. [10] [16]

### Beispiel 5

#### T3 Segment 1

Haupttopik I: Mitglieder der FCS-Fraktion besuchten das Ostsylvanische Parlament in Gladz.

Annotation I: Es ist das erste offizielle Treffen von Parlamentariern beider sylvanischer Staaten.

#### Segment 2

Subtopik I: In der Zwischenzeit äußerte M. Meyer, Geschäftsführer der CFS, Zweifel bezüglich Kontakten dieser Art.

Redewiederg.: Haupttopik: Der Besuch einer westsylvanischen Delegation könnte Statusfragen berühren, die nicht zu leicht genommen werden sollten.

Die Identifikation und Isolation von Redewiedergabe in Texten ist ein sehr wichtiger Punkt beim Parsen narrativer Texttypen, wie Nachricht oder Bericht. Ganz abgesehen von dem Umstand, daß die schemafundierte (sogenannte wissensbasierte) Text-Analyse hier auf Schwierigkeiten stoßen muß, wenn beim "Auffüllen" von Schemaleerstellen die Grenze zwischen Sachverhalten innerhalb

und außerhalb von Redewiedergaben überschritten werden, auch im Kontext von Information Retrieval, Wissensextraktion [2] und maschineller Übersetzung [15] ist die Unterscheidung zwischen "originären" und "zitierten" Textpassagen zu beachten.

Im Anschluß an Textzerlegung, Binnenanalyse und Etikettierung der Segmente wird eine Repräsentation der in den etikettierten Textteilen residenten hypothetischen Topiks vorgenommen. Das Format ist zur Zeit noch primitiv: Es handelt sich um ein "Topik-Profil", eine Liste des jeweils vorkommenden Vokabulars - reduziert um die Strukturwörter. So ergibt sich für das Segment Subtopik (Redewiedergabe (Haupttopik)) in T3 folgendes Topik-Profil.

### Beispiel 6

T3

Subtopik ...

Redewiedergabe:

**Haupttopik:** Besuch - westsylvanisch - Delegation - Statusfragen - berühren - leicht - nehmen

Natürlich kann ein solches "Topik-Profil" nicht als Repräsentation der Bedeutung eines Textteils angesehen werden wie eine "Makroproposition" etwa. Es handelt sich lediglich um die Oberfläche eines Konstrukts genannt "Topik", das Ergebnis der bisher genannten Analyseschritte ist. In Verbindung mit Text-Retrieval befindet sich diese Ebene der Topik-Beschreibung in einer Mittellage zwischen syntax-fundierten und schema-fundierten Strategien. Sie geht hinaus über die Betrachtung der Zusammengehörigkeit von Termen im Rahmen eines Syntagmas (z.B. Nominalgruppen, Präpositionalkonstruktionen) und sie verlangt nicht die Einbeziehung des ganzen Text-Vokabulars. Auch einige charakteristische Schwierigkeiten schema-fundierter Topik-Analysen lassen sich damit vermeiden. Diese treten dann zutage, wenn das Textmaterial den konzeptionellen Vorgaben nicht entspricht: Texte sind nicht immer stereotypen Inhalts und in schlichter Manier vertextet (wie die berühmten Schankschen Restaurant-Texte, um nur ein Beispiel zu nennen). Das gilt auch für Nachrichten. Im Hinblick auf die Nichtvorhersagbarkeit von Nachrichtenthemen und die vielfältigen Möglichkeiten der Vertextung ist es ratsamer, sich nicht allein auf vorfabrizierte Topik-Schemata zu verlassen, sondern genuin textliche Strukturelemente vorrangig zu verwenden.

Die Topik-Analyse schließt ab mit dem Aufbau eines Textbaums, einer Topik-Hierarchie. Zugrunde liegen dabei folgende Struktur-Regeln:

(1)  $TEXT \rightarrow$   
 $Haupttopik(\&Subtopik*)(\&Annotation*)$

(2)  $Haupt - /Subtopik \rightarrow$   
 $Topik(\&Annotation)(\&TEXT)$

Nachfolgender Textbaum repräsentiert die Topik-Hierarchie in T3.

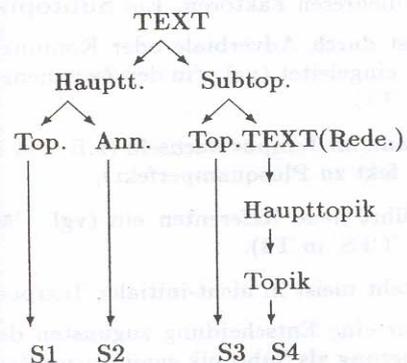


Abbildung 1: Topikhierarchie T3

### 3 Der Kollokationsgenerator

Durch die Generierung eines gewichteten Netzes, das die gegenseitige Abhängigkeit von Lexemen aufgrund von deren Gebrauch in den Texten repräsentiert, hat man ein Filter zu den Dokumenten zur Verfügung, das erlaubt:

- Das Retrieval von Dokumenten - aufgrund der subjektiven Anforderungen des jeweiligen Benutzers - mit einem höheren Genauigkeitsgrad zu bestimmen, als das beim booleschen Retrieval der Fall ist.
- Die Retrievalsuche durch Anklicken einer graphischen Repräsentation zu aktivieren, was den Vorteil hat, daß der Benutzer die für ihn relevanten Dokumente identifizieren kann, auch wenn er nicht in der Lage ist, eine entsprechende Abfrage zu formulieren.

Sei  $LEX$  eine Menge von Lexemen, die vom Benutzer eingegeben wird. Für  $A, B \in LEX$

und  $\mathcal{O}(A)$  Menge der Texte, in denen eine Instantiierung von  $A$  vorkommt, kann man eine Relation, welche die intertextuelle gegenseitige kollokationelle Abhängigkeit zwischen  $A$  und  $B$  modelliert wie folgt definieren:

$$\text{conf}(A, B) = \frac{|\mathcal{O}(A) \cap \mathcal{O}(B)|}{|\mathcal{O}(B)|}$$

Falls in jedem Text, in dem eine Instantiierung von  $A$  vorkommt, auch eine solche von  $B$  vorkommt, dann wird  $\text{conf}(A, B) = 1$  und falls  $A$  und  $B$  kein gemeinsames Vorkommen besitzen  $0^1$ . Dadurch wird eine

$$|\text{LEX}| \times |\text{LEX}| \text{ Matrix CONF}$$

berechnet, welche als Eintragungen die entsprechenden  $\text{conf}$  Werte hat.  $\text{conf}(L_i, L_j)$  ist das  $\langle i, j \rangle$ -te Element dieser Matrix und gibt die Gewichtung der Abhängigkeit  $f_{ij} : L_i \rightarrow L_j$  an.

Wenn die Dokumentenbasis durch neue Dokumente augmentiert wird, kann die CONF Matrix Neuberechnet werden, oder aber - falls die "Geschichte" der Dokumentenbasis mitberücksichtigt werden soll - kann  $\text{conf}$  zu einer inkrementellen Gewichtsbestimmung erweitert werden. Für die Neuberechnung von  $\text{conf}(A, B)$  gibt es nach der Aufnahme eines weiteren Textes drei Fälle:

1.  $A$  und  $B$  sind im neuen Text erhalten.

In diesem Fall muß sowohl die Anzahl der gemeinsamen Vorkommen als auch diejenige der Vorkommen von  $B$  um 1 erhöht werden, d.h.  $\text{conf}(A, B)$  wird

$$\text{conf}(A, B)_{\text{neu}} = \frac{|\mathcal{O}(A) \cap \mathcal{O}(B)| + 1}{|\mathcal{O}(B)| + 1}$$

Es gilt nun:

$$\text{conf}(A, B)_{\text{neu}} = \frac{\frac{|\mathcal{O}(A) \cap \mathcal{O}(B)| + 1}{|\mathcal{O}(B)| + 1}}{\frac{\text{conf}_{\text{alt}} + \frac{1}{|\mathcal{O}(B)|}}{1 + \frac{1}{|\mathcal{O}(B)|}}} =$$

2. Nur  $B$  kommt im neuen Text vor.

In diesem Fall wird nur  $|\mathcal{O}(B)|$  um 1 erhöht. Analog zum ersten Fall haben wir:

$$\text{conf}(A, B)_{\text{neu}} = \frac{\text{conf}_{\text{alt}}}{1 + \frac{1}{|\mathcal{O}(B)|}}$$

3. Nur  $A$  kommt im neuen Text vor.

In diesem Fall gilt:

$$\text{conf}(A, B)_{\text{neu}} = \text{conf}(A, B)_{\text{alt}}$$

<sup>1</sup> $\text{conf}$  ist analog zu Carnaps regulären Bestätigungsfunktion (Carnap, 1950) - hier in einem "Gebrauchsuniversum"- $c$  zwischen Aussagen  $e$  und  $h$ .

Um die jeweils aktuelle CONF Matrix zu bestimmen reicht es also wenn man die vorherige CONF Matrix und die Anzahl der Vorkommen der einzelnen Lexeme gespeichert hat.

Aus der CONF Matrix kann man nun die gewichteten Kanten eines gerichteten Graphen wie folgt bestimmen:

$$\text{conf}(A, B) \geq \text{conf}(B, A) \Rightarrow f : A \rightarrow B$$

$$\text{conf}(A, B) < \text{conf}(B, A) \Rightarrow f : B \rightarrow A$$

Dieser Graph kann als eine Hypertextstruktur aufgebaut werden, indem:

- jedes Lexem als eine Karte<sup>2</sup> implementiert wird und
- die Kanten, welche dieses Lexem mit anderen verbinden, als buttons implementiert werden, deren Name aus dem Ziellexem und der entsprechenden Gewichtung zusammengesetzt wird und die zu der Karte dieses Ziellexems führen.

Der Benutzer kann nun durch Anklicken der entsprechenden buttons durch das Netzwerk navigieren.

Um eine graphische Repräsentation des Netzwerks zu bekommen, wird ein *Hyperviewgenerator* eingesetzt, der aus der Hypertextstruktur eine *hyperview* erzeugt<sup>3</sup>. An diese *hyperview* werden die Dokumente angehängt, so daß man durch Anklicken eines Lexems (oder einer Menge von Lexemen) eine Hypertextstruktur bekommt, welche die Dokumente, in denen dieses Lexem vorkommt (oder diese Lexeme vorkommen), enthält.

## 4 Der Retrievalvorgang

Der Retrievalvorgang spielt sich in folgender Weise ab:

1. Eine Menge von gespeicherten Dokumenten wird in das System eingegeben, und es erfolgt sowohl eine Bestimmung der Topikhierarchien als auch der Kollokationsdependenzen. Die generierten Topikhierarchien werden an die *hyperview* des Kollokationsnetzes angehängt.

<sup>2</sup>Wir benutzen dabei die Terminologie von *hypercard*.

<sup>3</sup>Diese globale *hyperview* kann dann durch Interaktion mit dem Benutzer weiterhin restringiert werden bis der Benutzer eine Struktur erreicht, die für seine Zwecke ausreichend ist. Diese rekursive Restriktion basiert allerdings auf kategorientheoretische Formalismen, die hier nicht erörtert werden können, und die zu einer Lösung des Problems von "getting lost" führen ((Rieger & Thiopoulos, 1991) und (Bruza & Thiopoulos & van der Weide, 1992)). Diese Formalismen wurden im Rahmen einer Modellierung des phänomenologischen Bedeutungsbegriffs entwickelt ((Thiopoulos, 1990) und (Thiopoulos, 1992)).

2. Nachdem der Benutzer das Netz inspiziert hat, kann er für ihn interessante keywords anklicken. Damit gelangt er in einen korrespondierenden Topik, in dem die keywords vorkommen.
3. Er kann jetzt die entsprechende Topik-hierarchie inspizieren. Für den Fall, daß andere Teile dieser Topikhierarchie sein Interesse erregen, kann er durch Anklicken vorhandener *buttons* den Text durchmustern. Andernfalls kann er zum nächsten Text übergehen, genauer gesagt: zum nächsten Topik, in dem die spezifizierten keywords vorkommen. Von da aus hat er dieselben Möglichkeiten.
4. Während des Navigierens kann der Benutzer die für ihn interessanten Topiks in einer Hypertextstruktur ablegen, die das Ergebnis des Retrievalprozesses bildet.
10. Polanyi, Livia/Scha, Remko J .H. 1983. On the Recursive Structure of Discourse. In: Connectedness in Sentence, Discourse and Text. (Eds.) Ehlich, K. /van Riemsdijk, Henk, 141-178. Tilburg: University.
11. Rieger, B. *Unscharfe Semantik*. Frankfurt: Peter Lang, 1989.
12. Rieger, B. & Thiopoulos, C. A self organizing lexical system in hypertext. In *Proceedings of the first Conference on Quantitative Linguistics*. Trier, 1991. ( demnächst)
13. Thiopoulos, C. Meaning metamorphosis in the semiotic topos. *Theoretical Linguistics*, 16, 2j3, 255-274, 1990.
14. Thiopoulos, C. Towards a logic of semiotic systems. In *Revue Mathematiques, Informatiques et Sciences Humaines*, December, 1991.

## 5 Bibliographie

1. Barrett, E. (Ed.). *Text, ConText, and HyperText*. MIT: 1988.
2. Bergler, Sabine 1991. The Semantics of Collocational Patterns for Reporting Verbs. Proceedings of the 5th *EACL* Conference 1991 Berlin, 216-221
3. Bruza, P. D. & van der Weide T. P. Assessing the quality of hypertext views. *ACM SIGIR FORUM*, 24(3), 6-25, 1990.
4. Bruza, P. D. & Thiopoulos, C. & van der Weide, T. P. A categorial search in multi level hypermedia, 1992. (demnächst)
5. Carnap, R. *Logical foundations of probability*. Chicago: The University of Chicago Press, 1950.
6. Conklin, J. Hypertext: an introduction and survey. *Computer*, Vo1.20, No.9, September, 1987.
7. Furnas, G.W. Generalized fisheye views. *Human Factors in Computing Systems, CHI'86 Conference Proceedings*. Boston: 1986.
8. Jones, W. P. How do we distinguish the hyper from the hype in non-linear text? In: H.J. Bullinger & B. Shackel (Eds.). *INTERACT'87*. North Holland, 1987.
9. Nakhimovsky, Alexander/Rapaport, W.J.1988. Discontinuities in Narratives. Proceedings of COLING 1988 Budapest, 465-470
15. Weber, Heinz J. Faktoren einer textbezogenen Maschinellen Übersetzung: Satzstrukturen, Kohärenz- und Koreferenz-Relationen, Textorganisation. In: Batori, J. S. & Weber, H. J.(Eds.). *Neue Aufsätze in Maschineller Übersetzung: Wissensrepräsentation und Textbezug*. Tübingen: Niemeyer, 1986.
16. Weber, Heinz J. Text-Serien, Text-Ebenen und Text-Instanzen in journalistischen und amtlichen Nachrichten In: Fachsprachlichkeit in Texten. (Ed.) Kalverkämper, H., Tübingen: Narr. ( demnächst)

## Bericht über die GWAI '90

Hans-Dieter Lutz, Univ. Ko-La, CL  
(18.9.1990)

Die 14. GWAI (German Workshop on Artificial Intelligence) fand vom 10.-14. September 1990 wohl zum letzten Mal in der Abgeschlossenheit von Eringerfeld bei Geseke statt (nach allem, was "man" hörte, soll die nächste GWAI ins Gustav-Stresemann-Haus nach Bonn umziehen). Ca. 200 Teilnehmer waren gekommen, davon nur 11 (in Worten: elf) aus der (zur Zeit der Abfassung des Berichts) "Noch-DDR" (ich hoffe, daß Ingrid Matthäus-Meier Recht behält, die einmal anlässlich einer Parlamentsdebatte gesagt hat: "Ein Kennzeichen der deutschen Sprache ist ihre Unkaputtbarkeit.")).

Neben 34 Fachvorträgen (davon 6 zu 'Wissensrepräsentation', 2 zu 'Mensch-Maschine-Interaktion', 4 zu 'Expertensysteme', 4 zu 'Deduktion', 8 zu 'Natürlich-sprachliche Systeme', 3 zu 'Maschinelles Lernen', 2 zu 'Bildverstehen' und 1 zu 'KI-Programmiersprachen') standen 2 eingeladene Hauptvorträge (G. Cottrell, B. Bartell, C. Haupt: *Grounding Meaning in Perception* und B. Wielinga, G. Schreiber: *KADS: Model Based KBS Development*).

Die Verteilung der Fachvorträge läßt nicht mehr die Schlußfolgerungen zu, die Lena Bonsiepen noch bezüglich der GWAI '89 ziehen konnte (vgl. KI 2/90, S. 20f.). Dennoch läßt sich eine Dominanz von den Bereichen 'Natürlich-sprachliche Systeme' (dem man im Grunde noch einen der beiden MCI-Vorträge hätte zuschlagen können) und 'Expertensysteme' feststellen, dann nämlich, wenn man noch die Tutorien und die Workshops hinzunimmt. Tutorium 1 'Knowledge Engineering' umfaßte einen ganzen Tag (Mo nachmittags und Di vormittags), Tutorium 2 'Wissensbasierte Verarbeitung gesprochener Sprache' den Montagnachmittag. Die Work

shops fanden am Mittwoch und Donnerstag jeweils nachmittags statt und hatten die Themen: 'Semantisch-pragmatische Verarbeitung von Pluralen und Quantoren in NLP - Ansätze, Ergebnisse, Perspektiven', 'Logik als Programmiersprache', 'Robotik und KI: Paradigmen und Projekte', 'Problemklassen von Expertensystemen' und 'KI-Produkte im praktischen Einsatz'. Der Workshop 'Problemklassen von Expertensystemen' war eine closed-shop-Veranstaltung, die aber am Freitag zum Ende der Tagung für die Allgemeinheit in Form einer Podiumsdiskussion geöffnet wurde.

Die Empfehlung der Fachbereichsleitung (aufgrund der Erfahrungen mit der GWAI '89) an das Programmkomitee der GWAI '90, den Workshop-Charakter durch Beteiligung und organisatorische Einbindung der Fachausschüsse stärker hervorzuheben, hat teilweise gegriffen; "teilweise" deshalb, weil es wohl mehr Workshops gab, aber es m.E. nicht in ausreichendem Maße gelungen ist, "mehr themenübergreifende Workshops anzubieten, die ein breites Spektrum der GWAI-Teilnehmer interessieren" (Lena Bonsiepen, a.a.O.).

Am Mittwochabend gab es eine vergleichsweise "zahme" Sitzung der "Fachbereichsleitung" (gemeint war aber Fachbereichsgemeinde), deren wesentlicher Punkt wohl war, daß sich die finanzielle Situation des Fachbereichs 1 'Künstliche Intelligenz' der GI (und damit die Beziehung zur Dachorganisation) in Zukunft wohl erträglicher und freundlicher gestalten lassen wird, wenn die Zeitschrift KI nicht mehr vom Oldenbourg-Verlag, sondern (wie es derzeit den Anschein hat) vom FBO (welcher genaue Name sich dahinter versteckt, blieb ungewiß; er geht wohl

in die Richtung 'Fachverlag für Büro und Organisation') verlegt wird. Damit waren's alle zufrieden, und man konnte nach gut einer Stunde, wie es C.-R. Rollinger formulierte, "zum Biere schreiten".

Ja, und am Donnerstagabend - um bei den social events zu bleiben - gab es noch ein Bankett mit den diversen durchaus angebrachten Danksagungen und einer Neuigkeit: es wurde zum ersten Mal der 'best paper award' vergeben und zwar an Stefan Wermter (Fb Informatik der Universität Dortmund) für seinen Beitrag *Combining Symbolic and Connectionist Techniques for Coordination in Natural Language*. In guter deutscher Tradition wurde die materielle Seite des von SIEMENS gesponserten Preises nicht genannt.

Aus der Zahl der Fachvorträge, der Tutorien und der workshops dürfte schon ersichtlich sein, daß so einiges parallel stattfinden mußte. Ich habe besucht das Tutorium 2, den workshop 1 und die meisten Vorträge aus den Bereichen 'Mensch-Maschine-Interaktion' und 'Natürlichsprachliche Systeme'. Zu den Fachvorträgen und den beiden eingeladenen Vorträgen möchte ich hier nichts sagen, sie sind nachzulesen im Tagungsband:

Marburger, Heinz (Hrsg.):

GWAI-90. 14th German Workshop on Artificial Intelligence.

Proceedings. Eringerfeld, September 1990.

(Informatik-Fachberichte. 251)

Berlin: Springer 1990.

Nun aber einiges zu Tutorium 2 und workshop 1. Tutorium 2 'Wissensbasierte Verarbeitung gesprochener Sprache' war organisiert worden von Günter Görz (Universität Hamburg) und bestand aus vier Vorträgen:

1. Sprachsignalanalyse  
(Euler/Hübener, Univ. Hamburg, die für Noeth, Univ. Erlangen-Nürnberg eingesprungen waren)
2. Phonologie (, Morphologie) und Syntax  
(Gibbon, Univ. Bielefeld)
3. Verstehenskomponenten für Systeme gesprochener Sprache (Niedermayr, SIE MENS AG)
4. Sprachsynthese (Hess, Univ. Bonn)

Das für den Schluß des Montagnachmittags vorgesehene Panel 'Integrierte Architekturen für sprachverarbeitende Systeme' fiel dem Zeitmangel zum Opfer. Leider hatte

nicht nur ich den Eindruck, daß die Absprache der Referenten untereinander minimal war; ein roter Faden war nicht zu entdecken, eine einheitliche Terminologie auch nicht und das "Ziel, einen mit den Grundzügen der KI und der (text-orientierten) Computerlinguistik vertrauten Hörerkreis in die wichtigsten Begriffe und Methoden der Verarbeitung gesprochener Sprache einzuführen. Dabei soll insbesondere der Aspekt einer Integration der Verarbeitung von gesprochener und geschriebener Sprache betont werden" (Ankündigungstext) hatten sich auch nicht alle zueigen gemacht. Das war insofern bedauerlich, weil das strategische Ziel, "dem Themenfeld 'Gesprochene Sprache' zukünftig bei der GWAI etwas mehr Gewicht zu verleihen", um auch nicht den Anschluß an die internationale Forschung zu verlieren, dadurch nicht gerade näher rückte.

Das Eigentümliche war, daß einerseits öfter mal gesagt wurde, daß es bei speech understanding systems nicht darum gehen könne, einfach ein speech front end vor und ein speech back end hinter ein klassisches NL-System zu schalten, daß aber auf der anderen Seite von der Besonderheit und Eigenheit 'der' gesprochenen Sprache praktisch keine Rede war (Ausnahme: Niedermayr).

Euler/Hübener behandelten Abstraktions-ebenen in der Sprachverarbeitung [gemeint war: phonetischen Sprachanalyse] (Vorverarbeitung, Merkmalsextraktion, Vektorquantisierung, Segmentierung, Lexikonsuche), das Sprachsignal (die Entstehungsarten von Lauten anhand eines Modells des menschlichen Sprechapparates: wie schaut ein Sprachsignal aus, welche Spektren hat es), Analyseverfahren (gehörbezogene Vorverarbeitung, Analyseverfahren im Frequenzbereich: Kurzeitspektrum, linear predicting coding, homomorphe Analyse), stochastische Verarbeitung (zur Klassifizierung von Merkmalen aufgrund von damit gewonnenen diskreten Prototypen von akustischen Einheiten), regelbasierte Verarbeitung (zur Segmentierung und zum Labelling).

Gibbon setzte ausdrückliche 'wissensbasiert' mit 'linguistisch' gleich und sprach über Probleme der Phonologie (Hierarchie von Segmenten? Sind distinktive Eigenschaften nur den kleinsten Segmenten zuzuordnen? Ist Variation durch den Kontext oder durch abstraktere Faktoren bestimmt?), der Computational Phonology (phrase structure phonological parsing, two-level-Systeme, feature-based syllable and word phonotactics), der Morphologie (Zuordnung von Morphem-Inventar zu Wort-Inventar, Flexi-

onsmorphologie), der Syntax und der Lexik (letztere in bezug auf Repräsentations- und Regelformate).

Niedermayr referierte über Unterschiede und Probleme mit Bezug auf andere NLP-Systemtypen, über die Architektur (bottom-up mit dem Problem der 'kombinatorischen Explosion', top-down mit dem Generierungsproblem, up-down mit dem Kontrollproblem, seriell - integrativ (mit dem Kommunikationsproblem der einzelnen Komponenten untereinander), den 'unsauberen' input (Elliptik, was heißt hier 'Korrektheit?'), die Verarbeitungsgeschwindigkeit der akustischen Erkennung und der linguistischen Analyse (Echtzeit-Problematik), den Abdeckungsgrad (im Vokabular: je größer desto schlechter? in der Grammatik: je freier desto schöner?), human factors (restriktiver Dialog vs. 'Freistil', Kanalprobleme: Störungen aus dem Hintergrund) und die einzelnen Komponenten (Lexikon, Prosodie, Syntax, Semantik, Dialog, Pragmatik) und der damit verbundenen Suchproblematik.

Hess berichtete anhand von vielen wahrnehmbaren Beispielen über die bisherigen Erfolge, aber auch über die Probleme und derzeitigen Grenzen der Sprachsynthese, die eben nicht darin bestehen könne, eine einfache Graphem-Phonemzuordnung mit anschließender Verkettung und Transformation in kontinuierliche Signale vorzunehmen. Um Natürlichkeit und Verständlichkeit zu erreichen, ist schon für "einfache" text-to-speech-Systeme vor der eigentlichen Synthese eine semantische und syntaktische Analyse der Ausgangsäußerung notwendig, um Fragen der Intonation und Prosodie adäquat, d.h. Hörerzentriert behandeln zu können.

Wie gesagt: die eigentlich vorgesehene Diskussion über integrierte Architekturen kam nicht im vorgesehenen Rahmen zu Stande. Görz versuchte zwar, dieses Thema noch in die allgemeine Runde von Verständnis- und inhaltlichen Fragen im Anschluß an den Vortrag von Hess einzubringen, aber die fehlende Zeit ließ keine vertiefte Erörterung zu, zumal wir uns nicht darauf einigen konnten, ob wir diesen Problemkreis rein akademisch (d.h. losgelöst von realen Anwendungen) oder aber zweckorientiert diskutieren sollten (wenigstens wurden dabei die unterschiedlichen Interessenslagen ansatzweise offengelegt). So kam es nur zu zwei essentials (für mich wenigstens) :

1. Durch Zusammenstöpseln (Spracherkennung+NLPS+Sprachsyntetisierung) ist nichts gewonnen.

2. a) "Einzelworterkennung gibt es nicht."
- b) "Satzerkennung, als aus Einzelworterkennung bestehend, gibt es noch weniger." (Gibbon)  
(Die Frage, ob man diese Liste noch weiter fortsetzen müßte, wurde nicht mehr diskutiert.)

Noch eine Bemerkung zur Veranstaltungsorte 'Tutorium': Sie scheint sich dadurch auszuzeichnen, daß viele Referenten immer wieder zu folgender Metabemerkung sich genötigt sahen: "Ich möchte das jetzt nicht im Detail ausführen, dazu ist eine Menge {Mathematik - Linguistik - ...} notwendig."

Im workshop 1 'Semantisch-pragmatische Verarbeitung von Pluralen und Quantoren in NLP - Ansätze, Ergebnisse, Perspektiven', den Jürgen Allgayer (Univ. des Saarlandes) organisiert hatte, waren solche Einschränkungen nicht vonnöten.

Nerbonne (DfKI, Kaiserslautern) gab für den verhinderten Hinrichs - einen einführenden Überblick über die Arbeiten der letzten 10 (!) Jahre unter dem Titel Coping with Plurals: distributive vs. kollektive Lesart von pluralen NPs, Desambiguierung aufgrund von Prädikationen (meist durch Verben: distributives *schwimmen*, kollektives *sich treffen*, 'mixed' *tragen*), mengentheoretische vs. algebraische 'Ontologie', die Arbeit von Godehard Link 1983, die Weiterentwicklung bis Lonning 1989 (die zur Feststellung führte, daß eine Plurallogik in einer Prädikatenlogik zweiter Stufe landet), die kumulative Lesart von kardinal quantifizierten NPs (*Zwei Firmen kauften drei Maschinen*. Scha 1989).

Link (LMU München) stellte eine Axiomatisierung seines 1983-Ansatzes vor, der sich durch eine 'structured ontology' auszeichnet (wobei es nicht um Seins-Fragen, sondern um die Modellierung einer Domäne geht; d.h. etwa: ob ein individual als atom oder als einelementige Menge aufgefaßt wird, entscheidet sich nur in der Theorie; wobei Link das im ersten Augenblick verblüffende Prinzip vertrat: Guck *nicht zu genau* hin).

Nerbonne plädierte in seinem eigentlichen Referat *Comparative Determiner und Generalisierte Quantoren* für die Einführung von Meßfunktionen zur Abbildung von natürlichen und reellen Zahlen als parametrische Maße in solchen Fällen wie *weniger/mehr als, zweimal soviel wie* u.a. Und wurde dabei in der Diskussion von Link unterstützt mit der Bemerkung: "Zählen (Kardinalität), ist Messen."

Latecki (Univ. Hamburg)/Pinkal (Univ. des Saarlandes) referierten über *Syntactic and Semantic Conditions for Quantifier Scope* und stellte dabei einige Verfahren vor (Cooper Storage Treatment, Nested Cooper Storage, Hobbs/Shieber-Algorithmus, weak crossover effect von Tanja Reinhardt), beleuchteten die Schwächen der Verfahren, präsentierten Erweiterungen und technische Modifizierungen, um schließlich bei der DRT zu landen und den für mich eigentlichen Gag zu präsentieren: "man muß erst Semantik betreiben und sehen, in welchem Skopus eines Quantors eine Variable landet, und dann erst syntaktisch abtesten mittels einer entsprechend modifizierten c-control-Bindungsregel (die tunlichst asymmetrisch sein sollte !)."

Heinz (ÖFAI, Wien) berichtete unter dem Titel *Treating Non-Singular Terms in a NL Consulting System* über das neue, auf DIALOG aufbauende Projekt, das mehr Aktivität des Systems und mehr Diskurs zulassen solle (DIALOG war ein klassisches FAS), und konzentrierte sich dabei vor allem auf die Eigenschaften von NPs (wie deskriptiver Inhalt, Referentialität, Diskursabhängigkeit, Definitheit, Quantifikation, Kardinalität, Generizität), ihre Einbindung in eine diese Eigenschaften berücksichtigende Syntax (DP-Syntax) und ihre Behandlung im Rahmen einer situation semantics.

Fliegner (vormals WISBER, Hamburg, jetzt Software AG, Hamburg und nach eigenen Worten "Freizeitlinguist") beschäftigte sich mit Komplexitätstheorie- Betrachtungen bei der Repräsentation von Pluralphänomenen durch Quantifikation über Ereignisse. Mit Carola Eschenbachs (Univ. Hamburg) *Repräsentation von Pluralen in einer KL-ONE-ähnlichen Sprache* war der Bogen gespannt zur Wissensrepräsentation. Sie stellte einen modelltheoretischen Ansatz vor, der durch Definition von numerischen Primitiven in der A-box der bislang unstrukturierten A-box eine Struktur verleiht und damit z.B. die Kumulativität zu repräsentieren erlaubt; dafür gibt es dann aber Probleme mit *all, atleast, atmost*.

Carola Reddig-Siekmann (Univ. des Saarlandes versuchte sich in ihrem Referat *The Plural's Consequence: Legal and Illegal Inferences based on NL Quantifiers* an einer präzisen Definition von NLP-Anforderungen an eine KR/R-Sprache (in Anlehnung an Brachmann: knowledge representation and reasoning), die zu einer Modifizierung gegenwärtiger KR/R-Sprachen führen müsse; dafür zu sorgen, sei dann der Job der KR-Leute.

Damit kann ich gleich zu dem ersten meiner

Eindrücke übergehen, mit denen ich diesen workshop facettenartig kommentieren möchte: Gerade die letzten 30 (1) Jahre haben gezeigt, daß diese Art der Arbeitsteilung, wie sie sich Reddig-Siekmann vorstellt, nicht funktioniert; erst der Prozeß, daß Linguisten Logik machten und Logiker Linguistik (um es recht salopp zu formulieren), hat zu einer Verständigungsbasis geführt und zu gemeinsamen Prinzipien, die vorher Gegenstand heftiger Kontroversen waren (Beispiel: die Natürlichkeit der Repräsentation von NPs in Form von quantifizierten Termen).

Zweiter Eindruck: die Prädikatenlogik 1. Stufe (mit den darin möglichen Erweiterungen) hat durch die KL-ONE-Familie Konkurrenz bekommen, aber - soweit ich es bislang beurteilen kann - nur auf der Formalismus-Dimension; beide haben aber analoge Probleme bei der Behandlung nicht-klassischer Quantoren/Quantifikatoren. Vielleicht sollte man einmal darüber nachdenken, woran das liegen mag.

Dritter Eindruck: Mir wurde manchmal zu undifferenziert über Determinatoren geredet; es blieb oftmals völlig unklar oder unabgegrenzt, was an NL-Ausdrücken darunter fallen soll/kann. (Beispiel: *Jeder Student sah einen anderen Film*. Mal davon abgesehen, daß Einigkeit darüber bestand, daß man die 'Bedeutung' dieses Satzes bislang nicht angeben kann, blieb mir doch die Frage, ob *ander*-hier als Determinator gehandelt werden soll.)

Vierter Eindruck: Die Intention des workshops: "der Vergleich der formalen Apparate [zur Repräsentation, formalen Behandlung und Verarbeitung von Pluralen und zur formalen Beschreibung und (effizienten) Verarbeitung von Quantoren, eines der interessantesten Themen im Berührungsfeld zwischen Linguistik und KI], die Evaluation der Anwendbarkeit für bestimmte Anforderungen innerhalb der Sprachverarbeitung sowie Vorschläge zur Kombination, Variation oder Erweiterung der Theorien" wurde weitgehend eingelöst, vielleicht mit dem Abstrich, daß nur Heinz Ansätze zur "Evaluation der Anwendbarkeit für bestimmte Anforderungen innerhalb der Sprachverarbeitung" (wohl aufgrund ihrer Erfahrungen mit DIALOG) einbrachte. Fünfter Eindruck: (schon eine Sottise:) Zu wenig Zeit, um ins einzelne gehende Fragen (mit schwerwiegenden Folgen für das Ganze?) an- und ausdiskutieren zu können.

## Neue Produktionssysteme, Gruppenorientierte Arbeitsorganisation und Wettbewerbsfähigkeit

Magdalena Lutz-Hensel  
(14.11.1990)

### 1 Veranstalter und Umfeld der Konferenz

Auf Einladung des Instituts Arbeit und Technik des Wissenschaftszentrums NRW habe ich an der Internationalen Konferenz: "Neue Produktionssysteme, Gruppenorientierte Arbeitsorganisation und Wettbewerbsfähigkeit" am 26.-27.9.90 in Gelsenkirchen teilgenommen. Ausgerichtet wurde die Konferenz gemeinsam von

- . den Forschungsteams, die an dem internationalen Forschungsprojekt über neue Produktionssysteme und die sozialen und technischen Voraussetzungen für ihren Einsatz arbeiten und zwei Tage der Konferenz für den internen Austausch über ihre Arbeiten nutzten,
- . Koordinatoren dieser Arbeiten am Institut Arbeit und Technik und
- . dem Initiativkreis Ruhrgebiet, der das Ziel verfolgt, mit besonderen Veranstaltungen aus Kultur und Wissenschaft der Region neue Aufgabengebiete zu erschließen und auch so nach dem Rückgang von Bergbau und Schwerindustrie Initiativen für Investitionen auf breiter Basis zu setzen.

Auftraggeber für das Projekt ist die EG mit dem FAST-Programm einerseits und der Minister für Arbeit, Gesundheit und Soziales des Landes NRW mit dem Programm SoTech. Weiterhin ist das internationale Forschungsnetzwerk CAPIRN zu nennen, das ebenfalls unter der Rubrik "Veranstalter und Förderer der Konferenz" figuriert und wohl den Veranstaltern zuzurechnen ist.

Falls der Eindruck entstehen sollte, daß die Zusammenhänge zwischen den Aktivitäten der genannten Gruppen verzwickelt seien, so ist das beabsichtigt, wirft es doch ein Licht auf das Konglomerat von Interessen, dem zumindest die Outsider sich ausgesetzt sahen.

Bestärkt wird dieser Eindruck durch wechselseitige Übernahme von Aufgaben zwischen den Institutionen und die Anzahl

verschiedener Funktionen, in denen Wissenschaftler zu gleicher Zeit tätig sind.

Um den Blick auf die verschiedenartigen Interessen zu lenken, sei das folgende festgehalten: Das Forschungsprogramm FAST (Forecasting and Assessment in Science and Technology) der EG untersucht die Auswirkungen des wissenschaftlichen und technologischen Wandels für die langfristige wirtschaftliche und soziale Entwicklung der europäischen Gesellschaften. Sein Zweck ist es, in Forschung und technologischer Entwicklung Prioritäten für die EG zu setzen. Einige der derzeitigen Schwerpunkte seiner Forschung sind:

- . die Globalisierung der Wirtschaft und deren Folgeerscheinungen auf lokaler Ebene der EG-Staaten im Bereich Wissenschaft und Technik zu erfassen,
- . Wissenschafts- und Technologiepotentiale den gesellschaftlichen Bedürfnissen anzupassen.

Im Forschungsprogramm CAPIRN (International Network on Culture and Production) arbeiten Wissenschaftler aus den U.S.A., Japan, Großbritannien, Italien, Dänemark und der Bundesrepublik an Projekten zur Entwicklung, dem Transfer und der Anwendung von Automatisierungstechniken. Ländervergleichende Studien beobachten, wie neue Produktionstechnologien abhängig von nationalen Industriekulturen gestaltet werden, und widmen sich beispielsweise

- . der Bedeutung der Facharbeit im Produktionsprozeß,
- . den Ausbildungssystemen,
- . der innerbetrieblichen Arbeitsorganisation,
- . dem Ausmaß und der Art des Technikeinsatzes,
- . den Einflüssen nationaler, regionaler und kommunaler Politik auf Innovationsprozesse,

- besonderen Formen industrieller Beziehungen zwischen den Nationen.

**Das SoTech-Programm** (Mensch und Technik - Sozial verträgliche Technikgestaltung) ist Teil der Nordrhein- Westfalen-Initiative Zukunftstechnologien der Landesregierung. Das Programm soll technisch-ökonomische Innovationen mit solchen im sozialen Bereich verbinden und beispielsweise den öffentlichen Dialog über sozialverträgliche Technikentwicklung und -nutzung fördern, die individuellen Handlungsmöglichkeiten der Berufstätigen durch Weiterqualifizierung stärken und Mitbestimmungsmöglichkeiten neu eröffnen.

**Das Institut Arbeit und Technik** des Wissenschaftszentrums NRW beschäftigt sich zur Zeit besonders mit neuen Informationstechnologien und sozialverträglichen Technikkonzepten, d.h. es sollen Wege gefunden werden, qualifizierte Arbeit mit moderner Technologie so zu verknüpfen, daß sich der Strukturwandel in NRW zugunsten der arbeitenden Menschen auswirkt.

International war die Beteiligung der Forschungsteams und der Vortragenden mit Vertretern der Wirtschaft und Wirtschaftswissenschaft aus den U.S.A, Japan, der U.D.S.S.R., Australien und der EG-Staaten. Unter den Zuhörern waren zusätzlich Teilnehmer aus Schweden und der Schweiz, deutsche Teilnehmer waren jedoch in der Mehrheit, davon sicher über die Hälfte aus NRW.

Dem Ziel, den öffentlichen Dialog über sozialverträglichen Technikeinsatz zu fördern, entsprach die Zusammensetzung der Hörschaft, vertreten waren die Politik (Ministerien, Parteien, Kommunen), (Weiter-)Bildungsinstitutionen, Institutionen der Berufsförderung, Gewerkschaften, Arbeitgeberorganisationen, die Wissenschaft (als Wirtschafts- und Sozialforschung, Arbeitswissenschaft, Wissenschaftsforschung, Hochschuldidaktik u. a.), die Forschungsk Kooperation, die Informationswirtschaft und der Bereich des Informationstransfers.

Die Vorträge des ersten Tages befaßten sich gleich mit dem Thema, das sich als zentraler Begriff der Konferenz herausstellen sollte: den anthropozentrischen Produktionssystemen (APS), und zwar

1. als Ziel der EG-Politik für die Zukunft,
2. als Garantie für Produktivität und Qualität der Arbeitsleistung,
3. als Mittel zur Sicherung der

Wettbewerbsfähigkeit auf dem Weltmarkt.

Der vierte Vortrag führte Beispiele für solche Produktionssysteme aus der EG an.

## 2 Zusammenfassung der Vorträge des 1. Tages

### 2.1 Situation der achtziger Jahre in den Industrieländern

Das Prinzip, wonach sich der Markt durch Konkurrenz von selbst regelt, ist so unbrauchbar. Es zeigte sich, daß Preissenkung und Massenproduktion *allein* eine florierende Wirtschaft nicht garantieren: Preissenkung wurde z.B. erreicht durch Verlagerung der Fertigung in Billiglohn-Länder, was immense Transportkosten und den Verlust von einheimischen Arbeitsplätzen zur Folge hatte. Preissenkung wurde z.B. angestrebt durch hochgradige Automatisierung der Produktionsprozesse im Inland. Dabei entstanden erhöhte Kosten durch steigende Störanfälligkeit der Automaten und Produktionsstockungen infolge ausgedehnter Wartungszeiten. Außerdem traten etwa im Maschinenbau folgende Probleme auf:

- Das Material lagert im Fertigungsprozeß zu 95% der Durchlaufzeit unbearbeitet, was eine Kostensteigerung bedeutet.
- 45% des Betriebsvermögens ist in Lagerbeständen angelegt, - nur 13% in Produktionsmitteln, was bedeutet, daß dieses Kapital unproduktiv ist.
- 100 direkten Werkern stehen 144 Personen gegenüber, die nicht direkt im Fertigungsprozeß tätig sind, aber auf indirekte Weise für den Ablauf sorgen müssen, was das Endprodukt wegen hoher Lohnkosten verteuert.

Auch die computerintegrierte Produktion (CIM) brachte als technikzentrierte Produktion ihre Probleme mit sich:

Der hohe Software-Aufwand ist kapitalintensiv und risikoreich.

Die Produktion ist zu inflexibel bezüglich des Produkt- und des Produktteilespektrums.

Sie ist nur begrenzt flexibel bei Veränderungen im Fertigungsprozeß.

- . Sie bedeutet eine Verschwendung von Facharbeiterqualifikationen, da sich nicht benötigte Qualifikationen auf Dauer abbauen und bei Bedarf später nicht mehr nutzbar gemacht werden können. Auch die Speicherung von Facharbeiterwissen in Expertensystemen erscheint problematisch, weil das Wissen im Kontext individueller Erfahrungen über einen Zeitraum hinweg entstanden ist, die Erfahrungszusammenhänge so aber nicht mit-modelliert werden können.
  - . Das Management findet sich in Abhängigkeit von oft undurchsichtiger Software, was sich entscheidungshindernd auswirkt, wenn nicht dadurch Fehlentscheidungen entstehen.
- Gesetzt den Fall, Massengüter können billig produziert werden, so ist damit noch keine Absatzgarantie gegeben, weil der Markt in den Industrieländern für diese Produkte als gesättigt gilt und sich in den armen und verarmten Ländern keine zahlungsfähigen Interessenten finden.
- Dazu kommt, daß in den Industrieländern gesteigerte Ansprüche an Produkte gestellt werden, u. a. auch bezüglich ihrer Umweltverträglichkeit.
- In solcher Situation macht das Zusammenwachsen der EG-Staaten neue Überlegungen zur Wirtschaftlichkeit und sozial-ökonomischen Einbettung des Marktes nötig.
- . Diversifizierungs- und Nischenstrategien sind nur in Einzelfällen empfehlenswert, das Problem muß grundsätzlich, und zwar über veränderte Produktionsstrukturen angegangen werden.
  - . Da qualitativ hochwertige Waren hohes Innovationstempo erfordern und einen unsicheren und risikoreichen Markt bilden, auf dem beispielsweise auch alternative Produkte erprobt werden, ist eine hohe Flexibilität in der Produktion wünschenswert.
  - . Diese Flexibilität ist nicht durch bloße Arbeitsteiligkeit zu erreichen, sondern durch autonome Handlungsregelungen in Gruppen - anstelle der Anordnungen von außen. Gefragt ist also eine Umstellung von hierarchischen Organisationsformen auf dezentralisierte und dezentrale Organisationsformen.
  - . Die Herstellung von Produkten höchster Qualität muß in angemessener Zeit an immer neue Ansprüche angepaßt werden, das bringt kurze Produktzyklen mit sich, stellt hohe Anforderungen an Arbeitsorganisation und Management, verlangt ein hohes Ausbildungsniveau und zugleich ein breites Ausbildungsspektrum der Mitarbeiter und erfordert Arbeitskräfte, die hochmotiviert sind.

### 3 Änderungsvorschläge für den EG-Bereich

- . Standard- und Massenproduktion sollten auf geeignete Sparten begrenzt bleiben.
  - . Fertigungssysteme sind nicht aufgrund der zur Verfügung stehenden Technologie, sondern in erster Linie bezüglich ihrer Organisation zu optimieren. D.h. dem Ablaufplan-Dilemma, wonach Stückzahlen bestimmter Höhe die Durchlaufzeiten während der Fertigung unverhältnismäßig verlängern, ist durch Dezentralisierung zu begegnen; d.h. es müssen Teile-Familien konzipiert werden, deren Fertigung in einiger Unabhängigkeit voneinander erfolgen kann.
  - . Eine Umstellung auf qualitativ hochwertige Produkte, die den Kundenwünschen entsprechen, dauerhaft sind und mit Dienstleistungen gekoppelt verkauft werden, ist notwendig, um einheimische Märkte zu sichern.
  - . Um diese neuen Produktionsformen zu gewährleisten, ist es notwendig, die Vielzahl technischer Möglichkeiten zur Unterstützung menschlicher Fähigkeiten zu nutzen, nicht jedoch diese Fähigkeiten durch technische Systeme nachzuahmen und gar zu ersetzen.
- Der technikzentrierten Produktion wird so eine menschenzentrierte Produktion entgegengesetzt, die die Vermeidung von Problemen verspricht, die die technikzentrierte hervorbrachte. Demnach werden nicht entstandene Probleme gelöst, sondern durch Umdenken wird versucht, diese Probleme erst gar nicht entstehen zu lassen. (- eine zweifellos vielversprechende "Problemlösungsstrategie" ! -) Für die so veränderten Produktionsweisen, deren Folgen tiefgreifende Auswirkungen auf die gesamte Betriebs- und Wirtschaftsorganisation haben, ist der Begriff der ANTROPOZENTRISCHEN PRODUKTIONSSYSTEME geprägt worden.

## 4 Der Begriff Anthropozentrisches Produktionssystem (APS)

Der Begriff "anthropozentrisches Produktionssystem" wurde wechselnd mit "humanorientierter Produktion", "menschengerechter Produktion" u.a. verwendet und als "anthropocentric production systems" übersetzt. (Der Schöpfer des Begriffs ist offensichtlich ein Mitglied des Instituts Arbeit und Technik in Gelsenkirchen.)

Assoziationen an das anthropozentrische Weltbild stellen sich zwar selbstredend ein, sind aber ebenso fehl am Platz wie Assoziationen an klientenzentrierte Beratung. Überhaupt blieb unklar, was die Wortkonstruktion besagen soll; sogar, ob sie etwa zu paraphrasieren wäre als: Produktionssystem, in dessen Mittelpunkt (?1) ein Mensch, ... eine Arbeitsgruppe von Menschen, ... der Mensch, ... ein Menschenbild, ...die Humanität steht. - Irgendwie fühlte ich mich an artgerechte Tierhaltung erinnert, wobei die Frage aufkommt, ob der Art des gehaltenen Tieres oder der Art des Tierhalters entsprochen werden soll. Der Begriff "anthropozentrisches Produktionssystem"

- läßt offen, in welcher semantischen Beziehung das Adjektiv zum Substantiv steht,
- verwendet den Begriff "Mensch" changierend,
- ist auch bezüglich des Begriffs "Produktionssystem" unscharf.

Trotzdem - oder: deswegen (?) - wurde er so häufig benutzt, daß er als Konferenzthema hätte erscheinen können.

Man kann die Frage danach, worum es bei dieser Konferenz ging, aber darüber angehen, daß gefragt wird, in welchem Sinne APS von den Referenten der Vorträge und des Praxisforums, zuletzt des Panels gebraucht wurde. Die Spannbreite reichte dabei, um es pointiert zu sagen, von der Einrichtung von Kohlefiltern bei Verbrennungsanlagen bis zur Nutzung der menschlichen Kreativität schlechthin. Immerhin kam dabei heraus, daß neue Produktionssysteme nicht mehr produktorientiert noch technikorientiert sein sollen, oder anders: daß es nicht länger darauf ankommt, soviel wie möglich ohne Rücksicht auf den Absatzmarkt zu produzieren, und nicht darauf, Hochtechnologie ohne Berücksichtigung des ökonomischen, ökologischen und sozialen Nutzens einzusetzen. Aus verschiedenen Perspektiven und Interessenlagen fügten

sich im einzelnen etwa folgende Vorstellungen über die neuen Produktionssysteme zusammen:

- Anthropozentrik ist ein Wert an sich; wenn menschengerechte Arbeit unter dem Zwang der Erhaltung der Wettbewerbsfähigkeit gesehen wird, so beweist das, daß Menschen instrumentalisiert werden: das schließt eine Verkehrung des Begriffs "Anthropozentrik" in sich.
- Anthropozentrik ist als Zukunftschance, als Weg zur Anpassung an den raschen wirtschaftlichen und sozialen Wandel zu begreifen. Unter dem Prinzip der Wirtschaftlichkeit stehend wird Anthropozentrik soziale Erneuerung fördern. Dem Ziel einer menschengerechten Arbeit wird man sich jeweils nur annähern können. Die Arbeitenden sollten "sozial qualifiziert" werden, eine fachliche Weiterqualifizierung ist nicht von allen Betroffenen gewünscht.
- Anthropozentrik bedeutet, den Menschen stärker als ganzen zu sehen. Beispielsweise sind die Selbstverwirklichung während der Freizeit und die Fremdbestimmung während der Arbeitszeit durch größere Autonomie am Arbeitsplatz einander näher zu bringen. Es kommt weniger auf das Glück des Einzelnen als auf seine Motivation bei der Arbeit (z.B. durch Stärkung seines Selbstwertgefühls) und auf ein Akzeptieren der Arbeitsziele an. Nachdem die großen technischen Revolutionen hinter uns liegen, bewegen wir uns auf neue soziale Ziele zu: die partizipative Gesellschaft, die sich von bisher gültigen Befehlsstrukturen abwendet.
- Relativ autonomes Handeln am Arbeitsplatz, etwa bezogen auf die Reihenfolge der Tätigkeiten oder auf die Auswahl der Mittel und Wege zur Aufgabelösung, bedeutet noch nicht eine Realisierung von Anthropozentrik. Selbstbestimmung während der Arbeit ist in einer hochindustrialisierten Wirtschaft für das Gros der Arbeitenden unerreichbar. Es wird auch in Zukunft qualifizierte und nichtqualifizierte Werkstätige geben, deswegen sind Lohnsteigerungen und Arbeitszeitverkürzungen näherliegende Ziele im Interesse der Menschen.
- Wir sehen uns konfrontiert mit politischen, ökonomischen und sozialen

Veränderungen in Mittel- und Osteuropa, denen wir begegnen müssen. Nicht erwarten dürfen wir, daß sich Osteuropa dem westlichen Wirtschaftssystem nahtlos einpaßt. In den hochindustrialisierten Ländern haben wir dem Rechnung zu tragen, daß ältere Arbeitnehmer in den Arbeitsprozeß eingegliedert werden oder bleiben müssen. In solcher Situation bieten anthropozentrische Produktionssysteme die Möglichkeit zu einer Umstrukturierung der Ost/West-Wirtschaft, die sowohl hoch entwickelte Kulturen als auch hoch entwickelte Technologien einplant. Ziel der neuen Entwicklungen ist es, solche Produktionssysteme zu schaffen, die die Menschen glücklicher machen.

Differenzierte Sichtweisen der Produktionsstrukturen sind zukünftig nötig. Das bedeutet z.B., daß industriekulturelle Bedingungen mitberücksichtigt werden, daß ökologische Kriterien, wie die Wiederverwertbarkeit von Produkten, einzubeziehen sind. Die Informationstechnik spielt eine entscheidende Rolle bei der Ermittlung gemeinsamer Interessen.

- Anthropozentrisches Denken hat zur Voraussetzung, daß nicht länger nach bisheriger Art von Unternehmern einzelwirtschaftlich gedacht wird, gesucht ist die Beachtung einer Wirtschaftlichkeit, die auf die Gesamt-Gesellschaft bezogen ist.

- Anthropozentrische Produktion wird in der EG in Ansätzen bereits verwirklicht, sie beinhaltet:

- Teamarbeit auf allen Betriebs-ebenen,
- dezentrale Vernetzung aller Betriebsgruppen, die an der Produktion direkt oder mittelbar beteiligt sind,
- Förderung informeller Kommunikationsprozesse innerhalb der Teams und nach außen, um, den Gruppen sowohl innere Stabilität als auch ausreichend Flexibilität zu geben.
- Außerdem: Arbeit sollte soweit möglich an den Lebensort des Arbeitenden transportiert werden, da die sozialen Folgeschäden (z.B. für die Familienzusammengehörigkeit) auf Dauer gesamtgesellschaftlich gesehen zu hoch sind.

- Beispiele für dezentralisierte Produktionssysteme liefert etwa der Automobilbau, dessen Zulieferindustrie eigenständige Entwicklungen in die Produktionsprozesse einbringt und die an Forschungsdiskussionen aktiv beteiligt ist. Die Vernetzung von Absatzabteilungen, Fertigungsabteilungen und Entwicklungsabteilungen in Japan und eine geglückte Verbindung von Elektronik und Mechanik haben japanischen Firmen zu Wettbewerbsvorteilen verholfen. Nicht-unternehmerische Interessenverbände haben z.B. der norditalienischen Textilindustrie das Überleben gesichert.

- Besondere menschliche Fähigkeiten in der Arbeitswelt sollten nicht durch hochentwickelte Technik nachgeahmt und ersetzt, sondern stattdessen durch sie unterstützt werden. In diesem Zusammenhang ist etwa auf werkstatorientierte Programmiermethoden hinzuweisen oder auf das kognitive Verhalten von Facharbeitern, denen für Veränderungen an Objekten eher die graphische als die sprachliche Darstellung zusagt.

## 5 Probleme mit anthropozentrischen Produktionssystemen

Der zweite Tag zeigte in Vorträgen und Panel-Diskussionen stärker die Probleme auf, die mit der Einführung anthropozentrischer Produktionssysteme in speziellen industriekulturellen Umgebungen verbunden sind. Der Vermutung vom Vortag, daß die APS besonders gut für die Überleitung der osteuropäischen Wirtschaftssysteme in eine wie immer geartete Marktwirtschaft nach westlichem Muster geeignet sei, stellte der Experte für Südeuropa entgegen, daß erst Analysen der Wirtschaft in den entsprechenden Ländern notwendig seien, bevor man daran gehen könne, individuell angepaßte APS zu installieren. Für Spanien erinnerte er an den wirtschaftlichen Umbruch der siebziger Jahre, die darauffolgende Schaffung neuer Arbeitsplätze in den achtziger Jahren, und meinte, eine Sensibilität für anthropozentrische Produktion überhaupt sei zunächst vorzubereiten. Probleme wurden anläßlich der Fragestellung offenbar, ob denn nun nur hochqualifiziert Ausgebildete einen Arbeitsplatz mit weitreichender Handlungsfreiheit beanspruchen dürften, und

Zweifel geäußert, ob Weiterbildung von älteren Arbeitnehmern, Behinderten und Frauen und "Nichtweiter-Qualifizierbaren" wohl angenommen werde und als fachliche - oder vielleicht für manche nur als soziale? - Weiterbildung erfolgen solle. Der Gedanke an eine Zwei-Klassen-Arbeiterschaft ließ sich da nicht leicht verdrängen. Durchaus wirklichkeitsnah blieb man sich bewußt, daß Ausbildung Geld kostet; die vorgegebene Klassifizierung, - bei der übrigens die ausländischen Arbeitnehmer fehlten -, zeigt, wie weit der Begriff "Mensch" changiert. Auch wurden Schwierigkeiten vorausgesehen, die in unterschiedlichen Erziehungs- und Ausbildungssystemen begründet sind, im geänderten Sozialverhalten junger Leute, und kurz: im "Wertewandel" überhaupt, durch den das individuelle Wohlergehen an die Spitze der Werteliste geraten ist. Wie weit ökologische Krisen das Wirtschaftsgefüge beeinflussen werden, läßt sich zur Zeit ebensowenig absehen, wie die Menge der Reglements, die benötigt werden, um die neuen Produktionssysteme zum Laufen zu bringen.

Das Panel mit seinen fünf Teilnehmern aus der U.D.S.S.R., den U.S.A., Japan, Australien und Europa war etwas anspruchsvoll als "intercontinental" eingeführt worden, die Afrikaner wurden anscheinend dabei nicht vermißt. (Eine Begründung für die Auswahl der Repräsentanten der Kontinente wurde nicht gegeben - vielleicht erschien sie den Veranstaltern überflüssig.) Immerhin war erstaunlich, wie substantiell die meisten der Länder-Analysen ausfielen und wie unverblümt die Berichterstatter redeten:

Für die Sowjet-Union wurde die Abkehr von der Planwirtschaft und die Konvertierung der Militär-Industrie in eine zivile Industrie als Wendung zur Anthropozentrik gedeutet. Beteiligung der Belegschaft am Firmenkapital wird angestrebt. Durch den Teil-Rückzug des Staates aus der Finanzierung der Akademie der Wissenschaften in Moskau ergibt sich zugleich die Notwendigkeit und die Möglichkeit für die Wissenschaftler über die Gründung von Consult-Firmen die Industrie zu beeinflussen.

Die strukturelle Krise der U .s.- Wirtschaft wurde auf mangelnde Anpassung an die Weltwirtschaft sowie auf fehlende nationale Kooperation zurückgeführt: derzeit sei es für U.S.-Firmen leichter, mit japanischen Firmen zu kooperieren als mit ihres- gleichen. Nach wie vor herrscht eine Technologie-Gläubigkeit, die Neu-Organisationen verhindert: die innovativen Projekte von General Motors haben zu keinen tief tiefgreifenden Restrukturierungen

in der Organisation geführt. Darüber hinaus ist der Einfluß des Pentagon in der Wirtschaft weitreichend bestimmend. Eine bedeutende Veränderung im Wirtschaftsleben der U .S.A. ist weitgehend unbemerkt dadurch eingetreten, daß sich während der Reagan-Regierung die Beteiligung am Familieneinkommen auf alle arbeitsfähigen Familienmitglieder ausgedehnt hat. Die menschliche Arbeitskraft wird von der Wirtschaft jedoch eher als Kostenfaktor denn als Ressource eingeschätzt. Den Kern für Erneuerungen könnte der Einsatz von APS bilden, wenn nicht tiefverwurzelte Verhaltensmuster - wie zu befürchten sei - die Übernahme vernünftiger Lösungen verhindern.

Australiens Wirtschaft ist z. Zt. gekennzeichnet durch den Import industrieller Produktion anstelle von Innovationen in den Produktionsprozessen. Eigenes Konzeptdesign spielt daher eine geringere Rolle als die Übernahme fremder Konzepte, wobei darüber hinaus noch Teilaspekte bei der Übernahme sträflich vernachlässigt werden. Der Einfluß der Zusammenarbeit mit Japan könnte hier zu Neuerungen führen, die allerdings in einer sozialen und demokratischen Ausrichtung erfolgen sollten. Das Studium schwedischer Modelle von Produktionssystemen gibt dazu die Anleitung. Daß die australische Regierung solche Veränderungen initiiert, zeigt sich daran, daß Firmensubventionen nur mit Auflagen hinsichtlich beruflicher Fortbildung der Arbeitnehmer und hinsichtlich der Kooperation mit andern Firmen vergeben werden.

Besonders die großen, international arbeitenden japanischen Firmen sind an anthropozentrischer Produktion interessiert, was sich auch an der Tatsache ablesen läßt, daß in letzter Zeit vier internationale Konferenzen zu diesem Thema in Japan organisiert worden sind. Allerdings ist eine Anpassung an APS wegen der starken internen Institutionalisierung der japanischen Wirtschaft schwer einzuführen. Es gibt neben CIM-Adaptionen allerdings Versuche, althergebrachte und neue Produktionssysteme ineinander zu integrieren. Japanische Suggestion Systems, in denen ein einzelner Arbeiter jährlich bis zu 200 Verbesserungsvorschläge in seiner Firma einreicht, wurden unwidersprochen nicht als Weg zur Partizipation anerkannt, da die Vorschläge in der Regel in einer überstarken Firmenhierarchie versandt und so faktisch ins Leere gehen.

Insgesamt ergab sich aus den Erörterungen des zweiten Tages, daß es vorzugsweise in Europa, allenfalls Ansätze zu anthropozentrischen Produktionssystemen gibt, die je

doch vielversprechend sind. Zugleich lag aber in diesem Zusammenhang das Problem offen zutage, daß nämlich nicht nur individuell divergierende Konzepte des in Frage stehenden Begriffs auftreten, sondern daß Wirtschaftssysteme so sehr kulturell, auch an Mentalität, Wertsetzungen und latente Prinzipien des Handeins gebunden und entsprechend unterschiedlich sind, daß Übertragungen von anthropozentrischen Produktionssystemen vielfältige Hindernisse im Wege stehen. Ob sich Anthropozentrik in Zufriedenheit, Glück, Arbeitsmotivation, Gemäßheit oder Selbstbestimmung des arbeitenden Menschen äußert, ist letztlich nicht das eigentliche Problem, gewinnen doch auch diese Begriffe Bedeutungen erst in kulturellen Kontexten. Wohl aber erwies sich der Begriff der Anthropozentrischen Produktionssysteme als ein Begriff von ganz anderer Qualität, weil er selbst als bloßes Etikett einen Dialog über den Zusammenhang von ökonomischen, sozialen und ökologisch angepaßten Neuerungen hervorrief, weil er die Ausarbeitung von praktikablen Modellen evoziert und insofern zu Recht als ein Arbeitsbegriff bezeichnet werden darf.

## 6 Folgerungen für die Computerlinguistik

So interessant es für Linguisten-Ohren sein mag, an der Bedeutungsentwicklung eines Begriffs realiter teilzuhaben, möchte ich jetzt davon absehen und weiterreichende Folgerungen erörtern: Im Rahmen der Europäischen Gemeinschaft besteht Interesse daran, kleine und mittelständische Unternehmen zu erhalten und die Produktion in Großunternehmen effektiver zu machen. Dafür gleichermaßen bieten sich, wie zu erfahren war, anthropozentrische Produktions- und Dienstleistungssysteme an. Für den Aufbau solcher Systeme wird

- (a) der Einsatz von hochentwickelten Technologien geplant,
- (b) in einem auf Dezentralisation und Partizipation ausgelegten Betrieb mit fachlich hochqualifizierten und/oder bezüglich ihres sozialen Handeins versierten, entscheidungsfähigen und an ihrem Platz verantwortlichen Mitarbeitern gerechnet.

Solche Produktions- und Dienstleistungssysteme werden zunehmend abhängiger vom eingesetzten Wissen, das bedeutet: es werden

immer mehr Informations- und Kommunikations-"Güter" zwischen Einzelpersonen und Organisationen ausgetauscht anstelle materieller realer Güter. Der Handel mit Informationen (gesetzt den Fall, diese Vorstellung ist überhaupt zutreffend und nicht vielmehr eine Täuschung über die Wirklichkeit) und die daraus resultierende Bewertungs- und Entscheidungsfähigkeit der Mitarbeiter erfordert nicht eine bloße Akkumulation von möglichst viel Wissen, sondern setzt voraus, daß entscheidungsrelevantes, situationsabhängiges Wissen für spezielle Problemstellungen zur Verfügung gestellt wird, wobei Relevanz und Situationsabhängigkeit von den Problemlösern selbst fallweise beurteilt werden.

Die Ausweitung sozialer Vorgänge am Arbeitsplatz und die Änderung von Kommunikationsstrukturen in den Betrieben eröffnen ein Forschungs- und Arbeitsgebiet, an dem Natur-, Organisations-, Kognitions- und Sozialwissenschaften, im weitesten Sinne, zugleich beteiligt sind. Danach ist vorauszusehen, daß sprachliches Handeln am Arbeitsplatz, das unter Zuhilfenahme hochentwickelter Informationstechnik erfolgt, als Untersuchungs- und Entwicklungsobjekt von diesen Wissenschaften vereint bearbeitet werden muß, um die Beziehungen zwischen Technik, Organisation und Mensch im einzelnen zu klären.

Erwartungen an die Computerlinguistik treten dabei in dreierlei Hinsicht auf, bezüglich

1. des Forschungsobjekts,
2. der Forschungsweise,
3. der Ausbildungsziele und Studieninhalte.

Zu 1.:

Sprachliches kommunikatives Handeln in der Arbeitswelt sollte mehr als bisher im Interesse der CL-Forschung liegen. Die für die Arbeitswelt weithin natürliche Form der, gesprochenen Sprache verdient die Beachtung, die ihr gebührt. Der Komplex von Wissensproduktion, Wissensrepräsentation und Textverstehen sollte nicht als Zutat aus der Künstlichen Intelligenz und der Psychologie des Denkens mißverstanden werden. Computer-Imitation menschlichen sozialen Handeins als Forschungsziel wäre zu ersetzen durch Computer-Unterstützung, wobei dieser Begriff durch Forschung und Praxis gemeinsam zu interpretieren ist.

Zu 2.:

Da das sprachliche kommunikative Handeln in soziales nonverbales Handeln und

menschliches Verhalten eingebettet ist, ist Zusammenarbeit mit Human- und Sozialwissenschaften, im engeren Sinne, erforderlich. Da es um Arbeitshandeln geht, ist Kooperation mit Naturwissenschaften wie Biologie, Physik, Chemie gefragt. Die Organisations- und die Technikwissenschaften tragen ihren Teil zu den gesuchten Problemlösungen bei. (Aus der Schwierigkeit, Teilfragen an Einzeldisziplinen zu delegieren, mag ersichtlich sein, welche Anforderungen an ein problemorientiertes wissenschaftliches Arbeiten im Forschungsteam zukünftig gestellt werden.)

Zu 3.:

Die Computerlinguistik-Lehre hat primär mit **allen** Wissenschaften zu tun, die sich mit menschlichen natürlichen und formalen Sprachen befassen, und sich mit deren Sichtweisen auseinanderzusetzen. Da "sprachliches kommunikatives Handeln am Arbeitsplatz" die Ausrichtung des Faches bestimmt, müssen Grundlagenwissen vermittelt und Beobachtungsfähigkeit geschult werden, besonders bezüglich argumentativer Strukturen, die sich in der partizipativen Arbeitswelt etwa als Erklärungen, als Überzeugenshandlungen, als Verhandeln und Vereinbaren manifestieren. Nimmt man den kommunikativen Charakter sprachlichen Handelns ernst, verdienen die verschiedenen Arten der Verständnissicherung nicht nur theoretisch erörtert, sondern bewußt praktisch erprobt zu werden. So etwa könnte sich eine Computerlinguistik den auf der Tagung anvisierten Erwartungen stellen, will sie eine an den gesellschaftlichen Veränderungen erklärtermaßen mitwirkende Wissenschaft sein.

PS.: Wie sehr die Referenten bemüht waren, ihrem Gegenstand gerecht zu werden, ließ sich mindestens in zwei Situationen sehen: Ein Referent hatte Skizzen von arbeitenden Personen in die üblichen Pfeildiagramme eingezeichnet, um darauf hinzuweisen, daß er von Menschen rede. Ein anderer benutzte eine industriell bereits vorgefertigte schematisierte Hand mit ausgestrecktem Zeigefinger anstelle des sonst üblichen Papp-Pointers für Folien oder des Lichtpfeils,- nicht ohne diese Wahl ausdrücklich mit dem Anthropozentrik-Anliegen zu begründen.

## Information Retrieval

Christa Wormser-Hacker  
Universität Regensburg

### Tagungsbericht

Workshop "Information Retrieval" Gesellschaft für Informatik (GI) - GMD-IPSI 24. und 25. Juni 1991.

Am 24. und 25. Juni 1991 fand bei der GMD in Darmstadt ein Workshop mit dem generellen Thema "Information Retrieval" statt. Das große Interesse an diesem Workshop spiegelte sich in den nahezu 100 Teilnehmern wider, die sich zum größten Teil aus Wissenschaftlern und Vertretern der Industrie zusammensetzten. Die Teilnehmer kamen aus Deutschland, Österreich und der Schweiz. Die beiden eingeladenen Vortragenden (Prof. Keith van Rijsbergen und Prof. Bruce Croft) erweiterten dieses Spektrum auf Großbritannien und die USA.

Veranstalter des Workshops waren die Gesellschaft für Informatik und die Gesellschaft für Mathematik und Datenverarbeitung/Institut für Integrierte Publikations- und Informationssysteme GMD-IPSI, die auch ihre Infrastruktur zur Verfügung stellte. Der Workshop sollte den Start bilden für die neue Fachgruppe "Information Retrieval" innerhalb der GI.

Die zweitägige Veranstaltung gliederte sich in sieben Sitzungen mit unterschiedlichen inhaltlichen Schwerpunkten, in welchen insgesamt 15 Vorträge gehalten wurden. Zum Teil ergänzten entsprechende Systemvorführungen die Referate.

Höhepunkte des Workshops waren sicherlich die bei den eingeladenen Vorträge. Keith van Rijsbergen (Professor an der University of Glasgow) sprach über "Information Retrieval and its Logic: Inference versus Relevance". Ausgehend von der Tatsache, Information Retrieval mittlerweile nicht nur mit Dokumenten beschäftigt, die in Form von Texten vorliegen, sondern auch Bilder, Videos, Objekte aus der Büroumgebung etc.

(Hypermedia!) gespeichert und wiederaufgefunden werden sollen, müssen auch die Grundlagen des Information Retrievals, d.h. die Modelle, auf ihre Anwendbarkeit auf die multimedialen Anforderungen überprüft werden. Van Rijsbergen sieht bei der Verschmelzung eines Hypermedia-Systems mit dem probabilistischen IR-Modell das Problem darin, daß die Unsicherheit der Information durch das Hinzufügen von Information aus der Umgebung oder einem Thesaurus noch verstärkt wird. Van Rijsbergen zeigte, indem er auf das "logical uncertainty principle" zurückgriff, daß Information Retrieval durchaus als eine Form von Inferenz angesehen werden kann.

Bruce Croft (Professor an der University of Massachusetts, Amherst), der zweite eingeladene Redner, hielt einen Vortrag mit dem Titel "Retrieval from Large Databases". Zunächst stellte er heraus, daß sich Information Retrieval in den USA nach wie vor großen Interesses erfreut, vor allem aufgrund der neuen Größenordnungen im Bereich der Dokumentanzahl, auf seiten der Anfragen und last not least auch im Hinblick auf das individuelle Dokument ("What is large?"). Ausgehend von den amerikanischen IR-Forschungsaktivitäten der letzten Jahrzehnte zeigte er, inwieweit die erzielten Ergebnisse vereinbar seien mit den Anforderungen von Datenbanken, die den neuen Größenordnungen entsprechen. Er warf Fragen auf nach der Übertragbarkeit von Experimenten in bezug auf die Beurteilung von Effektivität und Effizienz bestimmter Algorithmen oder Funktionen, die auf der Grundlage kleiner Kollektionen bzgl. der einbezogenen Dokumente und Anfragen erzielt worden waren. Bruce Croft ging auf drei Projekte ein, an welchen er beteiligt ist (TIPSTER, WESTLAW und CONSTRUE), die beitragen sollen, dieses Defizit zu beheben. Durchaus

optimistisch für die Forschung seine abschließende Formel: "Large Databases Present Large Challenges" !

Die 13 regulären Vorträge des Workshops kreisten um folgende Schwerpunkte: Entwurf von IR-Systemen und -Komponenten, Repräsentation und Deskribierung im IR-Bereich, Evaluierung und Messung, Realisierung von IR-Systemen aus verschiedenen Anwendungsbereichen und Hypertext und IR.

Im Beitrag von H.P. Frei und S. Meienberg ging es um ein Verfahren, das automatisch Boolesche Anfragen aus Anfragen mit gewichteten Suchtermen generiert und die Antwortmenge durch Rankingverfahren ordnet, um so das Retrieval effizienter zu gestalten. Die Vorträge von K. Tzeras und U. Pfeifer beschäftigten sich mit automatischer Indexierung. K. Tzeras referierte über die Abschätzung des Aufwands in bezug auf die Entwicklung eines Indexierungswörterbuchs. U. Pfeifer stellte neue Ansätze vor, die Experimenten zufolge erfolgversprechende Wege zu besseren automatischen Indexierungsfunktionen, die über die Deskriptorzuteilung entscheiden, darstellen. Beide Vorträge sind im Zusammenhang mit dem Darmstädter Indexierungsansatz zu sehen. Anforderungen an das **IR** aus der Praxis stellten H. Amstutz und B. Holländer/Thönssen durch ihren Vortrag "Elektronische Ablage und Archivierung auf der Basis eines Database Management Information Retrieval Systems" dar. K. Englmeier präsentierte mit dem wissensbasierten System WIDAB (Wirtschaftsinformationen und Datenbank) ein System zur Informationsorganisation zur effizienten Nutzung von Wirtschaftsinformation, die bei statistischen Ämtern und Wirtschaftsforschungsinstituten gehalten wird. Ein weiterer Vortrag aus der anwendungsorientierten Richtung von M. Münzenberger stellte ein Beispiel für eine IR-Anwendung im Rahmen des "Forensischen Informationssystems Handschriften" (entwickelt beim Bundeskriminalamt) vor. Durch Kombination von Mustererkennungs- und Clustering-Verfahren soll eine effektive Recherche ermöglicht werden. Der Beitrag von W. Augsburg, H.K. Rieder und J. Schwab kam aus dem Bereich des Faktenretrievals. Hier ging es um ein System, das Unternehmenszahlen zu bewertbaren Kennzahlen umformt und eine automatische Vorbewertung vornimmt. Mit dem Berliner ATLAS-Projekt befaßten sich die Vorträge von J. Willenborg und T. Muhr.

Im ersten Referat ging es um einen Ansatz zur Unterstützung bei der Thesauruserstellung und -pflege, im zweiten um ein

Interpretations- Unterstützungs-System, das die Arbeit mit Texten z.B. im Hinblick auf Vergleich, Kommentierung, Kategorisierung etc. unterstützt. H.P. Frei, S. Meienberg und P. Schäuble stellten eine neues Effektivitätsmaß vor, das die Nachteile der klassischen Maße recall und precision kompensieren soll. Drei Vorträgen waren der Koppelung von Hypertext und IR gewidmet. J. Herrmann und D. Meiser stellten die Realisierung eines Thesaurus durch Hypertextverknüpfungen vor. D. Merkl, A. Min Tjoa und S. Vieweg setzten ein Hypertext-System ein zur Navigation in juristischen Dokumenten. Um HyperQuery, ein Anfragesystem mit graphischer Benutzeroberfläche am Beispiel eines Reiseauskunftsystems ging es im letzten Beitrag von J. Herczeg, H. Hohl und M. Ressel.

Die Vorträge werden in einem Tagungsband publiziert, der im Herbst bei Springer erscheinen wird (Fuhr, N. (Hrsg.) (1991), Information Retrieval. Proceedings des Workshops "Information Retrieval", 24.-25.6.1991 in Darmstadt).

Hinzuweisen ist noch auf die Gründung der neuen Fachgruppe "Information Retrieval" innerhalb der GI, die sich an Forscher, Entwickler und Anwender von Informationssystemen richtet. Dieser Workshop ist als erste Veranstaltung dieser neuen Fachgruppe zu sehen, der in regelmäßigen Abständen weitere Tagungen folgen sollen. Themenschwerpunkte liegen u.a. in folgenden Bereichen: Schnittstellengestaltung, Information System Management, Retrievalmodelle, Evaluierung von Retrievalverfahren usw.

Kontaktadresse für Interessierte: Prof. Dr.

Norbert Fuhr Lehrstuhl Informatik VI Universität Dortmund W-4600 Dortmund 50

## Notizen zur Herbstschule CL in Trier 1990

Christian Lieske  
Universität Koblenz-Landau

### Überblick

Bei Tagungen und Konferenzen ist klar, worum es geht und was am Ende steht: Es sollen neue Erkenntnisse präsentiert bzw. gesichtet werden. Am Schluß läßt sich meist vieles des Vorgefallenen als eine Sammlung der Beiträge in Buchform mit nach Hause nehmen. Welche Ansprüche werden jedoch an den Besuch einer Schule, genauer gesagt einer "Herbstschule Computerlinguistik", gestellt? Welche Eindrücke und Ergebnisse bleiben zurück? Der folgende Text beantwortet diese Fragen aus der Sicht eines Teilnehmers an der letztjährigen Veranstaltung, die vom 10. bis 13. Oktober 1990 an der Universität Trier stattfand.

dem Ausland angereisten Teilnehmer, die als Gesamtheit in einen kleineren Hörsaal paßte, war ideal. Sie war so groß, daß die Mensa jeden Tag mit anderen Personen aufgesucht werden konnte, und zugleich so übersichtlich, daß Kontakte auch außerhalb des Unterrichts leicht entstehen konnten. Nicht eindeutig zu bewerten ist allerdings der breitgefächerte Hintergrund der Teilnehmer, von denen die meisten noch studierten, und die unter anderem aus den Fachbereichen Linguistik, Informatik, Phonetik, Germanistik, Philologie und Übersetzungswissenschaft stammten. Dieses weite Spektrum ermöglichte zwar einerseits interessante Blicke über den Grenzen des eigenen Faches, wirkte sich aber andererseits nachteilig auf die Kurse aus, denn ein gemeinsames Niveau ließ sich nur schwer finden.

### Erwartungen

Nach dreieinhalb Tagen Schule kann nicht damit gerechnet werden, souverän die behandelten Themenkreise im Griff zu haben. Zumindest bei solch umfassenden Themen wie "Grammatik-Modelle", "Praktische Lexikographie", "Maschinelle Übersetzung", "Quantitative Linguistik", "Software Ergonomie" sowie "KI-Programmierung und Wissensrepräsentation" ist eine derartige Erwartung nicht angebracht. Ein- und Überblicke dürfen hingegen erwartet werden, ebenso Tips für ein weiteres Eindringen in die berührten Gebiete. Auch Anregungen für eigene Arbeit in einem der Bereiche im Rahmen einer Studien-, Magister- oder Diplomarbeit sind als legitime Hoffnungen zu nennen. Zudem sollte die Möglichkeit gegeben sein, Kontakte zu knüpfen und fachfremde Einstellungen gegenüber der Computerlinguistik kennenzulernen.

### Kurse

Schon während des ersten Schulungstags wurde indirekt der längliche Name des Veranstalters in Erinnerung gebracht. Durch die Betonung der Praxis und nahezu vollständige Ausklammerung theoretischer Hintergründe ergab sich ein Übergewicht von rein mechanistischer Verarbeitung von Sprache gegenüber linguistisch-theoretischer Durchleuchtung von sprachlichen Strukturen. Diesem Eindruck sowie allen anderen Bemerkungen liegen übrigens nur Erfahrungen aus den Kursen "Grammatik-Modelle", "Praktische Lexikographie" und "Maschinelle Übersetzung" zugrunde.

Was das bereits erwähnte Problem betraf, das richtige Kursniveau zu finden, so hatten der mit der maschinellen Übersetzung betraute Johann Haller sowie die für die praktische Lexikographie zuständigen Raimund Drewek und Andreas Lang mehr Glück als der für Grammatik-Modelle verantwortliche Hans Ulrich Block. Obwohl auch letzterer sich zu Beginn seines Kurses nach dem Wissensstand seiner Klientel erkundigte, ging rasch durch die Inhomogenität der Gruppe der gemeinsame rote Faden verloren. Als al

### Teilnehmer

Die Zahl der aus dem gesamten Bundesgebiet und zu geringem Prozentsatz auch aus

len drei Kursen gemeinsamer Schwachpunkt wurde die Fixierung auf bestimmte Software-Werkzeuge bemängelt, ein Punkt der mit der oben erwähnten Vernachlässigung der Theorie zusammenfällt. Die in diesem Zusammenhang vorgebrachte Argumentation läßt sich als Frage formulieren: Was nutzt die Einarbeitung in ein Programm, wenn dieses am Studienort nicht verwendbar ist? Weiterhin wurde bemängelt, daß mit drei verschiedenen Editoren gearbeitet werden mußte. In Reinhardts Köhlers Kurs über quantitative Linguistik gab es Kontroversen über den breiten Raum, den er wissenschaftstheoretischen Diskussionen einräumte. Einige waren davon angetan Argumente für die Rechtfertigung des eigenen Tuns an die Hand geliefert zu bekommen, während andere handfeste Inhalte vermißten.

## Organisation

Die Räume in Trier waren gut gewählt, denn die Wege zwischen den einzelnen Hörsälen, den Rechnerräumen und der Mensa waren kurz. Ebenfalls guten Eindruck hinterließ das Informationsmaterial, in dem sich Hinweise auf Literatur, genaue Kursinhalte und Unterbringungsmöglichkeiten fanden. Es fehlte allerdings ein Gebäudeplan der Universität. Hilfreich wäre ebenfalls eine Liste der vorangemeldeten Teilnehmer gewesen, die zur privaten Organisation von Mitfahrgelegenheiten hätte genutzt werden können.

Der Veranstaltungsplan erwies sich als zu dicht gepackt. Es blieb kaum Zeit für Gespräche und einzelne Teilnehmer waren zeitlich so belastet, daß sie der Weinprobe und der Podiumsdiskussion, die gemeinsam das Rahmenprogramm bildeten, fernblieben. Eine Belastung war auch die nicht korrekte Verbindung des Universitätsrechnernetzes mit dem für die maschinelle Übersetzung angeschlossenen Sun- Rechner durch die es Probleme beim Editieren gab.

## Fazit

Die Herbstschule ist Studenten zu empfehlen. Sie können die Bedeutung der aus ihrem Studium bekannten Theorie für die Praxis kennenlernen und somit Motivation tanken. Außerdem ist ein Einblick in die in der Praxis vorkommenden Problemstellungen möglich. Dies eröffnet die Chance, im eigenen Studium eine entsprechende Orientierung vorzunehmen, was ein Einwirken auf die Gestaltung der Lehrpläne einschließt.

Der Kontakt zu Studenten aus anderen Fachbereichen bewirkt ein besseres Verständnis derselben. Außerdem kann es durch diesen Kontakt zum Aufbau fächerübergreifender Projekte kommen und die Scheu vor dem Wechsel an eine andere Hochschule bzw. in ein anderes Fachgebiet abgebaut werden. Der Blick über den Tellerrand bewahrt ebenfalls davor, Energien in Arbeit zu investieren, die außerhalb des eigenen Faches schon geleistet worden ist.

## Anregungen

Die Dozenten sollten am Tagungsbrett über mögliche Praktika, sowie über Themen für Magister-, Studien- und Diplomarbeiten Informationen Geben, denn ein Großteil der Teilnehmer war danach auf der Suche. Dem kamen nicht alle Dozenten entgegen. Auch die Teilnehmer sollten Gelegenheit bekommen, Bedürfnisse hinsichtlich einer Zusammenarbeit zu formulieren. So könnten beispielsweise ein Philologe und ein Informatiker für die Entwicklung eines Transkriptionsprogramms zusammenfinden. Zu wenig Raum für persönliche Gespräche und fehlende Informationen über Interessen der anderen Teilnehmer hat ein solches Zusammenfinden in Trier verhindert.

Selbstverständlich sind die organisatorischen Schwachstellen zu beheben und zum Beispiel eine OEI, eine "one editor initiative", ins Leben zu rufen.

## The Joint Project WISBER

This is the complete list of all reports and memos. The papers are available and can be ordered free of charge from the above address:

Prof. Dr. W. Wahlster  
Verbundprojekt WISBER  
Universität des Saarlandes  
FR. 10 Informatik IV  
Im Stadtwald 15  
D-6600 Saarbrücken 11  
Electronic mail address:

wahlster@sbsvax.uucp@germany.csnet

- BI: H.-U. Block, M. Gehrke, H. Haugeneder, R. Hunze, Neuere Grammatiktheorien und Grammatikformalismen (Siemens)
- B2: R. Busche, S. op de Hipt, M.-J. Schachter-Radig, Entwurf eines Erhebungsschemas für Geldanlage (SCS)
- B3: H. Rösner, Generierung von Erklärungen aus formalen Wissensrepräsentationen (Nixdorf) in: LDV-Forum, Band 4, Nummer 1, Juni 1986, pp. 3-19
- B4: H.-U. Block, R. Hunze, Incremental Construction of c- and f- Structure in an LFG-Parser (Siemens) in: Proc. of the 11th Intern. Conference on Computational Linguistics, COLING '86, Bonn, pp. 490-493
- BS: H.-U. Block, H. Haugeneder, The Treatment of Movement Rules in an LFG-Parser (SIEMENS) in: Proc. of the 11th Intern. Conference on Computational Linguistics, COLING '86, Bonn, pp. 482-486
- B6: M. Gehrke, H.-U. Block, Morpheme-based Lexical Analysis (Siemens)
- B7: H.-U. Block, M. Gehrke, H. Haugeneder, R. Hunze, Probleme der Wissensrepräsentation in Beratungssystemen (Siemens)
- B8: H. Bergmann, M. Gerlach, W. Hoepfner, H. Marburger, Beratung und natürlichsprachlicher Dialog - eine Evaluation von Systemen der Künstlichen Intelligenz (Universität Hamburg)
- B9: M. Fliegner, M.-J. Schachter-Radig, Form der Ergebnisse der Wissensakquisition in WISBER (SCS)
- B10: R. Block, Lexical Functional Grammar and Natural Language Generation (Universität Hamburg)
- B11: R. Hunze, H.-U. Block, A Two Step Reference Problem Solver (Siemens)
- B12: H.-G. Siedka-Bauer, Repräsentation des Fachwissens einer komplexen Domäne in einer KL-ONE-artigen Repräsentationssprache (SCS)
- B13: S. op de Hipt, R. Busche, Analyse und Zusammenfassung der empirischen Studien, Erfahrungen bei der Erhebung von Wissen in einer komplexen Domäne (SCS)
- B14: H. Bergmann, M. Fliegner, M. Gerlach, H. Marburger, M. Poesio, IRS - The Internal Representation Language (Universität Hamburg)
- B15:** H. Bergmann, M. Gerlach, Semantisch-pragmatische Verarbeitung von Äußerungen im natürlichsprachlichen Beratungssystem WISBER (Universität Hamburg) in: W. Brauer, W. Wahlster (Eds.), Wissensbasierte Systeme - GI-Kongreß 1987, Springer Verlag, Berlin 1987, pp. 318-327
- B16: S. Schachtl, The Problem of Overgeneration in parsing Processes and the Aid of Linguistic Generalizations (Siemens AG)

- B11:** R. Frederking, M. Gehrke, Resolving Anaphoric References in a DRT-based Dialogue System: Part2: Focus and Taxonomie Inference (Siemens AG)
- B18:** J. Meßing, I. Liermann, M.-J. Schachter-Radig, Handlungsschemata in Beratungsdialogen, Am Gesprächsgegenstand orientierte Dialoganalyse (SCS)
- B19:** M. Sprenger, Konzeption des Partnermodells in WISBER (Universität Hamburg)
- B20:** R. Block, Papers on Reference and Knowledge Representation (Universität Hamburg)
- B21:** M. Poesio, Temporal Reasoning in a Hybrid System (Universität Hamburg)
- B22:** E. Werner, Toward a Theory of Communication and Cooperation for Multiagent Planning (Universität Hamburg) in: Theoretical Aspects of Reasoning about Knowledge, Proceedings of the 1988 Conference, Morgan Kaufman Publishers, 1988
- B23:** M. Gerlach, M. Sprenger, Semantic Interpretation of Pragmatic Clues: Connectives, Modal Verbs, and Indirect Speech Acts (Universität Hamburg) in: Proc. of the 12th Intern. Conference on Computation Linguistics, COLING '88, Budapest
- B24:** R. Hunze, Resolving Anaphoric References in a DRT-based Dialogue System: Part1: The Formal Discourse Model (Siemens)
- B25:** K. Jablonski, A. Rau, J. Ritzke, Anwendung einer logischen Grammatik zur Generierung deutscher Texte (Nixdorf) in: Proc. des Wiener Workshops Wissensbasierte Sprachverarbeitung (WWWS 88) der ÖGAI, Wien
- B26:** M. Poesio, Toward a Hybrid Representation of Time (Universität Hamburg) in: Proc. of the ECAI 88, München, pp. 247-252
- B21:** K. Jablonski, a. Rau, J. Ritzke, NUGGET - Ein DCG-basiertes Textgenerierungssystem (Nixdorf Computer AG)
- B28:** M. Sprenger, M. Gerlach, Expectations and Propositional Attitudes. - Pragmatic Issues in WISBER (Uni Hamburg)
- B29:** M. Fliegner, HOKUSPOKUS - Verwendung terminologischen Wissens bei der Analyse von Quantorenkopos und Distributivität (Universität Hamburg)
- B30:** H. Horacek, H. Bergmann, R. Block, M. Fliegner, M. Gerlach, M. Poesio, M. Sprenger, From Meaning to Meaning: A Walk Through WISBER's Semantic-Pragmatic Processing (Universität Hamburg)
- B31:** H.. Horacek, C. Pyka, Facets of Knowledge About Natural Language Syntax: Representation and Use in Parsing and Generation (Universität Hamburg)
- B32:** H. Horacek, C. Pyka, Towards Bridging Two Levels of Representation- Linking the Syntactic Functional and Object-Oriented Paradigms (Universität Hamburg)
- B33:** H. Horacek, C. Pyka, Anwendbarkeit von Unifikationsgrammatiken für effizientes Generieren (Universität Hamburg)
- B34:** E. Werner, A Formal Computational Semantics and Pragmatics of Speech Acts (Universität Hamburg)
- B35:** E. Werner, Social Intentions (Universität Hamburg)
- B36:** E. Werner, Socializing Robots: A Theory of Communication and Social Structure for Distributed Artificial Intelligence (Universität Hamburg)
- B31:** E. Werner, Information, Intention and Ability: Outline of a Unified Theory (Universität Hamburg)
- B38:** R. Hunze, The Treatment of Ambiguity within a DRT-Framework (Siemens)
- B39:** M. Gehrke, Neues aus der Beziehungskiste: Begriffsdefinitionen und Referenzidentifikation definiter Beschreibungen (Siemens)
- B40:** S. Schachtl, Der Akkusativ in den Medialkonstruktionen des Deutschen, erscheint in: Fanselow, G., Felix, S. (Hrsg.), Strukturen und Merkmale grammatischer Kategorien, Studien zur deutschen Grammatik, Narr, Tübingen 1989 (Siemens)
- B41:** S. Schachtl, Morphological Case and Abstract Case: Evidence from the German Genitive Construction (Siemens)

- B42: M. Poesio, Dialog-Oriented ABoxing (Universität Hamburg)
- B43: J. Meßing, M. Mankel, Wissensorganisation zur Dialogsteuerung (SCS)
- B44: E. Werner, Two Ways of Representing Knowledge (Universität Hamburg)
- B45: R. Block, Relation Grammar and Natural Language Generation (Universität Hamburg)
- B46: R. Block, Generating Referential Expressions (Universität Hamburg)
- B47: (not available)
- B48: E. Werner, The Modal Logic of Games (Universität Hamburg)
- B49: E. Werner, Semantics for Tensed Modal Logic (Universität Hamburg)
- B50: E. Werner, Cooperating Agents: A Unified Theory of Communication and Social Structure (Universität Hamburg)
- B51: E. Werner, Communication in WISBER (Universität Hamburg)
- MI: S. Savory, Untersuchung von Benutzerpräferenzen und daraus resultierende Spezifikation von Gestaltungsregeln für die Ausgabe von natürlichsprachlichen Texten am Bildschirm mit einer prototypischen Implementierung in TWAICE (Nixdorf)
- M2: R. Busche, M.-J. Schachter-Radig, Dialogführung im simulierten Szenario (SCS)
- M3: R. Heckl, M.-J. Schachter-Radig, Eingrenzung des Sachgebietes innerhalb der Geldanlage (SCS)
- M4: S. op de Hipt, Nutzerkreisfestlegung (SCS)
- M5: H. Rösner, Besuchsbericht Q&A (Nixdorf) (VERTRAULICH)
- M6: S. Saovry, Beurteilung von LeCourtier (Nixdorf) (VERTRAULICH)
- M7: M. Gerlach, WISBER Systemarchitektur - ein Entwurf (Universität Hamburg)
- M8: S. Savory, FENRIS: Eine Wissensdarstellungssprache für WISBER (einschließlich Quellcode) (Nixdorf) (VERTRAULICH)
- M9: K. Jablonski, Generierung konjugierter Verben mit TwAIce. Zur Erklärungs- und Instruktionsfähigkeit (Nixdorf)
- M10: M. Fliegner, Fragen der sprachpragmatischen Analyse am Beispiel der simulierten Beratungsdialoge (SCS)
- M11: H. Bergmann, M. Gerlach, QUIRK - Implementierung einer TBox zur Repräsentation begrifflichen Wissens (Universität Hamburg)
- M12: K. Jablonski, A. Rau, J. Ritzke, Konzeption und Architektur des taktischen Textgenerierungssystems NUGGET (Nixdorf)
- M13: H.-U. Block, H. Haugeneder, An Efficiency-Oriented LFG Parser (Siemens)
- M14: H.-G. Siedka-Bauer, Repräsentation des Fachwissens der Domäne Geldanlage in der KL-ONE-artigen Repräsentationssprache QUIRK (SCS)
- M15: R. Busche, S. op de Hipt, I. Liermann, J. Meßing, M.-J. Schachter-Radig, Ein Modell der Dialoganalyse, Handlungsschemata in Beratungsdialogen (SCS)
- M16: A. Winizki, INTERCOM - ein System zur Übersetzung von Interlisp-D nach CommonLisp (Universität Hamburg)
- M17: M. Gerlach, BNF - a Tool for Processing Formally Defined Syntactic Structures (Universität Hamburg)
- M18: M. Sprenger, Interpretation von Modalverben zur Konstruktion von Partnermodelleinträgen (Universität Hamburg)
- M19: I. Liermann, J. Meßing, Dialogstrukturen (SCS)
- M20: R. Busche, M.-J. Schachter-Radig, INF: Implementation eines Inferenzmechanismus - Spezifikation und Architektur des Regelinterpreters (SCS)
- M21: M. Treder, J. Pischke, B. Behr, Funktionale Beschreibung des Regelinterpreters (SCS)
- M22: M. Poesio, The QUARK Reference Manual (Universität Hamburg)
- M23: K. Lehner, Erstellung einer Lehrkomponente für TWAICE- Wissensbanken (Nixdorf)

- M24: S. Lowe, Detecting User Misconceptions - Approaches in WISBER (Universität Hamburg)
- M25: M. Gerlach, H. Horacek, Dialog Control in a Natural Language System (Universität Hamburg)
- M26: H. Horacek, Towards Principles of Ontology (Universität Hamburg)
- M27: R. Frederking, The WISBER Common Lisp i.s. Facility (FOR- macro) (Siemens)
- M28: A. Klär, M. Mankel, J. Meßing, A. Schnoor, WISBER- INF, Implementation der Inferenzkomponente (SCS)
- M29: H. Horacek, The Architecture of a Generation Component in a Complete Natural Language Dialog System (Universität Hamburg)
- M30: H. Bergmann, FTRANSLATE: Terminologische Transformationen in Analyse und Generierung (Universität Hamburg)
- M31: M. Gerlach, Implementierung der Dialogkontrolle (Universität Hamburg)
- M32: R. Sprengel, Funktionalität von QUARK (Universität Hamburg)
- M33: E. Werner, Money Talk: Hierarchical Consultation Dialogue (Universität Hamburg)
- M34: E. Werner, Temporal Modal Reasoning in Games (Universität Hamburg)
- M35: A. Winizki, M. Gerlach, Portierung der WISBER- Wissensrepräsentationswerkzeuge von Interlisp nach CommonLisp (Universität Hamburg)
- M36: H. Horacek, M. Gerlach, Goal Oriented Dialog Control in a Natural

Der neue

## **Studienführer LDV/CL für die deutschen Universitäten**

(Saarbrücken 1991)

informiert auf 95 Seiten mit gesicherten Auskünften. über  
Magister- und Diplom-Studiengänge LDV/CL  
Weiterbildungsstudiengänge LDV/CL und  
Sprach- und Textverarbeitung für Philologen  
Nebenfachstudienmöglichkeiten in LDV/CL.

Trotz großer Unterschiede zwischen den Studiengängen werden die Informationen in systematischer Aufbereitung präsentiert, um Übersichtlichkeit zu gewährleisten. Wo die für LDV/CL notwendigen Vorfelder Linguistik und Informatik im Studiengang bereits eingeplant sind, bzw. wo sie angeordnet werden können, ist in den entsprechenden Einzelfällen angegeben. Auch über vorgeschriebene oder mögliche Fächerkombinationen an den jeweiligen Universitäten gibt der Studienführer Auskunft.

Er ist erhältlich beim:

IAI, Martin-Luther-Str. 14  
6600 SAARBRÜCKEN 3

Bitte senden Sie zusammen mit Ihrer Bestellung einen frankierten (1,80 DM), mit der Aufschrift "Drucksache" und Ihrer Adresse versehenen DIN A5-Umschlag ein, und fügen Sie Briefmarken im Wert von 10,- DM bei Sammelbestellungen richten Sie bitte auch an das IAI.

# Veranstaltungen

## Veranstaltungskalender

- 19.11.-21.11.91 in Helsinki Congress on Cultural Identity and Integration in the New Media World.  
Auskunft: Marianne Fillilä, University of Industrial Arts, Centre for Advanced Studies, Hämeentie 135G, SF-00560 Helsinki.
- 21.11.-23.11.91 in Hempstead  
Conference on Money: Lure, Lore and Liquidity. Themen: Artistic Designs of Money, Money and Culture, Psychology of Money.  
Auskunft: Natalie Datlof, Hofstra Cultural Center, Hofstra University, Hempstead, NY 11550, U.S.A.
- 25.11.-27.11.91 in Bad Homburg  
Tagung über die Transkription der Kommunikation zwischen Gehörgeschädigten.  
Auskunft: Gudula List, Seminar für Heilpädagogische Psychologie, Universität zu Köln, Richard-Wagner-Straße 39, W-5000 Köln 1.
- 9.12.-12.91 in San Marino Workshop on Semantic Theory.  
Auskunft: Università di San Marino, Segretaria, Contrada delle Mura, I-47031 San Marino.
- 12.12.-14.12.91 in Berlin Internationales Bopp-Symposium. Auskunft: Tomas Milosch, Bopp-Symposium, Humboldt-Universität, Postfach 12 97, W-1000 Berlin.
- 31.01.1992 - **Göttingen**  
3. Symposium des GAWS Wissensbasierte Systeme in der Betriebswirtschaft - Anwendung und Integration mit Hypermedia Termine: Beiträge erbeten bis 15.11.1991 Veranstalter: Göttinger Arbeitskreis für Wissensbasierte Systeme(GAWS).
- Info: Dr. Ronald Bogaschewsky, Abteilung Unternehmensplanung der Universität Göttingen, Platz der Göttinger Sieben 3, W-3400 Göttingen
- Februar 1992 in Köln**  
Fachgespräch über Gehörlosengemeinschaften und Gebärdensprachen.  
Auskunft: Gudula List, Heilpädagogische Psychologie und Psychiatrie, Universität zu Köln, Richard-Wagner-Straße 39, W-5000 Köln 1.
- 24.02.-26.02.1992, Ulm Workshop  
Informationssysteme und Künstliche Intelligenz-Modellierung Veranstalter: GI-FG 1.1.4(Wissensrepräsentation), GI-FG 2.5.1 (Datenbanksysteme), GI-FG 2.5.2 (EMISA)  
Info: Prof. Dr. Rudi Studer, Universität Karlsruhe, Institut für Angewandte Informatik und Formale Beschreibungsverfahren,  
Postfach: 6980, W-7500 Karlsruhe
- 25.2.-28.2.92 in Bremen Tagung über sprachliche Universalien.  
Auskunft: Hartmut Günther, Albersallee 89, W-4190Kleve.
- März 1992 Stanford Univ, US AAAI Spring Symposium Series  
There will be 5 - 10 symposia, all with limited attendance. Info: Peter F. Patel-Schneider, AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974, US pfps@research.att.com
- 02.03.-06.03.1992 Monterey, CA US  
CAIA'92 IEEE Conf on AI for Applications  
Two kind of papers will be presented: on case-studies of knowledgebased applications and on AI techniques and principles that underlie knowledge based systems. Info: IEEE Computer Society, 1730 Massachusetts Avenue NW,

- Washington, DC 20036-1903. phone: +1 202371 1013
- 22.03.-25.03.1992, Las Vegas USA Symposium on Document Analysis and Information Retrieval  
Info: William S. Brogan, Dept. of Electrical Engineering, University of Nevada, Las Vegas, 4505 Maryland Parkway, Las Vegas, NV 89154,  
Email: eewlb@cs.univ.edu
- 23.3.-27.3.92 in Duisburg Symposium über interkulturelle Kommunikation. Auskunft: Rene Dirven, Universität Duisburg, Lotharstraße 65, W-4100 Duisburg.
- 24.3.-26.3.92 in Mannheim Jahrestagung des Instituts für Deutsche Sprache. Thema: Deutsch als Verkehrssprache in Europa. Auskunft: Gerhard Stickel, Institut für Deutsche Sprache, Friedrich-Karl-Straße 12, W-6800 Mannheim.
- 1.4.-3.4.92 in Trient Tagung über Angewandte Sprachverarbeitung.  
Auskunft: Lyn Bates, BBN Systems & Tech Corporation, 10 Mouton Street, Cambridge MA 02238, U.S.A.
- 1.4.-4.4.92 in Paris  
9th International Humor Conference  
Auskunft: Françoise Barioud, 39 rue Ste Croix de la Bretonnerie, B 23, F-75004 Paris.
- 03.04.-05.04.1992, Buffalo USA Cognition and Representation Info: Center for Cognitive Science,  
651 Baldy Hall, State University of New York at Buffalo, Buffalo, New York 14260  
Email: dcp@cs.buffalo.edu
- 05.04.-09.04.1992, Oxford GB ALLC-ACH'92  
Joint Conference This event covers literary, linguistic and humanities computing.  
Info: Centre for Humanities Computing, Oxford University Computing Service, 13 Banbury Road, Oxford OX2 6NN, England.  
voice: +44-8865-273 200  
Fax: ...273 275  
ALLCACH@VAX.OX.AC.UK
- 13.04.-14.04.1992, München  
6. Workshop Planen und Konfigurieren  
Veranstalter: Forschungszentrum für Wissensbasierte Systeme, GI-FG 1.5.3  
(Planen und Konfigurieren)  
Info: Tilo Messer, Forschungszentrum für Wissensbasierte Systeme, Orleansstr. 34, W-8000 München 80
- 21.04.-24.04.1992, Univ Vienna, AT  
Eur. Meeting on Cybernetics and Systems Research This is the 11 th meeting from a series starting in 1972. There will be a number of symposia. Submit 7 single spaced, max 50 lines A4 by Oct 15. Papers and info: Austrian Society for Cybernetic Studies, Schottengasse 3, A-1010 Wien 1, Austria.  
voice +43 1 5353 2810,  
fax: +43 1 630 652  
sec%ai-vie.uucp@relay.eu.net
- 11.05-15.05.1992, Vancouver, Canada AI'92  
Canadian AI Conf  
This is the 9th biennial conference on AI sponsored by the Canadian Society for Computational Studies of Intelligence. Papers and info: Janice Glasgow, Dept. of Computing and Information Science, Queens University, Kingston, Ontario, K7L 3N6, Canada.  
voice +16135456058  
fax:...545 6513  
janice@qucis.queensu.ca
- 27.5.-30.5.92 in Glottertal  
Arbeitstagung über biosemiotische Modelle in der Medizin. Thema: Subjektivität und Intersubjektivität von Krankheitszeichen.  
Auskunft: Jörg M. Herrmann, Klinik für Rehabilitation, Glotterbad, W-7804 Glottertal.
- 01.06.-05.06.1992, **Avignon, FR**  
Avignon'92 12th Int Conf AI-Expert Systems-Natural Language; TAMA'92  
This 12th conference adopts a new formula. It will comprise one scientific conference devoted to tools, techniques and applications used for Knowledge- Based Systems, one conference on Natural Language Processing and its Applications.  
Avignon'92 will also host TAMA'92, the second TERMNet symposium on terminological applications on micro-computers.  
Submission Nov 29.  
Papers and info: Jean Claude Rault, EC2, 269 Rue de la Garenne, 92024 Nanterre Cedex, France.  
voice +33 1 4780 7000  
fax: +33 1 4780 6629

**5.6.-7.6.92 in Stendal** Kolloquium

über die Beziehung zwischen Text und Bild im Werk .Stendals.

Auskunft: Michael Nerlich, Institut für französische Literaturwissenschaft, Technische Universität Berlin, Sekr. TEL 3, Straße des 17. Juni 135, W-I000 Berlin 12.

fax: + 46 8 796 9639  
voice: +46 8 789 6683  
coling@com.qz.se

**10.06.-12.06.1992, Montreal, Canada ITS****'92 2nd Int Conf on Intelligent Tutoring Systems**

The conf will focus on a board spectrum of research concerned with how AI and other advanced technologies can be applied to education and training.

Info: Prof. Claude Frasson,  
Dept d'IR.O., Universite de Montreal,  
C.P.6128, Succ. "A", Montreal (Quebec),  
Canada, H3C 3J7  
Tel.: +15143437019,  
Fax: ... 3435834,  
Email: frasson@iro.umontreal.ca

**10.6.-14.6.92 in Göttingen** Tagung über Medien des internationalen Literatur Transfers.

Thema: Übersetzungs-Anthologien.

Auskunft: Arnim Paul Frank, Sonderforschungsbereich literarische Übersetzung, Georg-August-Universität, Humboldtallee 17, W-3400 Göttingen.

**17.06.-20.06.1992, Nova Scotia, Canada****ICCAL'92 4th Int Conf on Computers and Learning**

ICCAL is devoted to theory and practice of computers and learning.

Submission by Sept 15 not accept.

Dec 15, crc Febr 15.

Info: Dr. Ivan Tomek, Jodrey School of Comp.Sc., Acadia Uni, Wolfville, Nova Scotia, BOPIXO, Canada.

ICCAL@AcadiaU.ca  
voice: +1 902542 2201  
fax: .. .542 7224

**22.06.-23.06.1992, Stockholm, SE****Transiation and the European**

**Communities** In particular, papers are welcomed on the following issues: A pluralingual community and its impact on the national languages; Terminological support and language control; The translation market after 1992;

Submit by Jan 15, not accept. March 1, crc May 15.

Info: Eurofat AB, Skeppsbron 26, S11130 Stockholm, Schweden.

**01.07.-03.07.1992, Aberdeen, Scotland**  
**9th Int Conf Maschine Learning**

subm: early January 1992

Info: ML-92, Dept. of Computing Science,  
King's College University of Aberdeen,  
Aberdeen, AB92UB Scotland, Tel: +44 224 272296

fax: +44 224 487048  
Janet: ml92@uk.ac.abdn.cs

**23.07.-28.07.1992, Nantes, France COLING - 92 14th Int Conf on Computational Linguistics**

The conf will last 5 full days (not counting sunday).

Pre-COLING tutorials: 22-22 July(2-1j2 days) .

All topics in Computational Linguistics are acceptable.

Submission: Nov 1. Send 6 A40r

8.5x11 inch copies of the full paper to the Chair: Prof. A.Zampolli, Universita di Pisa, ILC, via della Faggiola 32, 1-56100 Pisa, Italy;

voice:' +3950 560481

fax: +39 50 589055. Not acc March 1, crc May 1.

Info: GETA-IMAG, COLING-92, BP 53X, F-38041 Grenoble France Exhib-info: EC2, G .d' Aumale, 269 Rue de la Giuenne, N anterre Cedex, France. voice: +33 1 47 80 7000,

fax: .. .80 6629

**31. 7.-2.8.92 in Canterbury** Conference on Literary Semantics.

Auskunft: Trevor Eaton, Honeywood Cottage, 35 Seaton Avenue, Hythe, Kent, CT21 5HH, England.

# HYPertext media

R. Kuhlen

## **Hypertext**

*Ein nicht-lineares Medium zwischen Buch und Wissensbank*

1991. XV, 350 S. (Edition SEL-Stiftung) Brosch. DM 48,- ISBN 3-540-53566-7

In dieser Einführung in die Hypertextmethodik werden die wesentlichen Elemente von Hypertextsystemen, die Informationseinheiten und Verknüpfungsarten sowie die hypertextspezifischen Navigationsformen, ausführlich theoretisch und anschaulich am Beispiel existierender kommerzieller und experimenteller Hypertextsysteme behandelt. Besonderer Wert wird auf den Zusammenhang von Hypertext und Information Retrieval und die Einsatzmöglichkeiten von Hypertext in Lernumgebungen gelegt. In theoretischer Hinsicht wird untersucht, ob sich Hypertextbenutzer mit Vorteil gegenüber anderen Medien die Information erarbeiten können, die sie aktuell in kritischen Situationen benötigen, ob und wodurch also gegenüber traditionellen linearen Formen ein "informationeller Mehrwert" erzielt wird. Außerdem wird die Möglichkeit diskutiert, mit Hilfe von Textanalyseverfahren und Techniken der Wissensrepräsentation Hypertexte aus Texten automatisch aufzubauen.

P.A. Gloor, N.A. Streitz (Hrsg.) **Hypertext und Hypermedia**

*Von theoretischen Konzepten zur praktischen Anwendung*

1990. IX, 302 s. (Informatik-Fachberichte, Bd. 249) Brosch. DM 51,- ISBN 3-540-53089-4

Der Band enthält Beiträge zu den meisten momentan aktuellen Projekten im deutschen Sprachraum. Dabei werden einerseits grundlegende Konzepte sowie existierende Hypertextsysteme dargestellt. Andererseits werden die Beziehungen zu und die Integration von Datenbank- und Expertensystem-Technologien diskutiert. Schließlich werden praktische Anwendungen von Hypertext in der Ausbildung und in der technischen Dokumentation vorgestellt.

D.H. Jonassen, H. Mandl (Eds.)

## **Designing Hypermedia for Learning**

With the Assistance of S. Wang and P.M. Fischer

1990. XXV, 457 pp. 64 figs. 6 tabs.

(NATO ASI Series F, Vol. 67)

Hardcover DM 128,- ISBN 3-540-52958-6

This most unusual book results from a NATO Advanced Research Workshop (ARW) held in Rottenburg am Neckar, FRG, in July 1989, within the new special programme on Advanced Educational Technology. The book is an annotated collection of papers defining hypermedia in several contexts and discussing how to design the information structure, the user interface and the learning model. Hypermedia development is compared with instructional design processes, and development assumptions and processes for different hypermedia environments are discussed.

H. Maurer (Hrsg.)

## **Hypertext/Hypermedia '91**

Tagung der GI, SI und OCG Graz, Mai 1991 Proceedings 1991. VIII, 299 S. (Informatik-Fachberichte, Bd. 276)

Brosch. DM 54,- ISBN 3-540-54145-4

Das Buch Hypertext/Hypermedia '91 ist der Tagungsband des gleichnamigen Symposiums, das als erste Veranstaltung im deutschsprachigen Europa stattfand, die von den drei großen Informatikorganisationen Deutschlands (GI), der Schweiz (SI) und Österreichs (OCG) getragen wurde.

Der Band enthält 27 Beiträge, die aus einer großen Anzahl von eingereichten Arbeiten ausgewählt wurden, und einen eingeladenen Vortrag von N. Magnenat Thalmann zum Thema "Multi Media, Virtuelle Realitäten und Computeranimation".



dtp.30054/V/1

## Rechenschaftsbericht über die Tätigkeit der Arbeitstreffen Computerlinguistik-Studiengänge 1986-1991

von Bernhard Kelle  
(Freiburg i.Br.)

### 1 Vorbemerkung <sup>1</sup>

Im August 1986 luden P. Hellwig und H.D. Lutz zum "1. Arbeitstreffen 'LDV-Studiengänge'" ein. Auf Initiative der Arbeitskreise "Ausbildung und Berufsperspektiven" und "Linguistische Datenverarbeitung und Nachbarn" der Gesellschaft für Linguistische Datenverarbeitung e.V. (GLDV) sowie im Namen von deren Vorstand und Beirat trafen sich Vertreter 2 interessierter Gruppen und/oder Institutionen zu einem Meinungsaustausch. Obwohl "erste" Einladungen genauso wie "erste" Bände häufig auch die letzten bleiben, folgten diesem weitere 18 Arbeitstreffen (AT) in verschiedenen Städten der Bundesrepublik.

An dieser Stelle sei allen Teilnehmern der AT, die im folgenden summarisch genannt werden, für ihre Mitarbeit herzlich gedankt. Dem Vorstand und Beirat der GLDV ist für die finanzielle und institutionelle Hilfe zu danken, die das vom AT Erreichte erst hat gedruckte Wirklichkeit werden lassen.

Teilnehmer der AT waren: 3

I.S. Batori, R. Drewek, B. Endres-Niggemeyer, A. Franzke, T. Gardner, Ch. Habei, U. Hahn, P. Hellwig, M. Heyn, L. Hitzenberger,

<sup>1</sup> Der Rechenschaftsbericht wurde auf dem Arbeitstreffen in Trier (27.9.91) von den anwesenden Mitgliedern diskutiert und in der vorliegenden Form verabschiedet.

<sup>2</sup> Mit dem Ausdruck 'Vertreter' sind sowohl männliche als auch weibliche Menschen angesprochen; dies gilt gleichermaßen für alle anderen in diesem Bericht verwendeten generischen Personenbezeichnungen.

<sup>3</sup> Nicht aufgenommen sind die eingeladenen Referenten des 'Status-Seminars (s.u.)'.

K. Jakob, S. Kanngießner, B. Kelle, G. Knorz, W. Kreitmeier, W. Löser H. Lutz, M. Lutz-Hensel, J. Neuhaus, P. Ovenhausen, H.P. Pütz, W. Putschke, B. Rieger, U. Reyle, B. Schaedler, K.D. Schmitz, H. Schnelle, G. Schweisthal, G. Willee, K. Wothke.

Die Arbeit, die die Teilnehmer der AT in den vergangenen Jahren geleistet haben, schlug sich in folgenden Publikationen nieder: 4

#### 'Konturierungspapier'

U. Hahn, M. Heyn, L. Hitzenberger, B. Kelle, H. Lutz, K. Wothke: Empfehlungen der GLDV zur Konturierung des Faches Computerlinguistik.

Hg. v. Gesellschaft für Linguistische Datenverarbeitung. 0.0., 1989.

In leicht veränderter Version auch publiziert als: Dit: Zur Konturierung des Faches Computerlinguistik. In: Informatik-Spektrum 13 (1990): S. 31-37.

In einer unvollständigen Vorversion: Dit.: LDV-Forum 5 (1988): HA. S. 86-96.

#### 'Dokumentation "Status-Seminar" ,

Status-Seminar Ausbildungsprofile von Computerlinguistik-Studiengängen. 14.15. Juni 1990, Universität Bonn. Zusammenfassende Dokumentation. Hg. v. "Arbeitstreffen Computerlinguistik Studiengänge" im Auftrag der Gesellschaft für Linguistische Datenverarbeitung (GLDV) e.V. Saarbrücken, 1990.

<sup>4</sup> Vorangestellt sind die Kurzbezeichnungen, unter denen im weiteren Text auf die Publikationen referiert wird.

'Dokumentation "Ausbildungsprofile", Computerlinguistik in Deutschland. Ausbildungsprofile von CL-Lehrangeboten (demnächst).

Das 'Konturierungspapier' 5 stellt den Versuch dar, die Umriss des Faches CL in Hinblick auf die Inhalte und die Berufsqualifikation der Absolventen zu beschreiben. Darüber hinaus dient es dazu, formale und inhaltliche Minimalanforderungen sowie Lehr- und Lernziele für einen Diplomstudiengang zu formulieren, womit auf die Kompatibilität von erbrachten Leistungen zwischen verschiedenen Studiengängen in der BRD gezielt wird.

Die 'Dokumentation "Status-Seminar", 6 dient dazu, den gegenwärtigen Ausbaustand der Computerlinguistik (CL) in den alten Bundesländern darzustellen und beruht zum einen auf einer schriftlichen Erhebung, die allen potentiell CL-interessierten Institutionen zugegangen war. Zum zweiten geht sie auf eine Veranstaltung, das "Status-Seminar" zurück, wo Vertreter eingeladener Institutionen ihren Studiengang beschrieben. Die 'Dokumentation "Status-Seminar", ist der Vorläufer der nunmehr vorliegenden, vervollständigten 'Dokumentation "Ausbildungsprofile", mit deren Erscheinen die inhaltlichen Vorarbeiten für den 'Workshop "Computerlinguistik in Deutschland. Forschungs-, Ausbildungs- und Anwendungspotential einer innovativen Disziplin" ' 7 der inzwischen unter dem Titel 'Language Industries von der GLDV angestrebt wird, abgeschlossen sind. 8 Mit Anstößen für die Ausrichtung dieses Workshops endet die langjährige Arbeit der AT.

Welche Gründe für die rege Tätigkeit eines relativ informellen Gremiums vorlagen, und wie die Arbeit zu den o.e. Ergebnissen führte, soll im folgenden detaillierter nachgezeichnet werden.

5 Zu den Einzelheiten s. u. Pkt. 4.

6 Zu den Einzelheiten s. u. Pkt. 6.

7 Zu den Einzelheiten s. u. Pkt. 7.

8 Zu den Vorarbeiten für den Workshop zählt auch der 'Studienführer Linguistische Datenverarbeitung' Computerlinguistik für die deutschen Universitäten. Ermittelt und bearbeitet von Magdalene Lutz-Hensel. Gesellschaft für Linguistische Datenverarbeitung e.V. (GLDV). Saarbrücken 1991.

## 2 Motive für die Etablierung der Arbeitstreffen

Die große Leistungsbereitschaft des Teilnehmerkreises ergab sich erstens daraus, daß die Treffen offen, d.h. weder auf einen bestimmten Personenkreis noch auf bestimmte Institutionen beschränkt waren.

Zweitens verstärkte sich im Laufe der Arbeit die materielle und ideelle Unterstützung durch die GLDV, so daß - bei Fortbestehen des offenen Teilnehmerkreises - eine motivierende Perspektive für eine offizielle Weiterführung der an gefangenen Projekte sichtbar wurde.

Drittens konsolidierte sich mit zunehmender Zahl der AT der Teilnehmerkreis; die so entstandene personelle Kontinuität verhinderte weitgehend, daß bestimmte Probleme jedesmal von neuem exhumiert werden mußten.

## 3 Das 1. Arbeitstreffen (Bonn, 30.-31. 8. 1986)

Der Verlauf des 1. AT soll etwas näher beschrieben werden, da sich die Stoßrichtung des AT in den folgenden Jahren aus den dort erarbeiteten Grundlagen ergab.

Die Bemühungen um die Einrichtung und/oder den Ausbau des Faches Computerlinguistik hatte Mitte der 80er Jahre in der ganzen BRD ein so großes Ausmaß erreicht, daß es dringend geboten erschien, für einen Überblick über die entstehenden oder schon etablierten Studiengänge zu sorgen und, wenn möglich, Anstöße für eine Standardisierung der Studiengänge zu geben. Unter Schirmherrschaft der GLDV luden daher, wie einleitend erwähnt, P. Hellwig und H.D. Lutz zu einem Treffen der interessierten Hochschullehrer und Studenten ein. Erstens sollte durch das Treffen der Status-Quo bestimmt werden, zu dessen vorläufiger Bestimmung an alle Teilnehmer ein Fragebogen verschickt wurde, zweitens sollten Vorschläge für eine Vereinheitlichung der Studiengänge gesammelt werden, und als dritter Punkt waren allgemeine Probleme im Zusammenhang mit CL zur Diskussion aufgerufen. Diese Diskussion zog sich übrigens quer durch (fast) alle AT's vor allem, wenn neue Teilnehmer zum AT stießen - und konkretisierte sich jeweils in der Existenzfrage: 'Gibt es das Fach CL (überhaupt) und, falls ja, heißt es so?' Die Bereitschaft, diese Frage zu diskutieren" war anfangs leidenschaftlich vorhanden, nahm aber

mit wachsender Zahl der AT drastisch ab.

Als Ergebnis der o.e. Fragebogenaktion lag dem Plenum eine Reihe von Stellungnahmen und Beschreibungen von CL-Aktivitäten vor, die zum Auftakt des Treffens von den Verfassern erläutert wurden. Als wichtiges Problem trat in diesen Äußerungen mehrfach die Frage hervor, was die Aufgaben eines CL-Studiengangs und was die Aufgaben der Absolventen sein sollen. Zur ersten Teilfrage wurden folgende Aspekte genannt:

- Aus Sicht der Industrie seien Klagen über schlechte und falsche Ausbildung zu hören;
- gefordert werde die Fähigkeit zur Operationalisierung von Kommunikationsabläufen;
- weniger interessant sei die intensive Beherrschung traditioneller und strukturell-linguistischer Kenntnisse;
- eine wichtige Forderung ziele auf sehr gute Programmierkenntnisse.

Zum zweiten Teil der Frage - nach den Aufgaben der Absolventen - ergaben sich folgende Schwerpunkte:

- Besetzung von 'Nischen' (starke Spezialisierung);
- Prototyping;
- die gekonnte Benutzung von Anwenderprogrammen;
- die Vermittlung von Informatikkenntnissen in den Schulen durch Deutschlehrer;
- die Umsetzung der Kenntnisse über Strukturen von Kommunikationsabläufen in natürlicher Sprache in industriell nutzbare Abläufe.

Die Teilnehmer trennten sich, um die aufgeworfenen Fragen im Rahmen zweier Arbeitsgruppen weiter zu diskutieren: Gruppe 1 widmete sich den Inhalten des Studiums. Gruppe 2 erörterte die Frage nach den Fähigkeiten der Absolventen und den Berufsfeldern.

Gruppe 1 erarbeitete folgende Forderungen an die 'Inhalte des Studiums':

- Empirisch fundierte Methodenkenntnis, die theoriegeleitet und theorieproduzierend sein müsse;
- Kenntnis der Methoden zur Darstellung und Theorieentwicklung;

- ingenieurmäßige Durchführung in Hinblick auf Arbeitsfeld und Praxis;
- Kenntnisse im Bereich Wissenschaftstheorie;
- Kenntnisse im Bereich Gesellschaftsrelevanz.

Die zweite Gruppe versuchte zunächst zu erhellen, wie und wo die Schwerpunkte der Ausbildung zu liegen haben: Informatiker mit Linguistikkenntnissen vs. Linguist mit Informatikkenntnissen? Die kontroverse, z.T. heftige Diskussion ergab, daß vom Anforderungsprofil her eher der Linguist gesucht sei, da nur er Kenntnisse über Theorien der sprachlichen Kommunikation verfüge und er darüber hinaus - kommunikative Fähigkeiten auch aktiv besser einsetzen könne. In bezug auf Berufsfelder, in die Absolventen eintreten könnten, wurde festgestellt, daß CL als Anwendungsbereich in der Industrie so gut wie unbekannt sei, daß es erst gelte, einen Bedarf zu wecken. Ferner wurde davon berichtet, daß Absolventen, die eine Stelle besetzt hätten, Nischen ausfüllten, die zur CL zählten und damit weiteren Bedarf schafften, der durch die Anstellung weiterer Absolventen gefüllt werden müßte. Die Integration von Linguistik und Informatik, als weiteres Problem, leide vor allem unter der Vielzahl von Modellen und Theorien der Linguistik, die zu einem großen Teil in der Informatik unbekannt und kaum zu algorithmisieren seien.

Die Diskussion am zweiten Tag des AT versuchte Klarheit in grundsätzlicheren Problemen zu schaffen, wie z.B. der Frage nach dem Namen des Faches. Ein Einstieg in die Details der Lehrinhalte (z.B. Parsertechniken, Kompositumszerlegung etc.) wurde versucht, es ergab sich jedoch, daß diese Diskussionen besser im Rahmen der Arbeitsgruppen weitergeführt und auf einem 2. Treffen behandelt werden sollten.

Das 1. AT hatte hiermit wesentliche - und strittige - Fragen gestellt und mit der Begründung der beiden Arbeitsgruppen (AG1: Inhalte und Methoden der Computerlinguistik; AG2: Fähigkeiten von CL/LDV-Absolventen, Ausbildungsprofil, Berufsprofil) einen Grundkonsens gestiftet, der sowohl die inhaltlich/sachliche Ebene, als auch die personelle einschloß.

## 4 Die Erarbeitung des 'Konturierungspapiers'

AG 1 und AG2 trafen sich in der Folgezeit getrennt, und erarbeiteten schriftliche Vorschläge, die im weiteren bei wieder vereinigten Sitzungen 9 zu einer vorläufigen, nur die Punkte 'Einleitung', 'Inhalte und Methoden' sowie 'Fähigkeiten und Ausbildungsprofil' enthaltenden Fassung des 'Konturierungspapiers' zusammengeschmiedet wurden. Diese Version wurde im LDV-Forum publiziert (s. o.). Gleichzeitig war die Bitte zu einer Diskussion veröffentlicht worden, die aber nur geringes Echo erzeugte. Die Diskussion in der Gruppe selbst war dagegen sehr intensiv und kreiste bei der Ausarbeitung des Textes u. a. um folgende Themenbereiche:

- *Das Fachverständnis:* Erörtert wurde vor allem die Frage, wie sehr es - nach damaliger Terminologie - auf 'natürlichsprachliche Mensch-Maschine-Kommunikation' beschränkt sein sollte oder demgegenüber auch den natürlichsprachlich orientierten Teil der KI-Forschung und Anwendungsbereiche (Maschinelle Übersetzung, Lexikologie, Automatische Sprachkartographie) mit einbeziehen müsse;
- wissenschaftstheoretische Voraussetzungen des Faches;
- das anteilige Verhältnis von Informatik und Linguistik;
- Studieninhalte und Stundenzahlen eines Diplomstudiengangs;
- die Frage, welche Rahmenbedingungen für die Ansiedlung einer CL erfüllt sein müßten (z.B. grundständige Informatik).

Auf dem Treffen in Saarbrücken (11./12.3.88) wurde beschlossen, eine Kommission 'CL-Studiengänge' zu gründen (s.u.), die beim nächsten Treffen in Regensburg (12./13.5.88) das Papier in einen endredaktions-fähigen Zustand bringen sollte. In einer beinharten Sitzung gelang dies, so daß in Kiel (2./3.7.88) die endgültige Version verabschiedet werden konnte. Die Mitgliederversammlung der GLDV beschloß im März 1989, den erarbeiteten Text als offizielles Papier der GLDV in

9Folgende Treffen fanden bis zur endgültigen Verabschiedung des 'Konturierungspapiers' durch das AT statt: Bonn (22./23.11.86), Freiburg (14./15.2.87), Heidelberg (11./12.4.87), Koblenz (30./31.10.87), Bonn (28./29.11.87), Darmstadt (9./10.1.88), Saarbrücken (11./12.3.88), Regensburg (12./13.5.88), Kiel (2./3.5.88).

Form einer Broschüre zu veröffentlichen. Vorstand und Beirat der GLDV schlossen sich an, und im September 1989 erschienen die "Empfehlungen der GLDV zur Konturierung des Faches Computerlinguistik" (s. o.).

Um eine noch bessere Verbreitung des Textes zu erreichen, wurde eine in wenigen Details veränderte Fassung im 'Informatik-Spektrum' veröffentlicht (s. o.). Unter der Rubrik 'Zur Diskussion gestellt' erscheinend, rief diese Fassung als Reaktion einen Beitrag von G. Görz, Ch. Habel, W.v. Hahn und M. Pinkal hervor (s.u.).

## 5 Die Zusammenarbeit mit der 'Sektion Computerlinguistik' der Deutschen Gesellschaft für Sprachwissenschaft

Bereits erwähnt wurde die Konstruktion des AT als 'offene' Veranstaltung. So war es selbstverständlich, daß im Rahmen der Gründung der 'Sektion CL' der DGfS von beiden Seiten der Wunsch bestand, die Diskussion um eine Vereinheitlichung des Studiengangs CL gemeinsam zu führen. An dem AT in Bonn (28./29.11.87) nahmen daher zwei Vertreter der Sektion (Ch. Habel, S. Kanngießer) teil. Es wurde dabei erstmals deutlich, daß die lockere Organisationsform des AT auch Nachteile hatte, da den Vertretern der DGfS die Nicht-Institutionalisierung, der wechselnde Teilnehmerkreis und der deshalb nicht überblickbare Arbeitsaufwand Sorgen bereiteten. Hinzu trat wohl noch das Problem der Vermittelbarkeit der gegebenen Konstruktion in der heimischen Sektion.

Um hier ein formal definiertes Forum für die gemeinsame Arbeit zu schaffen, schlug daher H.D. Lutz die Einrichtung einer Kommission 'CL-Studiengänge' durch die GLDV vor.

Sie sollte ein offizielles Mandat für die Kooperation mit anderen Gremien haben. Der Vorstand der GLDV unterstützte das Vorhaben, wenngleich ein offizielles Mandat nicht zu erhalten war. Die Kommission - gegründet auf dem AT in Saarbrücken (11./12.3.88) - konnte daher den Erwartungen der Nachbarorganisationen nicht genügen und 'beschränkte' sich darauf, die Endredaktion des 'Konturierungspapiers' einzuleiten (s.o.). Weitere Treffen mit Vertretern der DGfS wurden geplant, kamen dann aber nicht mehr zu

stande, so daß auch die Kommission nicht mehr tätig werden mußte/konnte. Ein von der DGfS geplantes Papier 'Empfehlungen für Studiengänge des Faches Computerlinguistik' wurde von deren zuständigen Gremien im März 1990 verabschiedet und liegt publiziert vor (DGfS Bulletin. Nr. 31. (Dez. 1990): S.24-32). Als Reaktion aus Kreisen der DGfS auf das 'Konturierungspapier' kann die Antwort angesehen werden, die G. Görz, Ch. Habei, W.v. Hahn und M. Pinkal unter dem Titel 'Computerlinguistik als Studienfach' in Informatik-Spektrum 13 (1990): S. 276-279 veröffentlichten. Kernpunkte der Kritik waren u. a.: "Im Grundstudium zuviel und verfrühte Anwendungsorientierung und zu wenig systematische Grundlagenvermittlung; im Hauptstudium zu viele und zu heterogene Informatikthemen und zu wenig theoriebasierte computerlinguistische Anwendung." (ib.: S. 278).

## 6 Das Status-Seminar 'Ausbildungsprofile von CL- Studiengängen' (Bonn 14./15.6.1990)

Nachdem das 'Konturierungspapier' fertiggestellt war, zog H.D. Lutz in einem Resümee den Schluß, daß von den im I. AT anvisierten Zielen nur ein kleiner Teil als bewältigt gelten konnte. Dringlich schien nunmehr, das "Ausbildungsprofil existierender Studiengänge mit LDV /CL-Anteil und Berufsqualifikation ihrer Absolventen" zu erheben. Als weiteres, aber zunächst nur tentatives Ziel nannte er "die Formulierung von Lehr- und Lernzielen für einzelne Veranstaltungen eines CL-Curriculums".

Sein Vorschlag wurde von den Teilnehmern aufgegriffen und ein allgemeines Brain-Storming, auch in kleineren Gruppen, führte zu folgender Konkretisierung der bei den Teilziele, zunächst unter dem Arbeitstitel 'Memorandum zur Lage der CL in der BRD', später 'Dokumentation "Status-Seminar" ' (s. o.). 10

- M. Lutz-Hensel legte ein Positionspapier zur Frage 'Ausbildungsprofil' vor;

10 Das AT traf sich zur Ausarbeitung der hierzu anstehenden Fragen in Koblenz (21./22.7.89), Heidelberg (1./2.12.89), Trier (26./27.1.90), Siegen (29.3.90), schließlich anlässlich des Status-Seminars, Bonn (11./15.6.90).

- H.D. Lutz, M. Lutz-Hensel, B. Schaefer und G. Willee arbeiteten die Details zur Erhebung der 'deskriptiven' Teile der 'Dokumentation' aus;

- B. Rieger gab dem Unternehmen eine neue Stoßrichtung, indem er von einem gewissen Interesse des Bundesministeriums für Bildung und Wissenschaft (BMBW) berichten konnte, sich an einem 'Workshop' zu beteiligen, der - ähnlich einer von den Wirtschaftswissenschaften durchgeführten Veranstaltung Lehre/ Forschung, Wirtschaft und Politik informieren und zusammenführen sollte.

Aus den vorgenannten Berichten und Aktivitäten ergab sich folgende Planung:

Die 'Lage der CL' sollte mit Hilfe einer umfassenden Fragebogenaktion erkundet werden. Als Schwierigkeit dabei hatte sich schon bei früheren Erhebungsversuchen ergeben, daß bestimmte Details schriftlich nicht zu erhalten waren. Sie ließen sich entweder nicht präzise genug erfragen, solange keine Rückfragemöglichkeit gegeben war, oder es bestand bei den Befragten kein Interesse, schriftlich/öffentliche Erklärungen abzugeben. Um auch in solchen sensibleren Bereichen Einblicke gewinnen zu können, wurde beschlossen, ein 'Status-Seminar' abzuhalten, zu dem die Vertreter der in der BRD etablierten CL-haltigen Studiengänge eingeladen werden sollten. Auf dem AT in Trier (26./27.1.90) waren die Vorbereitungen für den Fragebogen so weit gediehen, daß er kurze Zeit später ausgesandt werden konnte. Angeschrieben wurden 59 Institutionen, von denen festgestellt und/oder vermutet werden konnte, sie seien mit CL verknüpft oder wesentlich daran interessiert.

Verlauf der Fragebogenaktion und des Status-Seminars sind ausführlich in der o. e. Veröffentlichung der GLDV ('Dokumentation "Status-Seminar" ') beschrieben. Es fand am 14./15.6.1990 in Bonn statt, wobei dem gastgebenden Institut für Kommunikationsforschung und Phonetik (IKP) an dieser Stelle - auch für alle anderen unterstützenden Maßnahmen - herzlich gedankt sei. 11

Innerhalb einer sehr kurzen Frist wurden Transkriptionen der Tonbandmitschnitte verfaßt, die Antworten auf das Umfrageschema bezogen (Heidelberg (3.8 - 4.8.90)),

11 Folgende AT-Sitzungen fanden nach dem Status-Seminar statt: Heidelberg (3./4.8.90), Wolfenbüttel (24./25.11.90), Bonn-Röttgen (20./21.4.91), Trier (27.9.91).

schließlich die Manuskripte den Referenten vorgelegt und deren Zustimmung (Ratifizierung) eingeholt - und schließlich nach einer Endredaktion (Wolfenbüttel (24./25.11.90)) noch 1990 als 'Dokumentation "Status-Seminar", (s. o.) vorgelegt. Sie ging allen Institutionen zu, die im Rahmen der ersten Umfrage positiv oder mit 'Fehlanzeige' geantwortet hatten. Gleichzeitig wurde die Bitte angeschlossen, bisher nicht berücksichtigte Aktivitäten zur Gründung eines Studienganges CL zu melden, damit die 'Dokumentation "Status-Seminar" " um diesen Kreis erweitert, als 'Dokumentation "Ausbildungsprofile" , neu aufgelegt werden könne.

die Gewinnung eines Überblicks über die CL-Aktivitäten im europäischen Ausland; 13

die Einrichtung einer Clearingstelle für CL in Forschung und Lehre sowie die Benennung/Bestellung einer dafür verantwortlichen Person.

## **7 Gegenwärtiger Stand - Planung der Veranstaltung mit dem BMBW**

Das AT in Bonn-Röttgen (20./21.4.91) war im wesentlichen der Planung des 'Workshops' gewidmet, der unter dem Thema 'Computerlinguistik in Deutschland. Forschungs-, Ausbildungs- und Anwendungspotential einer innovativen Disziplin' stehen soll. 12 Als Materialgrundlage soll ihm - neben dem 'Studienführer' (s. o.) - die vollständige 'Dokumentation "Ausbildungsprofile"', dienen.

## **8 Erkenntnisse, Zukunftsaussichten und Ende des AT**

Beim Treffen in Bonn- Röttgen wurde der Beschluß gefaßt, daß der AT mit Ende des Jahres 1991 seine Arbeit einstellt. Dem Vorstand und Beirat sollen neben diesem Arbeitsbericht folgende Empfehlungen für eine Weiterführung bzw. Kontinuitätswahrung anheimgestellt werden:

- Eine allgemeine, möglichst längerfristig gültige Beschreibung der Studieninhalte auf der Ebene einzelner Lehrveranstaltungen;
- damit in Zusammenhang der Entwurf eines Lehrbuches Einführung in die CL';
- die semesterweise Veröffentlichung eines Verzeichnisses der Lehrveranstaltungen zur CL an den deutschen Universitäten;

12 Zur geplanten Präzisierung des Titels s. o.

13 Hierzu haben einige Mitglieder des AT bereits Vorarbeiten geleistet.

# GLDV-Jahrestagung

## Protokoll der Mitgliederversammlung der GLDV vom 26.9.91 in Trier

Ch. Schneider

Beginn: 15.08 Uhr  
Ende: 18.50 Uhr  
Sitzungsleitung: B. Rieger

### Tagesordnung

1. Regularien
2. Bericht des Vorstands mit Kassenbericht
3. Entlastung des Vorstands
4. Wahl von Kassenprüfern
5. Bericht des Beirats
6. Berichte der Arbeitsgruppen/-kreise
7. Neuwahlen
  - (a) Kandidatenliste Vorstand
  - (b) Kandidatenliste Beirat
8. Mitgliedsbeiträge: Neufestsetzung nach §7
9. Satzungsänderung: Neuformulierung des §5(2)1
10. Nächste Jahrestagung
11. Arbeitsprogramm 1991/92
  - LDV-Forum
  - GLDV-Herbst/Sommer-Schule
  - Konferenz: Language Industry
12. Verschiedenes

### TOP 1 Regularien

B. Rieger stellt fest, daß die Einladung zur Mitgliederversammlung fristgerecht ergangen ist. Es sind 30 Mitglieder anwesend, 2 Anträge auf Stimmübertragung werden vorgelegt. Die Mitgliederversammlung (MV) läßt auf Antrag von B. Rieger einstimmig Gäste zu. Die Tagesordnung wird um den Punkt 12. Verschiedenes ergänzt.

Das Protokoll der letzten MV in Siegen, das durch das LDV-Forum allen Mitgliedern zugegangen ist, wird einstimmig von der MV genehmigt.

### TOP 2 Bericht des Vorstands mit Kassenbericht

Der Bericht des 1. Vorsitzenden der GLDV B. Rieger bezieht sich auf den Zeitraum seit der letzten MV am 29.3.90 in Siegen.

Es fanden drei Vorstands- und Beiratssitzungen (am 12.10.1990 in Trier, am 21. und 22.6.1991 in Koblenz und am 26.9.1991 in Trier) statt.

### LDV-Forum

Trotz der Lösungen, die für 1990 und 1991 gefunden wurden, ist das LDV-Forum weiterhin problematisch (vgl. dazu TOP 11).

### Mitgliederentwicklung

Die Anzahl der Mitglieder pendelt um ca. 400, davon sind ca. 50% Studenten. Dieser hohe Studentenanteil kann dazu beitragen, daß der wissenschaftliche Charakter der Gesellschaft gefährdet erscheint. Es ist zu

beobachten, daß trotz steigender Mitgliederzahlen die Aktivitäten der Mitglieder stagnieren bzw. sinken. Zentral für die GLDV sind die Aktivitäten der Arbeitskreise (AK). Eine aktive Mitarbeit v. a. der Studierenden ist hier erschwert, da der erforderliche finanzielle Aufwand von der GLDV nicht mitgetragen werden kann. Dies führt zu der paradoxen Situation, daß die niedrigen Mitgliedsbeiträge (der Studentenbeitrag blieb seit der Gründung des Vereins unverändert) gerade jene zu unterstützen verhindern, die bei aktiver Mitarbeit in AKs dieser finanziellen Unterstützung bedürfen: die Studierenden.

### Kooperationen - KONVENS

Die GLDV beteiligte sich weiter an der Initiative von G. Görz zur Veranstaltung gemeinsamer Tagungen aller computerlinguistischen Vereinigungen. Deren Einzeltagungen pro Jahr lassen qualifizierte Beiträge in kaum ausreichender Zahl erwarten. Diese Initiative zielt - bei aller Diversifikation der Vereine auf eine verstärkte Zusammenarbeit in der Zukunft. Mehrere Tagungen im Jahr könnten kaum ausreichend Beiträge auf akzeptablen Niveau erwarten.

Es fanden mehrere Arbeitstreffen statt, auf denen ein Entwurf für eine entsprechende Vereinbarung erarbeitet wurde. Beteiligt sind die folgenden Organisationen:

- die Deutsche Gesellschaft für Sprachwissenschaft (DGfS) vertreten durch die Sektion 'Computerlinguistik'
- die Gesellschaft für Informatik (GI) vertreten durch den Fachausschuß 1.3 "Natürliche Sprache" im Fachbereich 1 "Künstliche Intelligenz"
- die GLDV
- die Informationstechnische Gesellschaft (ITG im VDE) und die Deutsche Gesellschaft für Akustik (DEGA) vertreten durch den gemeinsamen Fachausschuß "Sprachakustik"
- die Österreichische Gesellschaft für Artificial Intelligence (ÖGAI)

Im zweijährigen Turnus werden KONVENS-Tagungen (*Konferenz* Verarbeitung natürlicher Sprache) veranstaltet. Die Ausrichtung der Tagungen wird abwechselnd von einer der beteiligten Gesellschaften übernommen. Die beteiligten Gesellschaften verpflichten sich, für die Tagung zu werben und in KONVENS-Jahren keine konkurrierenden Veranstaltungen zu planen.

Die GLDV tritt als ganze Gesellschaft bei, die anderen Fachgruppen und Sektionen sind kleiner. MVen der GLDV könnten auch im Rahmen von KONVENS-Tagungen abgehalten werden. Die GLDV würde ihre Jahrestagungen daher nur noch im zweijährigen Turnus selbst veranstalten.

### Veranstaltungen/Veröffentlichungen

Das *Status-Seminar* zu Ausbildungsprofilen - veranstaltet vom Arbeitstreffen Computerlinguistik-Studiengänge - fand in Bonn statt, es waren nahezu alle relevanten Institutionen vertreten. Die Ergebnisse wurden in Zusammenarbeit mit H. Haller in einer Broschüre veröffentlicht.

Eben fertiggestellt wurde auch die Neuauflage des *Studienführers LDV/CL*, eine Publikation der GLDV, die das respektable Resultat der Arbeiten des AK Ausbildungs- und Berufsperspektiven darstellt. Der Studienführer ist als Publikation über den Rahmen der GLDV hinaus von Interesse und kann zum Preis von 10,- DM über H. Haller (Saarbrücken) bezogen werden.

B. Rieger nimmt diese beiden Publikationen zum Anlaß, darauf hinzuweisen, daß sie erwachsen sind aus dem langjährigen Engagement von M. Lutz-Hensel und H. D. Lutz, denen er an dieser Stelle für ihre Arbeit sehr herzlich dankt.

### Herbstschule CLAS 90

Die Herbstschule hatte 70 Teilnehmer. Die abschließende Bewertung (ausgewertet von R. Köhler) zeigte ein durchwegs positives Echo der Teilnehmer. 1991 wurde keine Sommer/Herbstschule wegen der zu diesem Termin stattfindenden Saarbrücker Sommerschule angeboten. Im Gegensatz zu den großen Sommer/Herbstschulen (Leuven, Saarbrücken) soll die GLDV-Herbstschule eine überschaubare deutschsprachige Veranstaltung mit maximal 100 Teilnehmern bleiben, die den Akzent auf praktischen Übungen legt. Für 1992 ist von D. Rösner wieder eine Herbstschule in Ulm geplant. Wer an der Mitarbeit als Dozent oder evtl. Veranstalter einer (weiteren) Herbstschule Interesse hat, möge sich bitte mit H. Haller in Verbindung setzen.

### Qualico-91

Die Qualico hatte ca. 70 Vollzeit-Teilnehmer aus 16 Ländern. Die Tagung ist so gut gelaufen, daß eine Wiederholung in ca. 3 Jahren

geplant ist. Die GLDV trat hier als Mitorganisatorin auf, ein DFG-Zuschuß konnte aufgrund der kontinuierlichen Aktivitäten der GLDV gewährt werden.

#### Kassenbericht (B. Schaefer)

Das Finanzamt verlangt ebenfalls eine jährliche Abrechnung, von den auf der letzten MV bestellten Kassenprüfern wurde die Kasse geprüft, es wurden keine Beanstandungen erhoben.

Der aktuelle Kontostand beträgt 6.009,07 DM. 1990 stehen den Einnahmen von 11.403,78 DM (davon 9.771,- DM Mitgliedsbeiträge) die folgenden Ausgaben gegenüber:

Vorstand (Porto MV 90, Briefpapier)	1191,10 DM
CL-Statusseminar (Porto, Kopien, 100 DM Hilfskräfte, 50 Disketten)	424,17 DM
Jahrestagung 89 (15 Autorenenexemplare d. Tagungsbandes)	712,30 DM
Jahrestagung 90 (Druckkostenzuschuß)	2000,- DM
Kasse (Kontogebühren, Bankrückforderungen)	645,10 DM
LDV-Forum (erst 1991 gebucht)	6000,- DM
<b>Gesamt</b>	<b>10972,67 DM</b>

Die Einnahmen entsprechen den Ausgaben, es sind keine Reisekosten des Vorstands oder von AK-Mitgliedern enthalten. Zur Förderung weiterer Aktivitäten ist damit kein finanzieller Spielraum.

### TOP 3 Entlastung des Vorstands

Die MV beschließt die Entlastung des Vorstands ohne Gegenstimme, mit 5 Enthaltungen (Vorstand).

### TOP 4 Wahl von Kassenprüfern

Als Kassenprüfer für 1991 werden G. Willee und G. Frackenpohl einstimmig (ohne Enthaltung) gewählt.

### TOP 5 Bericht des Beirats

W. Lenders gibt den Bericht des Beirats. Er bedauert die geringe Beteiligung der Beitragsmitglieder an den Sitzungen (2-3 Beiratsmitglieder). Künftig sollten Beiräte verstärkt mitarbeiten. Die Arbeitsthemen entsprechen denen des Vorstands (vgl. TOP 2), zumal die Sitzungen meist kombinierte Beirats- und Vorstandssitzungen waren.

Der Beirat befürwortet eine Beitragserhöhung und die vorgeschlagene Satzungsänderung; zusammen mit dem Vorstand suchte er nach Alternativen für das LDV-Forum (vgl. TOP 11); weitere Stichworte für die Beitragsarbeit sind: Weiterentwicklung der Ergebnisse des Arbeitstreffens Computerlinguistik-Studiengänge, Studienführer LDV/CL, KONVENS-Vereinbarung.

W. Lenders regt an, den Mitgliedern eine vorbereitete Karte für Adressen- und Kontoveränderungen zukommen zu lassen.

### TOP 6 Bericht der Arbeitsgruppen/ -kreise

#### AK Textanalyse

B. Rieger berichtet für K. Haenelt, daß keine Sitzung oder andere Aktivitäten zustande kamen und das Fortbestehen des Ak aufgrund der geringen Resonanz nicht sinnvoll erscheint.

#### AK Kodierung und Normierung maschinenlesbarer Texte (Neugründung)

G. Koch berichtet, daß die konstituierende Sitzung am kommenden Freitag stattfinden wird. 20-30 Antworten waren auf das Rundschreiben eingegangen. W. Lenders plant ein Tutorium zum Thema.

#### AK Lexikographie

Hellmann berichtet von einer Veranstaltung (Präsentation des Systems Word Cruncher).

#### AK Maschinelle Übersetzung

Von AK-Leitung D. Rösner liegt kein Bericht vor.

## AK Sprachgenerierung und phonetische Datenbanken

Von der Arbeitsgruppe Schweisthal liegt kein Bericht vor, Ergebnisse werden jedoch im LDV-Forum publiziert.

## AK Ausbildung und Berufsperspektiven

M. Lutz-Hensel betrachtet nach der Neuauflage des Studienführers einen wesentlichen Arbeitsabschnitt als abgeschlossen. Ein Schlußbericht wird noch im LDV-Forum veröffentlicht, ebenso ein Bericht über die Tagung: Anthropozentrische Produktionssysteme. Herr Heyn wird voraussichtlich künftig die Koordination des AK übernehmen.

## AK Quantitative Linguistik

R. Köhler berichtet, daß auf der konstituierenden Sitzung 12 Teilnehmer anwesend waren. Über einen Rundbrief sollen Forschungs-Projekte und industrielle Anwendungen, sowie unterstützende Software und Corpora dokumentiert werden. Dreimal jährlich sind Treffen geplant. Ein Bericht über eine Reise von R. Köhler in die Sowjetunion (ca. 20 Zentren) wird im LDV-Forum erscheinen.

## AK LDV und Nachbarn

nicht mehr existent

## AT CL-Studiengänge

H.-D. Lutz verweist auf die drei Publikationen, die aus der Arbeit des AT hervorgegangen sind (eine 'Einführung in die LDV' ist noch geplant). An diesem Punkt ist auch ein Abschluß der Arbeiten des AT vorgesehen. Ein Tätigkeitsbericht wird dem Vorstand vorgelegt, ebenso wie eine Liste weiterer zu bearbeitender Themen. In den vergangenen 5 Jahren fanden 20 Treffen statt, ein engerer Kreis kontinuierlich arbeitender AT-Mitglieder hat diesen Erfolg ermöglicht. Ihnen und dem Vorstand (v.a. bezüglich der Publikation) sei hier gedankt.

## TOP 7 Neuwahlen

### (a) Kandidatenliste Vorstand

Die Kandidatenliste für den Vorstand wurde ergänzt um H. Haller (2. Vorsitzender).

Die Kandidatenliste für den 2. Vorstand wurde mit einer Enthaltung ohne Gegenstimme angenommen. Die Kandidatenlisten für die anderen Vorstandsämter wurden einstimmig angenommen.

### (b) Kandidatenliste Beirat

Die Kandidatenliste für den Beirat wurde ergänzt um Prof. B. Körnitz Universität Leipzig, Sektion Informatik und Prof. W.

Lenders Universität Bonn IKP.

Die Kandidatenliste für den Beirat wurde einstimmig angenommen.

Damit ist die Grundlage für die Durchführung der Briefwahl geschaffen. Für die nächste Wahl wurde beschlossen eine kurze Personenbeschreibung der Kandidaten beizufügen.

## TOP 8 Mitgliedsbeiträge: Neufestsetzung nach §7

Es liegt ein schriftlicher Antrag des Schatzmeisters B. Schaefer auf Erhöhung der Mitgliedsbeiträge vor. Die Neufestsetzung der Mitgliedsbeiträge von bisher 50,- DM (Studenten 10,- DM) auf in Zukunft 100,- DM (Studenten 50,- DM) wurde vom Vorstand einstimmig beschlossen, vom Beirat empfohlen und ist allen Mitgliedern mit der Einladung zur MV zugegangen.

B. Rieger verweist auf die Ausgaben der G LDV (vgl. Kassenbericht). Für die Arbeit in der Gesellschaft werden nur gelegentlich die notwendigsten Material- und Portokosten erstattet. Weder für Vorstand, Beirat noch AK-Mitglieder konnten Reisekosten erstattet werden. Diese müssen privat finanziert werden. Es sind keine Mittel für Büro, Sekretärin oder Hilfskräfte vorhanden. Diese Arbeiten fallen jedoch in nicht unerheblichem Ausmaß an (z.B. Vorbereitung der Jahrestagung - Call, Einladung zur MV, Briefwahl, Teilnahme an Sitzungen - z.B. erforderten die KONVENS-Vereinbarungen bisher 3 Sitzungen, an denen jeweils 1-2 GLDV-Vertreter teilnahmen). Auch der Zuschuß zum LDV-Forum deckt im Wesentlichen nur die Druckkosten ab und nicht die Herstellung der Druckvorlage. Somit liegt der Beitrag für das LDV-Forum mit 6000,- DM zu niedrig und müßte bei etwa 10.000 DM/Jahr liegen. Im Vergleich mit anderen Publikationen stellt dies keinen überzogenen Ansatz dar. Im Hinblick auf die Förderung der wissenschaftlichen Arbeit in den Arbeitskreisen wäre es wünschenswert, auch hier

schwerpunktmäßig auf Antrag der Arbeitskreise Mittel bereitstellen zu können.

In der Diskussion schlägt Hellmann vor, das LDV-Forum kostendeckend zu verkaufen. Dem ist entgegenzuhalten, daß dies die Verwaltungsarbeit verdoppeln würde (LDV-Forum Abonnenten vs. GLDV-Mitglieder). Die Vergabe des LDV-Forums an einen Verlag wurde immer wieder erwogen, hat sich bisher jedoch immer als noch teurer herausgestellt.

Auf den ersten Blick erscheint die Erhöhung des Studentenbeitrags überproportional. B. Schaefer weist demgegenüber daraufhin, daß der Studentenbeitrag seit der Gründung des Vereins vor ca. 20 Jahren unverändert geblieben sei und daß nun die damals bestehende Relation 1:2 (Studentenbeitrag: voller Mitgliedsbeitrag) wieder hergestellt wird.

Der niedrige Studentenbeitrag war als Nachwuchsförderung gedacht. Es ist jedoch zu beobachten, daß die Studenten häufig austreten, sobald sie ihr Studium abgeschlossen haben. Eine Förderung aktiver studentischer Mitglieder ist mit der gegenwärtigen Regelung nicht möglich, sie befördert im Gegenteil ein passives, konsumorientiertes Verhalten.

L. Hitzberger vertritt den Standpunkt, daß der Studentenbeitrag zumindest die unmittelbaren Kosten decken sollte. Dem ist hinzuzufügen, daß gerade die Verwaltung studentischer Mitglieder überproportional schwierig ist (häufige Änderungen, Nachweis des Studentenstatus).

G. Koch hält die Erhöhung v. a. für Studenten für zu unvermittelt und fragt, worin denn nun die erweiterten Leistungen der GLDV bestünden.

H. Haller weist daraufhin, daß doch einige Aktivitäten stattfinden. Diese Vorhaben wurden zum größten Teil *nicht* aus GLDVMitteln bestritten. Publikationen (vgl. AT Computerlinguistik) und ihre Verteilung oder Veranstaltungen wie die Herbstschule verursachen einen erheblichen Aufwand. Es ist auf die Dauer nicht möglich, dies ohne eine entsprechende finanzielle Basis durchzuhalten.

Ergänzend weist B. Rieger daraufhin, daß die Herbstschule nur realisiert werden konnte, weil SIEMENS als Sponsor aufgetreten ist. Es sei richtig, daß die Erhöhung des Studentenbeitrags recht unvermittelt erscheine. Dies liege jedoch an den Versäumnissen der letzten 10 Jahre, in denen keine vernünftige Beitragsentwicklung durchgeführt wurde. Der Vorstand habe sich den Entschluß zur Erhöhung der Beiträge nicht leicht gemacht, er ist jedoch der Ansicht, daß die gegenwärtige Beitragsstruktur im Interesse der Vereinsziele revidiert werden müsse.

Der Antrag auf Schluß der Aussprache wurde mit 2 Gegenstimmen und 2 Enthaltungen angenommen.

Der Antrag auf Erhöhung der Mitgliedsbeiträge wurde mit 23 Ja-Stimmen, 5 Enthaltungen und 4 Nein-Stimmen angenommen.

## TOP 9 Satzungsänderung: Neuformulierung des §5(2)1

Der Text der vorgeschlagenen Satzungsänderung (Satzung §5 (2)1): "Der Antrag um Aufnahme als Mitglied ist schriftlich und mit der Befürwortung von zwei Mitgliedern an den Vorstand zu richten, der über ihn entscheidet." ist satzungsgemäß mit der Einladung zur MV jedem Mitglied zugegangen.

B. Rieger weist auf die Verpflichtung des Vorstands hin, über Aufnahmeanträge zu entscheiden. Aufgrund der größeren Mitgliederzahl hat der Vorstand keinen Hinweis und mangels Informationsgrundlage keine Möglichkeit, um begründet über einen Antrag zu entscheiden. Referenzen würden auch die Suche nach verschollenen Mitgliedern erleichtern.

W. Lenders weist daraufhin, daß auch in anderen Vereinigungen Referenzen Bürgenfunktion gegen einen denkbaren Mißbrauch der Mitgliedschaft haben.

Die Diskussion drehte sich um die Anzahl der erforderlichen Referenzen und darum, ob auch studentische Mitglieder einen Aufnahmeantrag befürworten können. Zur Abstimmung gelangten die folgenden Varianten:

1. Befürwortung von zwei nicht-studentischen Mitgliedern
2. Befürwortung von einem nicht-studentischen Mitglied
3. Befürwortung von einem Mitglied
4. Befürwortung von zwei Mitgliedern

Antrag 2 wurde mit 24 Ja-Stimmen, 3 Enthaltungen und 3 Gegenstimmen angenommen. (Antrag 1 wurde 5:4:19, Antrag 3 0:1:29 und Antrag 4 4:1:25 abgelehnt)

## TOP 10 Nächste Jahrestagung

Für Herbst 1992 ist die erste KONVENS-Tagung in Erlangen geplant, dort kann auch eine GLDV-MV stattfinden.

Ursprünglich war für Frühjahr 1992 eine GLDV-Jahrestagung in Kiel vorgesehen. Der Abstand zwischen der Jahrestagung 91 und 92 wäre damit sehr kurz. Darüberhinaus würde in diesem Fall auch (entgegen den KONVENS-Vereinbarungen) 1992 neben KONVENS eine weitere Tagung stattfinden.

Nach Auskunft von H. Pütz (Kiel), der sich bereiterklärt hatte, die G LDV - Tagung in Kiel zu organisieren, wäre eine GLDV-Tagung in Kiel im Frühjahr 92 oder 1993 möglich. Eine Beteiligung an der KONVENS 92 wurde mit 24 Ja-Stimmen und 3 Enthaltungen ohne Gegenstimme akzeptiert.

Der Antrag, Kiel als Ort der Jahrestagung 1993 festzulegen, wurde angenommen (26:1:0).

Arbeit des AT CL-Studiengänge (Workshop: Zur Lage der CL in Deutschland).

Den Ministerien der Länder und des Bundes sollen die Informationsverarbeitungsprobleme der nächsten Jahrzehnte nähergebracht werden (Studenten-/ Berufsperspektiven). Die GLDV als thematisch breit orientierte Gesellschaft ist dafür prädestiniert, den Stellenwert von natural language processing zu verdeutlichen. Vorbereitende Planungen und Sondierungen bei Industrie und Behörden sind durchzuführen.

W. Lenders regt an, auch auf Messen, Ausstellungen etc. zu gehen und Produkte zu präsentieren, z.B. die Expolingua in Mainz oder die Infobase. Ein Roundtablegespräch dort wäre evtl. für Entscheidungsträger attraktiv.

## TOP 11 Arbeitsprogramm 1991/92

### LDV -Forum

Die bisherigen Bedingungen, unter denen G. Knorz das LDV-Forum realisiert hat, sind nicht länger zumutbar. Nur seine persönliche Belastbarkeit hat die Entwicklung des LDV-Forums unter den gegenwärtigen Bedingungen ermöglicht. B. Rieger hatte zugesagt, 1991 und 1992 die Herstellung des LDV-Forums in Trier zu übernehmen. Da die Konditionen für die Beschäftigung von ABM-Kräften geändert wurden, stellt dies keine längerfristige Lösung dar. Eine langfristige Lösung könnte darin bestehen, für das LDVForum einen persönlich verantwortlichen Herausgeber zu finden, der in Bezug auf die Herstellung entlastet ist und sich in erster Linie mit Autorenakquisition und Autorenpflege befassen kann. Unter diesen Bedingungen ist G. Knorz prinzipiell bereit, seine Arbeit mit dem LDV-Forum fortzuführen. Die Herstellung könnte ab 1993 evtl. am IAI in Saarbrücken stattfinden.

### GLDV - Herbst/Sommer-Schule

(vgl. dazu auch TOP 2) Eine Finanzierung von Herbst/Sommer-Schulen allein über Teilnehmergebühren ist nicht möglich. Die Teilnehmerzahl sollte bei maximal 100 liegen. Potentielle Dozenten und Veranstalter mögen sich bitte bei H. Haller melden.

### Konferenz: Language Industry

Die Idee zu einer Konferenz zur Thematik Language Industry ist entstanden aus der

## TOP 12 Verschiedenes

Zu Punkt 12 der 1'0 liegt nichts vor.

Regensburg, den 17.10.91

Ch. Schneider  
Schriftführerin

B. Rieger  
Sitzungsleiter

# Mitteilungen aus der GLDV

## Bitte um update

Bei den Postversand-Aktionen der letzten Monate (Einladung zur Mitgliederversammlung, Neuwahlen, etc.) ergab sich ein erhöhter Postrücklauf wegen Unzustellbarkeit. Die/Derjenige, die/der die neue Anschrift eines der nachfolgend aufgeführten GLDV-Mitglieder kennt, möge sie bitte der Redaktion bekanntgeben:

- BARKEY , Reinhild, In der Vogelstang 1,  
6900 Heidelberg
- BARTH , Thomas, Rohrbacherstr. 75,  
6900 Heidelberg
- BLUMENTHAL , Andreas, M.A., Univer-  
sität Heidelberg, Germanistisches Semi-  
nar, Hauptstr. 207-209, 6900 Heidelberg
- BOHNERT , Stephanie, Marienborner  
Str. 15, 6500 Mainz
- BRUN , Georg, Kanzlerrain 19,  
CH-5430 Wettingen, SCHWEIZ
- BUTT , Matthias, Sophie-Charlotten-  
Str. 32, 1000 Berlin 19
- CASPARY , Christoph, Vogelsbergstr. 22,  
6000 Frankfurt 1
- DIERKS , Dr. Karin, Franckstr. 35,  
A-8010 Graz/Österreich
- DIESTELMANN , Martin, Arno-Ass-  
mann-Str. 15, 8000 München 83
- DWEHUS , Dagmar, Zur Mühle 9,  
5500 Trier
- ECKERT , Karin, Saarbrücker Str. 208,  
6602 Dudweiler
- EHLICH , Ute, Universität Erlangen-Nürn-  
berg, IMMD5, Markusstr. 3, 8520 Erlan-  
gen
- EICKE , Christine, Koppel 23,  
2000 Hamburg 1
- ERBACH , Gregor, Lessingstr. 36,  
6600 Saarbrücken
- ESA , Mohamed, Friedrich-Ebert-Anlage 47,  
6900 Heidelberg
- FIRZLAFF , Beate, Roonstr. 34,  
5400 Koblenz
- GREGOR , Dr. Bernd, Institut für deut-  
sche Philologie, Universität München,  
Schellingstr. 3, 8000 München 40
- GUCKLER , Dr. Gudrun, Software AG,  
Haardtring 100, 6100 Darmstadt
- HANKE , Manfred, Nordstr. 10,  
5300 Bonn 1
- HAUSER , Roland, Hüttenstr. 37,  
1000 Berlin 21
- HEIDEMANN , Matias, Merzhauserstr.  
14, 7800 Freiburg
- HIRSCHMANN , Astrid, Augsburgener Str.  
93a, 8400 Regensburg
- JACOB , Daniel, Schroederstr. 37,  
6900 Heidelberg
- KELLNER , Marianne, Sterbergstr. 26,  
8400 Regensburg
- KRAISS , Martin, Möhlstr. 14,  
6800 Mannheim 1
- KRAUSS , Margit, Rohrbacher Str. 75,  
6900 Heidelberg
- KRENN , Monika, Therese-Giese-Allee 50,  
8000 München 50
- KROUPA , Edith, Hesselauweg 103,  
7000 Stuttgart 80

- KUHNERT , Klaus-Dieter, Dipl.-Ing.,  
Robert-Koch-Str. 14, 8012 Ottobrunn
- LEMNITZER , Lothar, Fahrtgasse 3-5,  
6900 Heidelberg
- MALCHOW , Carsten, Grosse Barlinge 43,  
3000 Hannover
- MARHENKE , Ralf, Dyroffstr. 3,  
5300 Bonn 1
- MEISBERGER , Rike, Apothekergasse 5,  
6900 Heidelberg
- MEYER , Ralf, Haseterwall 9,  
4500 Osnabrück
- MÜLLER-ZANTOP , Susanne, Mauerkir-  
cherstr. 29, 8000 München 80
- VON OETTINGEN , Edgar, c/o Oettingen  
GmbH, Friedrich-Ebert-Str. 27,  
4000 Düsseldorf
- ORTMANN , W.D., Dr., Goethe-Institut,  
Mailand, ITALIEN
- PRÖPPER , Ellen, Georg-Treber-Str. 59,  
6090 Rüsselsheim
- RHEIN , Ulrike, Roonstr. 34, 5400 Koblenz
- RIPP , Volker, Birkenstr. 12, 1000 Berlin 21
- RITZKE , Johannes, Nixdorf Computer  
AG, Abt. E 054, Pontanusstr. 55, 4790  
Paderborn
- ROTHKEGEL , Dr. Annely, Waldwiese 9,  
6600 Saarbrücken
- RUDOLF , Klaus, IABG/SzFF, Einstein-  
str., 8012 Ottobrunn
- RUGE , Nina, Volksgartenstr. 22,  
4000 Düsseldorf 1
- RYCHLY , Vladislav, Mühlheimer Str. 63,  
4300 Essen 1
- SAGAWA , Helmut, Obere Büttengasse 12,  
6900 Heidelberg
- SCHLÖGELL , Volker, Sternstr. 39,  
4400 Münster
- SCHMIDT , Arno, Manteifelstr. 6,  
4600 Dortmund 1
- SCHREIBER-SCHWENKGLA , Almut,  
M.A., Kleiststr. 1, 6200 Wiesbaden
- STEFFENS , Petra, Taunusstr. 72,  
7030 Böblingen
- STREHLITZ , Birgit, Am Gonsenheimer  
Spiess 87, 6500 Mainz
- THÖNSSSEN , Barbara, Frankenallee 228,  
6000 Frankfurt
- WAGNER , Franc, Postfach 10 61 51,  
6900 Heidelberg
- WESCHE , Birgit, Sulzauer Str. 17,  
7000 Stuttgart 80
- WIRTZ , Guido, Endenither Allee 110,  
5300 Bonn 1
- ZAMUROVIC-HELLER , Nada, Fraunho-  
fer Institut, Holzgartenstr. 17,  
7000 Stuttgart 1