

Dialogschnittstellen an
Online-Informationssystemen: Notwendigkeit,
Leistungsfähigkeit und
Entwicklungsmöglichkeiten am Beispiel des
OSIRIS-Systems

Dissertation zur Erlangung des philosophischen Doktorgrades
am Fachbereich Sprach- und Literaturwissenschaft
der Universität Osnabrück

vorgelegt von
Marc Ronthaler
aus Hamburg

Osnabrück, im Oktober 2000

Inhaltsverzeichnis

1	Einleitung	1
2	Bedeutung von Dialogschnittstellen	5
2.1	Wachsende Bedeutung von Online-Informationen	5
2.1.1	Änderung der Nutzerschichten	11
2.1.2	Bedarf an besseren Schnittstellen	12
2.2	Bibliotheken als exemplarische Anwendung	24
2.3	Fazit	41
3	Das Bibliothekssystem OSIRIS	44
3.1	Konzeption des OSIRIS-Systems	44
3.2	Aufbau und Funktion der Dialogschnittstelle	56
3.2.1	Überblick über die Systemkomponenten	56
3.2.2	Der Task Manager	56
3.2.3	Das Lexikon	58
3.2.4	Der Parser	61
3.2.5	Morphologie	64
3.2.6	Kompositazerlegung	65
3.2.7	Fehlerkorrektur	66
3.2.8	Semantik	67
3.2.9	Phonetik	68
3.2.10	Bewertung und Ergebnisübergabe	71
3.3	Verbesserung der Recherche durch OSIRIS	74
3.3.1	Die Retrievalmöglichkeiten im direkten Vergleich zum OPAC	75
3.4	OSIRIS im Kontext anderer Bibliothekssysteme	79
3.4.1	MILOS	79
3.4.2	CANAL/LS	82
3.4.3	VILIB	84
3.4.4	KASCADE	84
3.4.5	Die OKAPI-Projekte	87
3.5	Fazit	88
4	Eingabemöglichkeiten an Dialogschnittstellen	96
4.1	Dialogschnittstellen im Vergleich	96
4.1.1	Fazit	107
4.2	Linguistische Phänomene	109
4.2.1	Diskursstrukturen	109
4.2.2	Anaphorische Referenz	118
4.2.3	Ellipsen	122
4.2.4	Präsuppositionen	127
4.2.5	Implikaturen	137

4.2.6	(Indirekte) Sprechakte	139
4.2.7	Indirekte Reaktionen	141
4.2.8	Deiktische Ausdrücke	142
4.2.9	Kontext	142
4.2.10	Turn taking	143
4.2.11	Einschübe	144
4.2.12	Conversational fillers	144
4.2.13	Ambiguität	145
4.2.14	Disjunktive Verwendung von <i>und</i>	146
4.2.15	Quantoren	148
4.2.16	Komposita	149
4.2.17	Fazit	150
5	Verbesserungen und Erweiterungen von OSIRIS	153
5.1	Erweiterung der Möglichkeiten der Eingabeverarbeitung	153
5.1.1	Erweiterung der OSIRIS-Eingabemöglichkeiten	154
5.1.2	Erweiterung der linguistischen Komponenten in OSIRIS	161
5.1.3	Anfragebeispiele	166
5.1.4	Benutzerinteraktivität	172
5.1.5	Fazit	174
5.2	Weitere Verbesserungen von OSIRIS	176
5.2.1	Integration weiterer (heterogener) Datenbestände	176
5.2.2	Verbesserungen im Bereich des Interfaces	185
5.2.3	Unterstützung und Verbesserungen der Suchverfahren	197
6	Die Wirkung der OSIRIS-Schnittstelle auf den Benutzer	206
6.1	Psychologische Untersuchungsergebnisse zu OSIRIS	206
6.1.1	Fazit	215
6.2	Natürlichsprachliche Schnittstellen in der Literatur	219
6.2.1	Gibt es generelle Alternativen zu natürlicher Sprache ?	219
6.2.2	Natürliche Sprache im Vergleich zu formalen Sprachen	221
6.2.3	Linguistische Phänomene an NL-Schnittstellen	231
6.2.4	Wie einfach sind NL-Schnittstellen zu bedienen ?	235
6.2.5	Erleichterung und Steuerung der Bedienung	240
6.2.6	Benutzerverhalten an natürlichsprachlichen Schnittstellen	242
6.2.7	Fazit	246
7	Ausblick	249
7.1	Informationen zur Systemerweiterung	249
7.2	Computer als soziale Agenten	257
	Literatur	268

1 Einleitung

Die immer stärkere Verlagerung von Geschäftsprozessen ins Internet, die wachsende wirtschaftliche und gesellschaftliche Bedeutung von Information und der rasante Ausbau der Netze sorgen dafür, daß immer mehr Menschen in zunehmend heterogenen, unstrukturierten und verteilten Daten suchen. Immer mehr Nutzer dieses Informationsangebotes sind, z.B. im Wettbewerb auf dem Arbeitsmarkt, darauf angewiesen, online verfügbare Informationen selbständig für sich nutzen zu können. Dabei fehlt den allermeisten dieser auf gute Rechercheergebnisse angewiesenen Personen eine besondere Ausbildung im Information Retrieval.

Die Nutzung dynamischer und unübersichtlicher Datenbestände durch unerfahrene Personen verlangt aber nach einfach zu bedienenden Schnittstellen, die einen intuitiven und einheitlichen Zugriff auf die verteilt vorliegenden Daten ermöglichen.

In dieser Arbeit wird mit OSIRIS¹ eine Schnittstelle vorgestellt, die es sich in der beschränkten, aber repräsentativen Domäne der Online-Bibliothekskataloge zum Ziel gesetzt hat, die Einschränkungen der bestehenden Systeme zu überwinden, indem unerfahrenen Nutzern des Datenbestandes eine einfache Interaktionsmöglichkeit geboten wird. Zudem werden bessere Suchergebnisse durch ein anderes internes Suchverfahren erzielt.

OSIRIS wurde 1996 bis 1999 im Rahmen eines gemeinsamen Projektes der Universitätsbibliothek und des Instituts für Semantische Informationsverarbeitung der Universität Osnabrück (ISIV) mit finanzieller Unterstützung der Deutschen Forschungsgemeinschaft (DFG) entwickelt. Das Projekt wurde geleitet von Prof. Dr. Claus Rainer Rollinger (Direktor des ISIV) und Dr. Eilhardt Cordes (Bibliotheksdirektor). Verantwortlicher Projektleiter war Dr. Hartmut Zillmann (Universitätsbibliothek).

Ausgangspunkt des Antrages auf Förderung an das Bibliotheksreferat der DFG war die Einsicht, daß die Suchmöglichkeiten im Online-Katalog der Universitätsbibliothek einerseits stark verbesserungswürdig und andererseits auch deutlich verbesserungsfähig sind. Darüber hinaus wurde der Antrag von der Erkenntnis getragen, daß die in Osnabrück erkannten Probleme und auch deren mögliche Lösung Modellcharakter für andere Bibliotheken vergleichbarer Größe haben können.

Dabei stellte sich die Situation zu Beginn des Projektes wie folgt dar: Die Buchwerbungen der Universitätsbibliothek Osnabrück ab 1980 sind seit Anfang 1993 in einem Online-Katalog² recherchierbar (ca. 600.000 Datensätze). Seit 1996 stellt dieser – über eine VT100-Schnittstelle auch per Telnet im Internet verfügbare – Katalog, die einzige Möglichkeit dar, auf aktuellen Daten zu recherchieren. Neben den allgemein bekannten Schwierigkeiten mit der VT100-Oberfläche kam hinzu, daß der vorhandene Datenbestand (wie in anderen Bibliotheken auch) nicht durchgängig verbal erschlossen ist, d.h. nicht durchgängig mit Schlagwörtern versehen ist. Durch das

¹Osnabrück Intelligent Research Information System

²Online Public Access Catalog (OPAC)

Fehlen von Schlagwörtern ist ein Endbenutzer aber gezwungen, auf die Stichwortsuche auszuweichen.

Die Mehrzahl der Benutzer war also durch mangelnde Erschließung der Daten und inadäquate Gestaltung der Katalogoberfläche nicht in der Lage, sinnvoll nach Literatur zu recherchieren. Der Prozentsatz der Benutzer, die in der Lage sind, am OPAC bspw. eine Autor-Stichwort-Verknüpfung durchzuführen, wird von Bibliotheksangestellten auf unter 10 % geschätzt. Weiterführende Rechercheinstrumente wie Einschränkungen mit Erscheinungsjahren oder Materialarten (z.B. Zeitschrift) werden nicht einmal wahrgenommen.

Demzufolge werden für einen kleinen Prozentsatz qualifizierter Endbenutzer durchaus wichtige Indizes im Interesse einer für den Durchschnittsbenutzer überschaubaren Schnittstelle gar nicht erst angeboten.

Diese Situation wird nicht durch zukünftige Investitionen in eine bessere Erschließung der bibliothekarischen Daten verbessert werden können: Einerseits lassen die enormen Kosten eine solche Lösung, die erst in vielleicht 10 Jahren deutliche Retrievalverbesserungen erwarten lassen würde, illusorisch erscheinen. Andererseits fehlt diesem Lösungsansatz eine rückwärtige Perspektive für den Altbestand (das sind die Buchwerbungen bis zum Beginn der erschließungstechnischen Offensive).

Zielsetzung des OSIRIS-Projektes war es deshalb, auf der Basis des vorhandenen Daten- und Informationsbestandes eine Verbesserung der Recherchemöglichkeiten des Endbenutzers zu erreichen.

Nach dem Projektbeginn im Sommer 1996 konnte bereits im Frühjahr 1997 auf der CeBIT ein erster Prototyp vorgestellt werden. Im Herbst 1997 stand das OSIRIS-System allen Mitgliedern der Universität Osnabrück in der Version 1.1 zur Literaturrecherche zur Verfügung. Im April 1998 folgte dann die deutlich verbesserte Version 2.0, deren Nachfolger, Version 2.1, anlässlich der Präsentation des OSIRIS-Systems auf der Frankfurter Buchmesse 1998 für den weltweiten Zugriff geöffnet wurde.

Von weiterem bibliothekspolitischen Interesse ist, daß neben der beschriebenen Verbesserung der Literaturrecherche für Endbenutzer OSIRIS auch bei der Katalogisierung eingesetzt werden kann. Für neuerworbene Titel macht das System dem Bibliothekar einen Vorschlag, wie das betreffende Buch lokal klassifiziert werden soll. Statt einen hochspezialisierten Fachreferenten aufwendig über die Klassifikation eines neuerworbenen Buches befinden zu lassen, zieht das System Informationen aus der OSIRIS-Wissensbasis über den vorhandenen Bestand sowie eventuell vorliegende Fremddaten für das neu erworbene Buch heran, um dem Bibliothekar eine Handvoll Vorschläge zu unterbreiten, aus denen er den ihm korrekt erscheinenden auswählen kann.

Nach Ende der Förderung durch die DFG ist OSIRIS vollständig in den Alltagsbetrieb der Universitätsbibliothek integriert worden. Im Gegensatz zu vielen anderen Projekten wird es in der Praxis eingesetzt, aktiv gepflegt und weiterentwickelt. Aktuell (Oktober 2000) ist die Version 4.1 im Einsatz.

Seit Beginn des öffentlichen Betriebes wird das OSIRIS-System von Angehörigen des Fachbereiches Psychologie der Universität Osnabrück unter der Leitung von Dr. Kai-

Christoph Hamborg nach software-ergonomischen Gesichtspunkten und hinsichtlich der ISO 9241/10 evaluiert.

Diese Arbeit behandelt die in OSIRIS vorgenommenen Verbesserungen im Vergleich zu herkömmlichen Schnittstellen an Bibliothekskatalogen, die Bedeutung dieser Verbesserungen im Kontext des Bedarfs an besseren Schnittstellen allgemein und die Möglichkeiten eines weiteren Ausbaus des Systems.

In Kapitel 2.1 wird die steigende Bedeutung von Online Information ausgeführt und anhand einiger aktueller Beispiele der dringende Verbesserungsbedarf für Online-Dialogschnittstellen demonstriert.

2.2 belegt, daß Schnittstellen zu Online-Bibliothekskatalogen ein repräsentativer Sonderfall von Online-Dialogschnittstellen sind, anhand derer sich die generelle Problematik der Online-Recherche gerade auf textuellen Daten sehr gut zeigen läßt.

Kapitel 3 beschreibt OSIRIS konzeptionell (3.1) und im Detail aus Sicht der Entwickler (3.2). OSIRIS zeichnet sich durch zwei wesentliche Merkmale aus, die beschränkt natürlichsprachliche Eingabemöglichkeit und die automatische Optimierung des zu durchsuchenden Datenbestandes. Die Vorteile dieser Aspekte werden kontrastiv in Hinblick auf die Rechercheergebnisse (3.3) sowie in Hinblick auf andere Bemühungen im Kontext wissenschaftlicher Bibliotheken, die Recherche im Online-Katalog zu verbessern (3.4), beschrieben.

Um die Möglichkeiten von OSIRIS künftig weiter ausbauen zu können, werden in 4 die Eigenschaften verschiedener Typen von Dialogschnittstellen verglichen (4.1). Es wird der Versuch unternommen, relativ zu den möglichen Benutzereingaben eine Klassifikation von Dialogschnittstellen vorzunehmen (von einer passiven Schnittstelle über Menüstrukturen, beschränkter oder kontrollierter Eingabesprache bis hin zu freier Eingabe). Im Anschluß (4.2) wird ein Überblick über die bei einem Ausbau der Eingabemöglichkeiten an der Schnittstelle zu erwartenden linguistischen Phänomene gegeben.

Nach diesen allgemeinen Erwägungen zum Ausbau einer natürlichsprachlichen Schnittstelle werden in 5 für OSIRIS spezifische Möglichkeiten der Verbesserung und Erweiterung dargestellt.

Kapitel 6 widmet sich der Frage, wie die beschriebenen Verbesserungen durch OSIRIS vom Benutzer wahrgenommen werden. Dazu werden in 6.1 die vielfältigen Untersuchungen beschrieben, mit denen die Arbeits- und Organisationspsychologen der Osnabrücker Universität die Wirkung von OSIRIS auf den Benutzer untersucht haben. 6.2 geht auf die in der Literatur dargestellten Ergebnisse anderer Untersuchungen natürlichsprachlicher Schnittstellen ein, versucht Gemeinsamkeiten zu OSIRIS herauszuarbeiten und bislang von den Entwicklern nicht berücksichtigte Aspekte der Wirkung natürlichsprachlicher Schnittstellen aufzuzeigen. Natürlichsprachliche Schnittstellen haben seit jeher in Entwicklerkreisen einen hervorragenden Ruf und werden häufig als universelle Lösung für die am Anfang dieser Arbeit beschriebenen Probleme gesehen. Dennoch schneiden sie in eher benutzerorientierten oder psychologischen Untersuchungen schlechter ab als erwartet.

Kapitel 7 schließlich beschreibt auf der Grundlage der vorangegangenen Kapitel die möglicherweise limitierenden Faktoren eines weiteren Ausbaus einer natürlichsprachlichen Dialogschnittstelle wie OSIRIS: 7.1 widmet sich der Frage, woher die flächendeckend benötigten (linguistischen) Informationen kommen sollen, die für den Ausbau eines Systems wie OSIRIS benötigt werden. Schwierigkeiten bereiten hier die ungeheuer großen Datenmengen, die für einen Praxiseinsatz vonnöten sind und die jede manuelle Arbeit von vornherein ausschließen. In 7.2 wird noch einmal ein Punkt aus der Betrachtung der Literatur zur Wirkung von Dialogschnittstellen auf die Benutzer aufgegriffen. Hier geht es um die Frage, inwieweit bestimmte, an Menschen erinnernde Eigenschaften des Systems (z.B. Sprache zu verstehen) die Wahrnehmung der Maschine durch den menschlichen Benutzer verändern.

2 Bedeutung von Dialogschnittstellen

2.1 Wachsende Bedeutung von Online-Informationen

Seit einigen Jahren wird die Online-Verfügbarkeit von Information aus verschiedenen Gründen verstärkt von politischer Seite gefördert. Von Vertretern der Europäischen Union wurde wiederholt die Metapher verwendet, Information sei das Gold der Zukunft, auf das jeder Bürger Zugriff haben solle. Im Zuge einer aufkommenden weltweiten Informationsgesellschaft sei es wichtig, daß alle Bürger der Europäischen Union an dieser teilnehmen und von ihr profitieren könnten (Roukens, 1997). Hinter diesen Überlegungen steht die Erwartung, daß die allgemeine Zugänglichkeit von online-verfügbaren Dienstleistungen und Informationen europäische Unternehmen leistungsfähiger und flexibler im globalen Wettbewerb macht (Brinkhoff, 1997). Dies erscheint möglich durch direkten Zugriff der Unternehmen auf relevante Informationen und die stärkere Vernetzung der Unternehmen untereinander. So zitiert das Grünbuch der EU zu Informationen des öffentlichen Sektors den *Binnenmarktanzeiger* vom Oktober 1998: „58% der Unternehmen sind der Auffassung, daß der Zugang zu Informationen ihnen unter Umständen eine Ausweitung ihrer Tätigkeit innerhalb der Union ermöglichen würde. 66% der Firmen geben an, Bedarf an präzisen Informationen über Verwaltungsverfahren zu haben. 25% sind der Auffassung, daß das Fortbestehen von Schranken für Handel und Geschäftstätigkeit auf fehlende Informationen über EU-Vorschriften zurückzuführen sein könnte.“(Europäische Kommission, 1998, :5) Als Beispiel für benötigte Informationen nennt das Grünbuch den Zugang zu Patentinformationen:

„Schätzungen des Europäischen Patentamtes zufolge werden Jahr für Jahr über 18 Mrd. EUR für Forschungsarbeiten ausgegeben, die schon einmal durchgeführt wurden. Ein besserer Zugang zu Informationen über den Stand der Forschung könnte hier Abhilfe schaffen.“(Europäische Kommission, 1998, :6)

Auch Finanzinformationen sowie Informationen über das Rechtssystem benachbarter Länder könnten wettbewerbsentscheidend sein. Dabei geht es der EU-Kommission nicht darum, „daß die Mitgliedstaaten mehr Information produzieren, sondern daß die Daten, auf die die Allgemeinheit bereits zugreifen kann, klarer und für potentielle Interessenten leichter nutzbar werden.“(Europäische Kommission, 1998, :1) Dabei wird die Zahl der bereits vorhandenen, aber noch nicht online verfügbaren öffentlichen Datenbanken allein für die Niederlande mit etwa 36.000 angegeben.

Auch Brinkhoff (1997) betont die Verbesserung der Zugänglichkeit vorhandener Informationen und legt dabei den Akzent auf neu zu entwickelnde Schnittstellen. Wie Klavans und Schauble (1998) sieht auch Brinkhoff, daß diese Schnittstellen dazu beitragen sollten, daß nicht nur Benutzer mit Computer-Erfahrung und guten Englischkenntnissen Online-Angebote nutzen könnten.

Gesellschaftspolitische Ziele der Förderung von Online-Informationen und dem einfachen Zugriff auf diese sind u.a. die im EU-Vertrag garantierten Rechte der Bürger durch mehr Transparenz bei Informationen des öffentlichen Sektors zu stärken und die Europäische Integration sowie die Akzeptanz der politischen Strukturen der EU zu fördern (Europäische Kommission, 1998). Auch wird die Befürchtung geäußert, es könne ohne Anstrengungen der Politik zu einer Spaltung der Gesellschaft kommen in solche, die (online) Zugriff auf Informationen haben und solche, die hiervon nicht profitieren können.³

Ein weiteres erklärtes Ziel der EU-Kommission ist es, die öffentliche Verwaltung leistungsfähiger und bürgernäher zu gestalten. Derzeit sind bspw. in Deutschland nur wenige Prozent der Kommunen überhaupt im Internet vertreten. Hinzu kommt, daß sowohl in der Kommunikation zwischen Bürger und Verwaltung als auch innerhalb der Verwaltung selbst es offenbar vielfach der Fall ist, daß Informationen aus verschiedenen Datenbanken zusammengesucht werden müssen. Dies könnte durch die bessere Verfügbarkeit solcher Informationen in Online-Datenbanken und deren konsequente Nutzung vermieden werden. Dadurch würde auch verhindert, daß weiterhin Daten mehrfach erhoben werden. Allein dies würde nach Ansicht der EU-Kommission zu einer beträchtlichen Verringerung des Verwaltungsaufwandes für Bürger und Unternehmen führen. Mit gemeinsam genutzten und online verfügbaren Daten wäre es aber auch möglich und wünschenswert, Informationen und zugehörige Dienstleistungen gemeinsam anzubieten. So könnten nicht nur Informationsbedürfnisse schnell und mit nur einem Ansprechpartner befriedigt werden, sondern auch bspw. zugehörige Formulare für einen Verwaltungsakt mittels eines angegliederten Transaktionsdienstes ebenfalls online bereitgestellt werden. In diese Richtung zielte auch der Multimediawettbewerb *Media@Komm* des Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF), bei dem 1998 drei Kommunen (Bremen, Esslingen und Nürnberg) für ihre Internet-Präsenz prämiert und mit je 20 Millionen Mark für den weiteren Ausbau dieser Dienstleistungen ausgestattet wurden.

Mit der Förderung einher geht die wachsende Akzeptanz von Online-Information. Als Ergebnis einer mehrjährigen Untersuchung kommt das Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) (Einsporn, 1995) zu dem Schluß, daß in Deutschland nun auch kleinere und mittlere Unternehmen beginnen, Online-Datenbanken zu nutzen. Als Gründe hierfür wurden von den Teilnehmern der Untersuchung im Bereich Forschung und Entwicklung die Nutzung von Patentinformationen, kürzere Entwicklungszeiten und die Vermeidung von Fehlentscheidungen genannt. Daneben bestehe die Notwendigkeit, angesichts schneller Marktveränderungen und der Vereinheitlichung des europäischen Marktes Wettbewerbsvorteile durch bessere und schnellere Information zu erringen. Auch die Kostenreduktion ist ein wichtiger Faktor für den Einsatz von Online-Information: Die durchschnittliche Kosteneinsparung einer selbständig durchgeführten Recherche (im Vergleich zu einer

³Bundesaußenminister Fischer auf der Eröffnung der 4. Deutsch-Niederländischen Konferenz am 9.9.1999 laut einer DPA-Meldung vom selben Tag.

Recherche durch einen gewerblichen Rechercheur) wird von den Teilnehmern der Untersuchung mit 4.000 DM angegeben. Nach Ende der Untersuchung nutzen fast 90 % der Teilnehmer weiterhin regelmäßig Online-Datenbanken, um aktuelle betriebliche Fragestellungen zu klären.

Der wachsenden Akzeptanz steht in direkter Wechselwirkung ein verstärktes Angebot an Online-Informationen gegenüber. Dies äußert sich einerseits in einer steigenden Zahl von Angeboten. Etablierte Datenbanken wie *Genios*, *Juris*, *Roche*, das *Verzeichnis Lieferbarer Bücher (VLB)*, das *Deutsche und Europäische Patentamt*, Telefonbücher, Börsenkurse usw., Archive von Zeitungen und Zeitschriften sowie bestimmte Publikationen von Verlagen wie *Springer*, *Elsevier* oder *ACM Press* können im Internet online befragt werden. Andererseits aber müssen zum wachsenden Angebot auch diejenigen Quellen gerechnet werden, die firmeninterne Informationen für einen geschlossenen Benutzerkreis oder aber verkaufsfördernde Informationen z.B. in einem Internet-Kaufhaus zu bestimmten Produkten bereithalten. Die hier konstatierte wachsende wirtschaftliche Bedeutung von online-verfügbarer Information und der daraus resultierenden Untersuchung der Zugänglichkeit dieser Information mittels Schnittstellen beschränkt sich natürlich nicht auf die Angebote, bei denen Information als Handelsgut an sich vermarktet wird. Vielmehr betreffen den gesamten Wirtschaftsverkehr, der über das Internet abgewickelt wird, direkt oder indirekt die gleichen Fragen, die auch für Online-Datenbanken wichtig sind: Wie nämlich Informationen aktuell und in verständlicher Form aufbereitet unter einer einfach zu bedienenden Oberfläche möglichst ohne großen Aufwand schnell und präzise verfügbar gehalten werden kann. Die wachsende wirtschaftliche Bedeutung des Internet macht dabei die einfache und verlässliche Verfügbarkeit von Information zunehmend zum Wettbewerbsfaktor. Die Geschäfte, die derzeit unter dem schillernden Begriff *E-Commerce* subsumiert werden, erreichen in Deutschland nach Aussagen unterschiedlicher Untersuchungen 1999 Umsätze in zweistelliger Milliardenhöhe, wobei für die kommenden vier Jahre ein Zuwachs von mindestens 50 % pro Jahr erwartet wird. Die International Data Corp. (IDC) sagt für den europäischen Wirtschaftsraum insgesamt sogar jährliche Zuwachsraten von 138 % mit einem Umsatzziel von 430 Milliarden Dollar für 2003 voraus.⁴ Nach Aussagen der Bundesregierung werden im Bereich *Informationstechnologie* (ein Bereich, dessen Grenzen zwar nicht fest definiert sind, zu dem aber als ein wichtiger Bestandteil der Handel über das Internet gehört) bereits fünf Prozent des deutschen Bruttoinlandsprodukts erzeugt.⁵

Der eigentliche Handel mit Informationen in Form kostenpflichtiger Informationsdienste oder Datenbankrecherchen macht derzeit keinen nennenswerten Anteil an den Gesamtumsätzen im Internet aus. Kostenpflichtige Internet-Informationendienste wie *Genios*, *GBI*, *Reuters Business Briefing* oder *Handelsblatt Online* haben es schwer, sich gegen die aus den Anfangstagen des Internet herrührende Erwartungshaltung der Nutzer, Information sei ein kostenloses Allgemeingut („Information wants to be free“),

⁴www.heise.de/newsticker/data/cp-25.08.99-005/

⁵Bundeswirtschaftsminister Müller laut einer DPA-Meldung vom 27.8.1999

durchzusetzen. So sieht das Marktforschungsunternehmen Diebold im bloßen Verkauf von Informationen auch keine Perspektive. Informationen an sich würden über kurz oder lang gar nichts mehr kosten. Die einzige Möglichkeit, mit Informationen Geld zu verdienen, sei, sie individualisiert und früher als andere Anbieter dem Kunden zu liefern.⁶ Zielgerichtete und einfache Recherchemöglichkeiten und Lieferdienste werden offenbar den Handel mit Informationen bestimmen, nicht das bloße Angebot. Auch die EU-Kommission erwartet, daß der Trend, Daten des öffentlichen Sektors im Internet kostenlos bereitzustellen, Auswirkungen auf die Preisgestaltung anderer Anbieter haben wird.

Online-Verfügbarkeit von Information steht neben dem direkten Handel mit Daten auch in unmittelbarem Zusammenhang mit allen Aspekten des sogenannten E-Commerce. Revolutionär an der Nutzung des Internet (vor allem in seiner hypertext-basierten Form, dem sogenannten World Wide Web, kurz WWW) für die Geschäftsbeziehungen zwischen Unternehmen und ihren Kunden ist die Möglichkeit, die eigenen Produkte direkt, relativ leicht vergleichbar mit denen der Konkurrenz und vor allem interaktiv zu präsentieren. So ist es möglich, bislang nicht oder nur mit unverhältnismäßig hohem Aufwand erreichbare Märkte zu erschließen. Entscheidend dabei ist aber, daß die online verfügbaren Informationen über das Produkt auch den Käufer erreichen. Wihofszki (1999) zitiert den Leiter einer Untersuchung, in der 13.000 erfahrene Internet-Besucher zu Ihren Problemen beim Online-Einkauf befragt wurden: „Interessanterweise ist eines der markantesten Probleme das Internet selbst. Die unübersichtliche Struktur, das dynamische Chaos der Millionen Dokumente, die online zugänglich sind und an denen sich selbst schwergewichtige Suchmaschinen verschlucken, führt dazu, daß viele Internet-Nutzer sagen: Es ist gar nicht so einfach, die Produkte zu finden, die man sucht.“ So rangiert das Problem zu langwieriger Produktsuche in den Antworten auch auf Rang drei, gleich hinter der mangelnden Sicherheit beim Zahlungsverkehr und dem mangelnden Schutz persönlicher Daten. Auch der Handel mit materiellen Gütern (im Gegensatz zum Handel mit Informationen) beruht letzten Endes also auf der Verfügbarkeit von Online-Information.

Nach einem Bericht der Los Angeles Times (Piller, 1999) ist die Zeit, in der sich neue Angebote im Internet schnell durchsetzen konnten, bereits vorbei. Es seien bereits jetzt starke Konzentrationsprozesse im Gange, bei denen immer mehr Nutzer des Internet ihre Zeit bei immer weniger Online-Angeboten verbringen. So kommen die Analysten von Media Metrix zu der Einschätzung, daß, trotz vermutlich mehrerer hundert Millionen Seiten im WWW weltweit, im Juni 1999 in den USA fast 35 % der Online-Zeit auf den 50 meistbesuchtesten Sites des WWW verbracht wurde. Erstaunlich ist dabei die Beschleunigung des Prozesses: Vor 12 Monaten wurden laut Los Angeles Times noch 27 % der Online-Zeit auf den Top-50 Sites verbracht. Die Top-10 Sites konnten im Juni 1999 über 19 % der Online-Zeit für sich verbuchen (12 Monate zuvor waren es noch 16 %). Andere Untersuchungen kommen laut Piller (1999) zu noch drastischeren Ergebnissen: Hier ist von 21 % oder sogar 32 % der Online-Zeit die Rede, die bei

⁶Computer Zeitung Nr.50 vom 10.12.98

den Top-10 Sites verbracht werde. Unter diesen befinden sich die großen Portale (allen voran Yahoo), die im Umfeld der Suchmaschinen die Aufmerksamkeit der Benutzer auf ihr Angebot zu lenken versuchen. Nach den Ergebnissen der in Piller (1999) zitierten Untersuchungen haben zumindest in den USA die Internet-Nutzer also längst die Konsequenz aus der Unübersichtlichkeit des Netzes gezogen, indem sie überwiegend nur noch bestimmte Angebote frequentieren. Neue und bislang unbekannte Anbieter im WWW müssen daher stetig steigende Anstrengungen unternehmen, um wahrgenommen zu werden.

Neben den reinen Umsätzen durch Geschäftstätigkeit mit Endkunden oder zwischen Unternehmen, kommt dem Einsparpotential durch E-Commerce eine ebenso große Bedeutung zu: Der Wegfall von Zwischenhändlern, die direktere Rückkopplung von Hersteller und Lieferant und die Beschleunigung des Vertriebes werden allgemein als kostensparende Faktoren und damit als Wettbewerbsvorteil betrachtet. Diese Vorteile beruhen bei näherer Betrachtung aber auch auf der Online-Verfügbarkeit von Information. Die Beschleunigung der Geschäftsprozesse gründet sich wesentlich auf eine Beschleunigung des Informationsaustausches, der Wegfall von Vertriebsstufen ist nur möglich, weil Informationen zwischen Teilnehmern der Wertschöpfungskette präziser und schneller als bisher ausgetauscht werden können, und die Rückkopplung von Hersteller und Lieferant setzt die (teil)automatisierte Erhebung und Auswertung von Information aus den laufenden Geschäftsprozessen voraus.

All diese Informationen aber, die im Zuge eines durch E-Commerce geprägten Geschäftsganges erhoben und ausgetauscht werden, müssen in irgendeiner Form entweder sofort ausgewertet oder aber für eine spätere Auswertung gespeichert werden. Gespeicherte Informationen aber müssen recherchierbar sein, d.h. mit der wachsenden Bedeutung von E-Commerce entstehen große Datensammlungen, für die letztlich wieder die genannte Ausgangsproblematik gilt, wie nämlich in ihnen mit einer einfach zu bedienenden Oberfläche möglichst ohne großen Aufwand schnell und präzise gesucht werden kann.

Wie bereits beim E-Commerce gezeigt, hat auch das wachsende Angebot an online verfügbaren Datenbankschnittstellen und Auskunftssystemen in vielen Fällen eine tieferliegende Motivation im steigenden Kostendruck. So sind beispielsweise Kommunen, aber auch der gesamte Dienstleistungssektor gezwungen, langfristig über Möglichkeiten der Kosteneinsparung, zumeist im Personalbereich, nachzudenken. Hier bietet sich das Internet als Rationalisierungsmöglichkeit an. Das steigende Angebot gerade der Banken an Online-Transaktionsmöglichkeiten, ist somit immer auch als Vorbereitung flächendeckender Einsparungen zu verstehen. Was derzeit als Alternative zum herkömmlichen Geschäftsgang erprobt wird, kann in absehbarer Zeit bereits ausschließliche Interaktionsform oder einzige Verkehrsform ohne Aufpreis sein. Die Benutzerfreundlichkeit aktueller Online-Schnittstellen entscheidet damit nicht nur über den Zeitpunkt einer eventuell anstehenden Einsparung, sondern auch über den langfristigen Nutzen einer solchen Maßnahme. Denn eine solche (auf Marktmacht und Kostenvorteilen bauende) Umstellung des Kundenverkehrs ist nur dann erfolgreich, wenn es zu ihr später keine zwar teurere, aber für den Kunden deutlich attraktivere

Alternative, z.B. in Form eines Call Centers gibt. Daß dies ein dynamischer Prozeß ist, in dem immer wieder aus veränderten Kundenerwartungen eine neue Bewertung des vorhandenen Online-Angebots entsteht, belegen Beobachtungen, wie sie für die Internet-Präsentation von Karstadt von deren Pressesprecherin in Wihofszki (1999) berichtet werden: „Die Ansprüche der Kunden sind in den vergangenen drei Jahren rasant gestiegen. Gefragt ist nunmehr auch Lifestyle als Einkaufskonzept. Das heißt, im Internet müssen vermehrt Zusatzwerte wie Entertainment und Information angeboten werden.“(Wihofszki, 1999) Verkennen oder ignorieren Online-Anbieter eine solche Entwicklung, beispielsweise mit einer Schnittstelle, die an den Kundenwünschen vorbeigeht, haben wieder jene Zwischenhändler oder Alternativenanbieter eine Chance, die, obwohl teurer, aus Sicht des Kunden über einen geldwerten Zusatznutzen verfügen. Über die bloßen Zahlen zu den erwarteten Umsatzsteigerungen im Online-Handel hinaus kann als Grundlage für eine Einschätzung der längerfristigen Entwicklung von Online-Informationssystemen die Meinung der in der Delphi-Studie'98 (FhG-ISI, 1998) zu diesem Thema befragten Experten dienen. Im Rahmen der Delphi-Studie'98 wurden 1997/98 im Auftrag des Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie (BMBF) bundesweit über 2.000 Experten aller Wissensgebiete zu den Realisierungschancen von insgesamt mehr als 1.000 Visionen befragt. Diese Visionen wiederum waren zuvor von über 100 Experten aus Forschung und Lehre sowie der Industrie erarbeitet worden. Die Experten, die über die in Frage stehenden Visionen zu urteilen hatten, gehen mehrheitlich von einer deutlichen Zunahme der Informationssuche und des Wissenserwerbs über das Netz aus. Im einzelnen erwarten sie, daß

- bis zum Jahr 2002 der Handel in Netzwerken durch elektronische Abrechnungs- oder Geldsysteme weit verbreitet sein wird. Diese Einschätzung deckt sich mit den anderen bereits angeführten Prognosen zur weiteren Entwicklung des E-Commerce.
- bis 2003 sich ein System durchgesetzt hat, das verschiedene überregionale Telemetrie-Informationen für Wettervorhersage, Überwachung der Umwelt und des Straßenverkehrs usw. per Satellit sammelt und zur Verfügung stellt. Neben den Fällen, in denen solche Systeme z.B. für ein fahrendes Auto den entsprechend der aktuellen geographischen Position relevanten Ausschnitt der Informationen automatisch über einen definierten Kanal zur Verfügung stellt, sind zahlreiche andere Situationen vorstellbar, in denen ein Benutzer interaktiv und individuell diese Datenquellen befragen will.
- bis 2004 sich Strukturen im Bildungssystem durchgesetzt haben, die individuelle Anpassungsfähigkeit und wettbewerbliche Dynamik unterstützen. Weiterhin erwarten die Experten, daß bis 2007 ein Fernunterrichtssystem zur Aus- und Weiterbildung der Bevölkerung von zu Hause aus allgemeine Verwendung findet und bis 2008 virtuelle Prüfungen und umfassende Leistungsbeurteilungen

über öffentliche Netze möglich sind. Diese Entwicklungen setzen auf jeder Stufe natürlich einen adäquaten und konkurrenzfähigen Zugang zu den in Frage stehenden Lehrinhalten voraus. Dies gilt umso mehr für die Prognosen, daß bis 2009 virtuelle Weltuniversitäten weit verbreitet sein werden und bis 2010 für die Mehrheit der Arbeitnehmer die Nutzung von Bildungsleistungen mit Hilfe der Telekommunikation vollständig in den Ablauf der beruflichen Arbeitszeit integriert ist.

- bis 2005 die allgemeine Büroarbeit, ausgenommen Besprechungen und Verhandlungen, unter Nutzung von E-Mail, WWW und anderen Netzdiensten überwiegend von zu Hause aus erledigt wird. Auch bei verschärften Rahmenrichtlinien für den Datenschutz ist für eine solche Entwicklung das Entstehen größerer Informationspools zu Informations- und Dokumentationszwecken zwangsläufig. Solche Datenbestände sind aber wiederum nicht reine Ablagen, sondern müssen der schnellen Wiederauffindbarkeit von Informationen dienen.
- bis 2012 alle elektronisch verfügbaren Informationen in den gängigsten Welt-sprachen vorliegen.

In den kommenden 10 bis 12 Jahren ist also mit einem steigenden Bedarf an Lösungen für Online-Informationendienste zu rechnen.

2.1.1 Änderung der Nutzerschichten

Bei wachsenden Nutzerzahlen für Online-Angebote wächst zugleich die Zahl der „naiven“ Nutzer, d.h. innerhalb der Zielgruppe von Online-Diensten wird der Anteil der Nutzer ohne einschlägige Erfahrungen im Information Retrieval (IR) immer größer. Gleichzeitig ändert sich die Einstellung der Nutzer gegenüber den Retrievalsystemen: je selbstverständlicher der Umgang mit dem Internet wird, desto weniger sind Benutzer bereit, die Schwierigkeiten im Umgang mit Schnittstellen sich selbst und ihren mangelnden Kenntnissen anzulasten. Verfügbarkeit von Informationen und Bedienbarkeit von Schnittstellen wird in den Augen der Nutzer immer stärker zu einer Bringschuld der Anbieter. Dies ist aber nicht nur eine neue Anforderung der Nutzer an die Anbieter, vielmehr versuchen Online-Informationsanbieter neue Nutzerschichten zu erschließen und umwerben dabei auch direkt eine Zielgruppe, die bislang keine Erfahrungen mit Online-Informationsangeboten hat. Dies gilt umso stärker vor dem Hintergrund des oben angeführten Kostendrucks. Die konstatierte Änderung der Nutzerschichten wird durch die beginnende Verschmelzung der Medien wie z.B. Fernsehen, Hörfunk und Internet noch verstärkt. Abgesehen von der Vermischung der Inhalte werden dabei natürlich medientypische Charakteristika übernommen, so zum Beispiel erste Möglichkeiten der Personalisierung des Fernsehens durch einen Rückkanal. Umgekehrt beeinflusst aber auch das Fernsehen die Gestaltung des WWW und den Umgang mit demselben. So stellt Lieberman (1995) mit Letizia ein „Zapping-Interface für das WWW“ vor (siehe auch 5.2.3). Nicht alle Web-Inhalte müssen in Konkurrenz zum Fernsehen (oder

dessen Anteil an der Verschmelzung der Medien) stehen. Aber wo diese Konkurrenz besteht, sind die Web-Seiten in den meisten Fällen noch zu kompliziert.

Wihofszki (1999) gibt in diesem Zusammenhang für die bereits erwähnte Befragung von über 13.000 Internet-Besuchern, bei der 38,8 %⁷ eine zu langwierige Produktsuche als Problem nannten, zu bedenken, daß nur Personen mit relativ viel Erfahrung befragt wurden. Wenn aber bereits der geübte Online-Käufer Probleme habe, was könne dann von einem Internet-Laien erwartet werden ?

Die Kritik der Teilnehmer der BMBF-Untersuchung (Einsporn, 1995) geht in eine ähnliche Richtung: Kritisiert wurde von diesen Personen, daß nur ausgebildetes Personal in den Unternehmen in der Lage war, die oben erwähnten Einsparungen zu realisieren. Insbesondere dieses ausgebildete Personal wiederum bemängelte dabei die Vielzahl und Heterogenität der existierenden Datenbankabfragesprachen und Datenbankstrukturen. Dieser Punkt wurde übereinstimmend als großes Hemmnis bei der unternehmensweiten Einführung der Datenbanknutzung betrachtet. Daneben wurden Fremdsprachen und verwirrende Ergebnispräsentation als problematisch angesehen. Generell gilt nach dieser Untersuchung auch für die erfahrenen Anwender: je höher die Qualifizierung und Einarbeitung, desto größer der wirtschaftliche Wert der elektronisch recherchierten Informationen. Darüber hinaus stellt die BMBF-Untersuchung fest, daß Unternehmen die zunächst entstehenden Kosten akzeptieren, wenn die Qualität der gefundenen Information überzeugt.

Weiterhin geben die Autoren der BMBF-Untersuchung zu bedenken, daß die Öffnung der Datenbanken insbesondere für Gelegenheitsnutzer, wie es Angehörige der Geschäftsleitung zumeist sind, nicht nur die innerbetrieblichen Voraussetzungen für die häufig geforderte Vernetzung aller Informationsquellen darstellt. Darüber hinaus würden durch die Integration der Geschäftsleitung auch die zumeist kostenintensiven innerbetrieblichen Akzeptanzprobleme der Datenbankeinführung entfallen.

2.1.2 Bedarf an besseren Schnittstellen

Vor diesem Hintergrund kommt den Schnittstellen, mit denen die Bürger und Unternehmen die Informationen suchen sollen, die sie nach eigenem Bekunden und dem der EU-Kommission brauchen, eine herausragende Stellung zu (Europäische Kommission, 1998, :1,;14). Die derzeit im Einsatz befindlichen Benutzerschnittstellen lassen die dringend notwendige Vereinfachung der Recherche aber noch vermissen. Um diese Schnittstellen bedienen zu können, bedarf es immer noch eines erheblichen intellektuellen Aufwandes bei der Konstruktion der Anfrage.

Der nötige Aufwand wird mancherorts explizit angesprochen, wie z.B. auf der Seite des Europäischen Patentamtes (Abbildung 1). Bemerkenswert ist, daß das Angebot des Europäischen Patentamtes sich ausdrücklich nicht an Rechercheprofis, sondern an kleine und mittlere Unternehmen richtet, denen hier die einfache Suche in über 30 Millionen Patentdokumenten ermöglicht werden soll.

⁷Mehrfachnennungen waren möglich.

Daß dies trotz der an sich sehr einfachen Interaktionsform (es steht nur eine einzige Zeile zur Verfügung, um Eingaben zu tätigen) sehr schwierig sein wird, macht die zugehörige Anleitung deutlich.


Angesichts der vorgeschlagenen Sucheingaben („T_1022/92“) erscheint es unmöglich, ohne detaillierte Kenntnisse der Struktur der Daten erfolgreich zu Suchen. Inwieweit das im allgemeinen Fall überhaupt erreichbar ist, wird in 6.2 eingehender diskutiert. Hier ist jedenfalls festzuhalten, daß eine Schnittstelle wie die in Abbildung 1 gezeigte ihr Ziel, einen einfachen Zugang zu Informationen zu bieten, verfehlt.

Search interface for the Decisions of the EPO Boards of Appeal

The search engine processes the comprehensive text of each decision of the EPO Boards of Appeal contained in the database (despatched after 1st of January 1997). The decisions are always available in the language of the proceedings (English or French or German).

>> [General information on data coverage and updates.](#)

To search simply enter words and phrases into the search field.

search terms: search 

Search tips

- Searches for case numbers have the following syntax: "T_1022/92" or "T_1022/92 <NEAR> 3.2.2" or "T_0012/*"
- Your search terms do not need to be complete words: "comput*", for example, will match occurrences of computer and computers etc. (bi-lateral truncation is supported)
- Boolean operators are supported [AND, OR, NEAR]
- Phrases (if containing no parentheses!) can be searched by simply entering the terms **without** apostrophes, e.g. *inventive step*
- If you want to have implicit stemming disabled, please use quotation marks around your term, e.g. "remittal" will NOT display documents containing words like remitted etc.
- for more detailed information and/or questions, please select the "help" button on the left side

Results

- 1 The first level result consists of the hit-list containing the links to the document matching your query
- 2 Selecting a document out of the hit-list will open the corresponding bibliographic data in a **separate**
- 3 The full text (and images) of the decision is available in mixed mode PDF. You need to have the freeware [Acrobat Reader 3.x](#) installed on your work-station.

General information

- The collection of the T-decisions of the EPO's technical boards of appeal dispatched to the parties after 1st January 1997 are now available free of charge on this site
- The full text of the decision is searchable and can be viewed and downloaded in Adobe Portable Document Format (PDF). The freeware Adobe Acrobat reader is therefore required (www.adobe.com).
- Links to the International Patent Classification (IPC) and to the European Patent Convention (EPC) articles and rules relevant for each decision are also available.
- The G-, J- and W-decisions as well as T-decisions dispatched before 1997 will be made available step by step in the next months.
- We look forward to receiving [your comments](#) on this new enhancement on European Patent Office homepage.

[EPO Home Page](#) | [Index](#) | [Comments](#)

Copyright © 1998 European Patent Office . All Rights Reserved.

Abbildung 1: Esp@cenet: Europäisches Patentamt

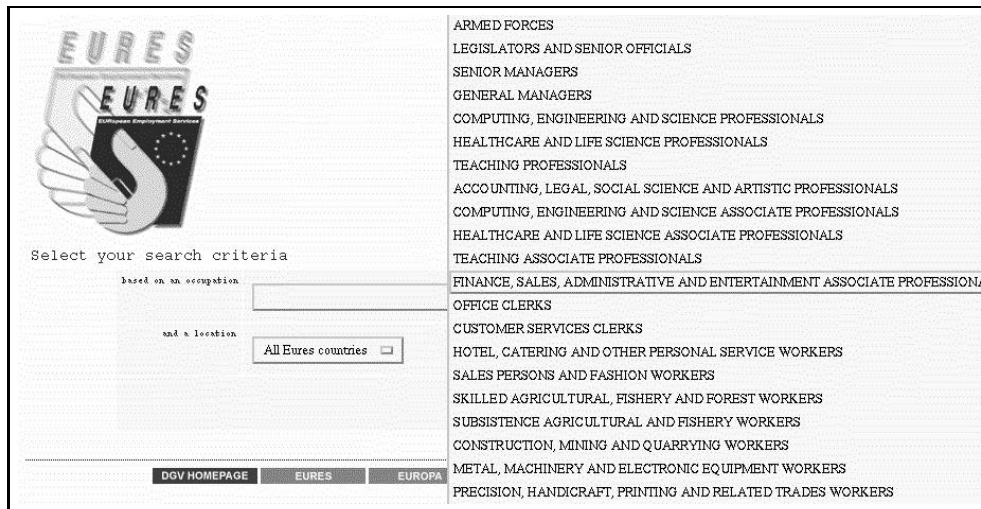
Andere Anbieter, wie z.B. das Deutsche Patentblatt (Abbildung 2), setzen Ihre Kunden direkt einer komplexen Suchmaske aus:

einsteiger experte assistent			
Dokumentennummer	<input type="text"/>	Publikationswoche	<input type="text"/>
Schriftart	<input type="text"/>		
Hauptklasse (IPC)	<input type="text"/>	Nebenklasse	<input type="text"/>
Anmelder	<input type="text"/>		
Erfinder	<input type="text"/>		
Vertreter	<input type="text"/>		
Titel	<input type="text"/>		
Zusammenfassung	<input type="text"/>		
Teil/Abschnitt	<input type="text"/>	Grund	<input type="text"/>
Prioritätsnummer	<input type="text"/>		

Abbildung 2: Maske des Deutschen Patentblattes

Der Suchende hat nicht mehr nur ein Feld für beliebige Eingaben zur Verfügung, sondern verschiedene, durch Bezeichner in ihren Aufgaben beschriebene Felder. Durch den Aufbau der Seite, so offenbar die Idee der Entwickler, ist die in Abbildung 1 nötige ausführliche Benutzungserklärung überflüssig geworden. Dennoch wird beim ersten Hinsehen bereits deutlich, daß auch die Seite des Deutschen Patentblattes explizites Wissen über die zugrundeliegende Datenstruktur voraussetzt. Es darf bezweifelt werden, daß dies dem eigenen Anspruch, eine Recherchemöglichkeit für Suchende ohne Vorkenntnisse („Einsteiger“) zu sein, gerecht wird.

Einen dritten Weg gehen Anbieter, die eine Reihe von Kategorien vorgeben, innerhalb derer der Benutzer seine Frage einzuordnen hat. Statt es dem Benutzer wie in Abbildung 2 zu überlassen, einzelnen Attributen einer Datenbankanfrage explizit oder durch Übernahme eines Defaults einen Wert zuzuweisen, in der Hoffnung, daß die Kombination eine nicht-leere Extension besitzt, wird eine aus Sicht des Datenanbieters sinnvolle endliche Kategorienliste vorgegeben. Diese Technik der Vorgabe einer Kategorienliste ist z.B. auf der EURES-Seite zu sehen, die Arbeitssuchenden Stellenangebote in den verschiedenen EU-Ländern sowie zugehörige Informationen über das betreffende EU-Land vermitteln soll:

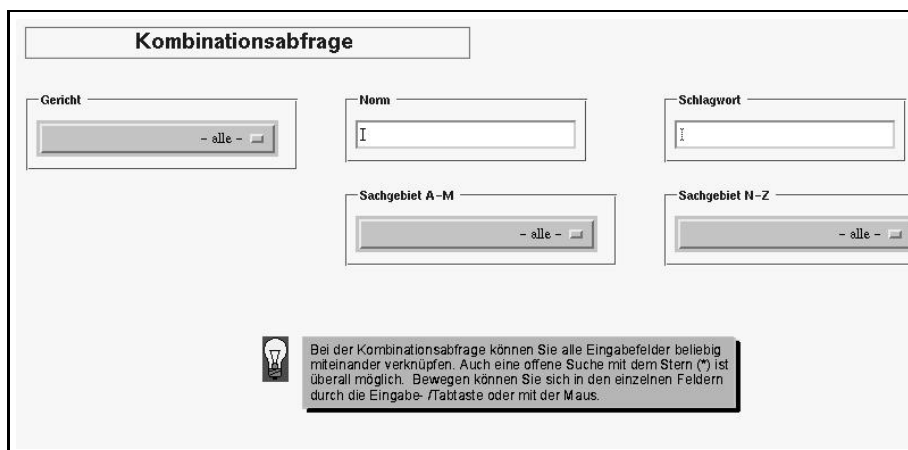


The screenshot shows the EURES website interface. On the left, there is a logo for EURES (European Employment Services) and a section titled "Select your search criteria". Below this, there are two input fields: "Based on an occupation" and "and a location". The "and a location" field has a dropdown menu with "All Eures countries" selected. At the bottom left, there are three buttons: "DGV HOMEPAGE", "EURES", and "EUROPA". On the right side, there is a vertical list of professions, including: ARMED FORCES, LEGISLATORS AND SENIOR OFFICIALS, SENIOR MANAGERS, GENERAL MANAGERS, COMPUTING, ENGINEERING AND SCIENCE PROFESSIONALS, HEALTHCARE AND LIFE SCIENCE PROFESSIONALS, TEACHING PROFESSIONALS, ACCOUNTING, LEGAL, SOCIAL SCIENCE AND ARTISTIC PROFESSIONALS, COMPUTING, ENGINEERING AND SCIENCE ASSOCIATE PROFESSIONALS, HEALTHCARE AND LIFE SCIENCE ASSOCIATE PROFESSIONALS, TEACHING ASSOCIATE PROFESSIONALS, FINANCE, SALES, ADMINISTRATIVE AND ENTERTAINMENT ASSOCIATE PROFESSIONALS, OFFICE CLERKS, CUSTOMER SERVICES CLERKS, HOTEL, CATERING AND OTHER PERSONAL SERVICE WORKERS, SALES PERSONS AND FASHION WORKERS, SKILLED AGRICULTURAL, FISHERY AND FOREST WORKERS, SUBSISTENCE AGRICULTURAL AND FISHERY WORKERS, CONSTRUCTION, MINING AND QUARRYING WORKERS, METAL, MACHINERY AND ELECTRONIC EQUIPMENT WORKERS, and PRECISION, HANDICRAFT, PRINTING AND RELATED TRADES WORKERS.

Abbildung 3: EURES: Stellenangebote in Ländern der EU

Die Probleme, die mit einem solchen Ansatz verknüpft sind, stellen sich in allen Bereichen, in denen eine Taxonomie vorgegeben und vom Benutzer verlangt wird, sich und sein Anliegen dort einzuordnen, z.B. bei Branchenbüchern, in Behördenformularen sowie bei allen Arten der statistischen Erfassung. Neben der mangelnden Flexibilität der Kategorienbezeichnungen und der fehlenden Dynamik des Kategorienschemas ist es schwierig, die Balance zu halten zwischen einem kleinen und übersichtlichen Schema, das für ein komplexes Anwendungsgebiet eine relativ hohe intellektuelle Leistung des Benutzers zur Eigenpositionierung verlangt und einem breiten und fein differenzierten Schema, in dem der Suchende sich leicht einordnen kann, allerdings um den Preis eines hohen Suchaufwandes.

Dabei sind natürlich Mischformen möglich, wie sie z.B. das Oberverwaltungsgericht Rheinland-Pfalz mit seiner Entscheidungssammlung praktiziert:



The screenshot shows the "Kombinationsabfrage" (Combination Search) interface. It features four search criteria: "Gericht" (Court), "Norm" (Norm), "Schlagwort" (Keyword), and "Sachgebiet A-M" (Subject Area A-M). Each criterion has a dropdown menu with "- alle -" (all) selected. Below the search criteria, there is a lightbulb icon and a text box that reads: "Bei der Kombinationsabfrage können Sie alle Eingabefelder beliebig miteinander verknüpfen. Auch eine offene Suche mit dem Stern (*) ist überall möglich. Bewegen können Sie sich in den einzelnen Feldern durch die Eingabe-/Tabtaste oder mit der Maus." (In the combination search, you can link all input fields arbitrarily. An open search with the asterisk (*) is also possible everywhere. You can move in the individual fields by pressing the input/tab key or with the mouse.)

Abbildung 4: Entscheidungssammlung des OVG Rheinland-Pfalz

Als Ergänzung wird häufig der alphabetische Index angeboten. Dahinter steht die Vorstellung, daß ein Benutzer statt einer inhaltlich orientierten Einordnung seines Suchbegriffes in ein Kategorienschema lieber anhand der äußeren Form überprüft, ob sein Suchbegriff im Kategorienschema vorkommt. Wird der Suchbegriff dort gefunden, dann muß dazu kein inhaltliches Verständnis der Begriffshierarchie vorhanden gewesen sein. Wird der Suchbegriff nicht gefunden, so wird vom Benutzer erwartet, daß er selbständig Synonyme bildet und testet oder aber im angebotenen Wortmaterial etwas für ihn passendes entdeckt. Allerdings zeigen Untersuchungen zu Online-Bibliothekskatalogen, daß die Verwendung solcher Indizes genau durchdacht sein muß, wenn man Abwehrreaktionen der Nutzer vermeiden will (siehe auch 2.2, Seite 35).

Das Fraunhofer Institut bietet bereits auf der Einstiegsseite dem Suchenden die Möglichkeit, sich zwischen einer freien Volltextsuche und einer Suche im alphabetischen Index zu entscheiden:



Abbildung 5: Einstiegsseite des Fraunhofer Institutes

Nach Anwahl der Stichwortsuche erfolgt eine umfangreiche Einführung in den Umgang mit der gewählten Suchoption (Abbildung 6). Die Auswahl eines Buchstabens per Mausklick führt dann in den eigentlichen Index, der einen alphabetischen Einstieg mit einer rudimentären Begriffshierarchie kombiniert.

Stichwortsuche

Stichwortsuche in den Profilen der Fraunhofer-Institute
(Forschungsgebiete, Leistungen, Ansprechpartner)

Die vorliegende Stichwortliste soll Ihnen helfen, sich einen Überblick über die vielfältigen Forschungsgebiete der Fraunhofer-Institute zu verschaffen und direkt den zuständigen Ansprechpartner zu finden. Klicken Sie hierzu einfach auf die bei den Stichwörtern befindlichen Verweise (hypertext links). Die kursiv gedruckten Stichwörter bezeichnen besondere Versuchseinrichtungen oder Geräte, die in den Instituten zur Verfügung stehen.
Zur Wahrung der Übersichtlichkeit gibt es für jeden Anfangsbuchstaben ein eigenes Stichwortverzeichnis.

Bitte wählen Sie den Anfangsbuchstaben Ihres Suchbegriffs:

A B C D E F G H I J K L
M N O P Q R S T U V W X Y Z

© 1999
Fraunhofer-Gesellschaft

Abbildung 6: Stichwortsuche des Fraunhofer Institutes

Auf den Seiten des Deutschen Arbeitsamtes wird dagegen wie bereits beim Deutschen Patentblatt (Abbildung 2) darauf gesetzt, daß durch die Vorgaben im Aufbau der HTML-Maske der Benutzer seine Anfrage entsprechend strukturiert und so eine Datenbankabfrage vorbereitet:

Stellen Informations-Service **sis**

Suchen

Beruf

Anzeige der Berufsbezeichnung in weiblicher Schreibweise männlicher Schreibweise

Tätigkeit Berufliste herunterladen

BKZ (falls bekannt)

Arbeitszeit Vollzeit Teilzeit Heimarbeit

Stellen der letzten wochen

Wirtschaftsraum

Regionen

Postleitzahl

Anzuzeigende Angebote

Anzahl der Treffer 0

Abbildung 7: Suchmaske des Deutschen Arbeitsamtes

Nach diesem Prinzip, nämlich über eine strukturierte Maske den Suchenden zur Identifikation der formalen Teile seiner Anfrage zu bewegen, funktionieren eine ganze Reihe von Schnittstellen, z.B. auch die Eingabemaske des Deutschen Patentamtes (siehe Abbildung 8).

DEPanet
Patentserver des Deutschen Patent- und Markenamts

DPMA Patentblatt
Deutsches Patentblatt - Online Schriften bestellen

Recherche in deutschen Veröffentlichungen

- Bitte geben sie die suchbegriffe in die Maske ein.
- Wenn sie mehrere Eingabefelder ausfüllen, werden die suchbegriffe mit den Booleschen "AND" verknüpft.
- Hilfe zu den einzelnen Eingabefeldern erhalten sie, wenn sie auf die jeweiligen Feldbeschriftungen klicken.
- allgemeine Hilfe erhalten sie, wenn sie den HILFE-Button anklicken.
- Die hier dargestellten Daten geben den aktuellen stand der Datenbank DEPanet des Deutschen Patent- und Markenamts wieder. Bei der Fülle der Daten sind Fehler und Lücken jedoch nicht vollständig auszuschließen. Haftungsrechtliche Ansprüche jeglicher Art sind ausgeschlossen. [\(weitere Hinweise\)](#)

Veröffentlichungsnummer	<input type="text" value="I"/>	z.B.: DE19750978
Aktenzeichen	<input type="text"/>	z.B.: DE19971050978
Prioritätsaktenzeichen	<input type="text"/>	z.B.: DE19971050978
Veröffentlichungsdatum	<input type="text"/>	z.B.: 19980402
Anmelder	<input type="text"/>	z.B.: Mayr
Erfinder	<input type="text"/>	z.B.: mueller
Klassifikationssymbol (IPC)	<input type="text"/>	z.B.: F16B5/06
Titel	<input type="text"/>	z.B.: Fahrrad or Rahmen

Suchen Löschen

Abbildung 8: Suchmaske des Deutschen Patentamtes

Ein Nachteil dieses Verfahrens, das auch auf kommerziellen Seiten häufig anzutreffen ist, liegt in der Begrenzung auf die in der Seitenstruktur vorgegebenen Suchmerkmale. Es ist also wie in vielen anderen Bereichen der Programmgestaltung (z.B. Sprachlehrsystemen) die Antizipationsleistung des Programmierers, die über die funktionale Angemessenheit der Suchmöglichkeiten entscheidet. Dabei kann der intellektuelle Aufwand, die Möglichkeiten, die die Schnittstelle zur gezielten Suche bietet, mit dem eigenen Informationsbedürfnis in Einklang zu bringen, sehr hoch sein. Kunisch (1999) beschreibt dies sehr eindrücklich anhand einer Studie zu Online-Stellenangeboten in den USA. Dieser Studie von Nielsen Norman zufolge scheiterten 42 % der untersuchten Bewerber⁸ vollständig bei dem Versuch, sich online auf eine online ausgeschriebene Stelle zu bewerben. Als Grund für das Scheitern wird eine derartig schlechte Bildschirmgestaltung genannt, die jeden Versuch der Kontaktaufnahme verhinderte. Weiteren 32 % gelang es zwar, mit dem Unternehmen Kontakt aufzunehmen, scheiterten jedoch im weiteren Verlauf der Bewerbung, weil sie aufgrund unverständlicher Berufsbezeichnungen sich auf die falsche Stelle bewarben, irrtümlich falsche Angaben im Online-Dialog machten oder aber der Online-Dialog an sich zusammenbrach und nur über andere Kanäle (Brief, Telefon) ein weiterer Kontakt möglich war. Lediglich 26 % der Versuchspersonen gelang es, den gesamten Vorgang der Online-Bewerbung erfolgreich abzuschließen. Kunisch hält die Ergebnisse dieser Studie für durchaus übertragbar auf deutsche Verhältnisse und führt in diesem Zusammenhang explizit die Seite des Arbeitsamtes an, auf der gleich an dritter Position nach der "Berufsgruppenkennzahl" (BKZ), einer internen Kennziffer, gefragt wird (siehe Abbildung 7). Den intellektuellen Aufwand, das eigene Informationsbedürfnis auf die Möglichkeiten

⁸ Alle Probanden waren mit dem WWW vertraut.

der Schnittstelle abzubilden, demonstriert auch die Maske der Internetsuchmaschine FAST sehr eindrücklich (Abbildung 9).

Abbildung 9: Maske der Internetsuchmaschine FAST

Man muß sich vor Augen führen, daß graphische Schnittstellen zu Datenbanken noch nicht sehr alt sind. Neben den Strategien der HTML-basierten Seiten, die Benutzbarkeit zu erleichtern, gibt es eine sehr viel längere Tradition des an der Kommandozeile orientierten Datenbankzugriffs. Hier wurde und wird überwiegend mit Booleschen Operatoren, regulären Ausdrücken, Programmiersprachen oder mit Sprachen wie SQL gearbeitet, die speziell für die Abfrage (relationaler) Datenbanken entwickelt wurden. Gerade Boolesche Operatoren und auch reguläre Ausdrücke sind dabei für WWW-Schnittstellen immer noch im Einsatz. Im simpelsten Fall sind es die verschiedenen Eingabefelder einer Suchmaske, die implizit durch AND oder OR verknüpft sind. Viele Datenbankschnittstellen bieten darüber hinaus explizit den Gebrauch Boolescher Operatoren an, und an den allermeisten Eingabeschnittstellen mit Volltextretrievalangebot gibt es die Möglichkeit, wenigstens mit Trunkierung, wenn nicht sogar mit regulären Ausdrücken zu arbeiten. Auf diese Weise werden die Suchmöglichkeiten des Nutzers erheblich erweitert, weil er nun nicht mehr auf die explizit in der Struktur der Eingabemaske angelegten Möglichkeiten festgelegt wird. Insbesondere kann im Gegensatz zu den oben erwähnten Möglichkeiten, das eigene Informationsbedürfnis in einer vorge-

gebenen Begriffshierarchie inhaltlich oder über einen alphabetischen Stichwortindex aufgrund der äußeren Form auf den Datenbestand abzubilden, nun auch mit Teilen eines Ausdrucks wie Präfixen oder Suffixen gesucht werden. Der Suchende bekommt so sehr mächtige Möglichkeiten an die Hand, seine Anfrage zu formulieren, und kann durch eine wohl durchdachte graphische Schnittstelle zudem erhebliche Vorteile hinsichtlich der Benutzbarkeit erwarten. Deshalb stellt die Kombination aus einer WWW-Oberfläche und der Verwendung Boolescher Operatoren für viele auch die optimale Voraussetzung für eine einfache Suche in Datenbanken dar.

Dem steht die Erkenntnis gegenüber, daß Boolesche Operatoren und reguläre Ausdrücke für Datenbankabfragen zwar mächtig, aber für eine Suche auf *sprachlichen* Daten keineswegs besonders gut geeignet sind (Turtle, 1994). Ein großer Teil der online verfügbaren Information wird so nicht oder nicht adäquat erreicht. Detailliert wird dieser Punkt anhand der Online-Bibliothekskataloge (2.2, Seite 35) erläutert werden. Zudem haben Benutzer offenbar große Probleme mit der Verwendung Boolescher Operatoren (Peters, 1989; Borgman, 1986; Frei und Qiu, 1993). Gattung (1991) bemerkt, daß OPAC-Nutzer am Düsseldorfer Katalog nicht selbständig Boolesche Operatoren einsetzen, wengleich sie sich der impliziten Verknüpfung der Eingabefelder mit AND offenbar bewußt seien. Glöckner-Rist et al. (1989) berichten von ihren Untersuchungen an CD-ROM-Datenbanken, daß die Suchenden auf die Verwendung Boolescher Operatoren offenbar aus Unkenntnis über deren Anwendung verzichteten. Eine Ausnahme würde AND machen, das regelmäßig eingesetzt würde – allerdings in vielen Fällen falsch, d.h. AND wird anstelle von OR verwendet. Das-Gupta (1987) macht diese Beobachtung ebenfalls und auch aus den Protokollen eigener Untersuchungen (Moranz, 1998) an der Universität Osnabrück geht hervor, daß Nutzer AND einsetzen, um eine OR-Verknüpfung zu erreichen. Glöckner-Rist merkt hierzu an, daß die „Verwendung Bool’scher Operatoren [...] Kenntnisse über die Abbildung logischer Beziehungen voraus[setzt], über die Laien bzw. Endnutzer i.d.R. nicht verfügen“ (Glöckner-Rist, 1993b, :70), eine Einschätzung, die durch die Untersuchungen von Tversky und Kahneman (1982, :15) bestätigt wird. Ogden und Kaplan (1986) zeigen überdies, daß viele Nutzer auch an einer Datenbankschnittstelle, die eingeschränkt natürlichsprachliche Ausdrücke versteht, regelmäßig *und* sowohl für die Bildung der Schnittmenge als auch der Vereinigungsmenge verwenden (vgl. auch 6.2). Das in der Theorie mächtige Werkzeug erweist sich in der Praxis als mindestens für den Laien nicht handhabbar. Da aber der unbedarfte Endnutzer für viele Schnittstellen zur Zielgruppe gehört, ist das Vorhandensein Boolescher Operatoren kein Argument für eine besonders gute Benutzbarkeit einer Schnittstelle.

Ebenso wichtig wie die Eingabe ist aber auch die Verbesserung der Qualität der Rechercheergebnisse. Ein generelles Problem liegt in der mangelnden Robustheit der zugrundeliegenden Systeme, über die eine Oberfläche, wie übersichtlich sie auch sein mag, nicht hinweghelfen kann. Unsicherheiten in der Schreibweise, in vielen Fällen sogar bereits Schreibvarianten, lassen Suchanfragen regelmäßig scheitern. Als Beispiel möge die erst wenige Monate alte WWW-Seite des 2001-Versands dienen (Abbildung 10), bei der die Suche nach „Schostakovitsch“ scheitert (nur „Schostakowitsch“ hat

Zweitausendeins zeigt: So war der Wilde Westen. Wirklich.

Guten Tag!	Liste aller Interpret/inn/en mit: <u>A B C D E F G H I J K L M N O P Q R S T U V W X Y Z</u>
Bestellzettel.	Liste aller Titel mit: <u>A B C D E F G H I J K L M N O P Q R S T U V W X Y Z</u>
CDs.	Suche per Eingabe: <input type="checkbox"/> Nach Interpret/in <input type="text" value="Schostakovitsch"/> <input type="button" value="suchen"/>
Bücher.	Es wurde nichts gefunden. Kein Treffer
Merkheft.	
Läden.	
Zu uns.	

Abbildung 10: Online Angebot des 2001-Versand

Erfolg). Auch die Suche mit der gebräuchlichen (englischen) Namensform „Handel“ des in Deutschland geborenen Komponisten, der die zweite Hälfte seines Lebens in England lebte und wirkte, scheitert.

Gerade bei der Schreibung von Eigennamen sowie bei fremdsprachigen Begriffen müssen Schnittstellen robuster werden. Das gilt insbesondere vor dem Hintergrund eines länder- und sprachenübergreifenden Einsatzes, wie er oben im Zusammenhang mit den Bemühungen der EU-Kommission beschrieben wurde. Auch andere, von Menschen als trivial empfundene Aspekte der menschlichen Sprache (z.B. anaphorischer Bezug auf eine vorangegangene Eingabe, Synonymverwendung usw.), können von Datenbankschnittstellen heute in der Regel nicht behandelt werden (vgl. 4.2).

Die Qualität der Rechercheergebnisse umfaßt aber auch die allgemeinen Systemmeldungen, insbesondere die Fehlermeldungen. Die in Abbildung 11 gezeigte Fehlermeldung ist für normale Benutzer kaum nachvollziehbar. Auch der Autor selbst kann nur vermuten, daß es sich um die öffnende Klammer handelt, die für Suchanfragen auf dem Informationsserver der Europäischen Union⁹ nicht zugelassen ist.

Angesichts dieser Beispiele für Schnittstellen und unter Berücksichtigung des eingeschränkten Nutzens Booleschen Retrievals entsteht der Eindruck, daß jeder Person die einen Internet-Zugang hat, theoretisch der direkte Zugang zu zahllosen Datenbanken mit wertvollen Informationen aus den Bereichen Wissenschaft, Wirtschaft, Politik und Gesellschaft eröffnet wird. In der Praxis aber stehen diese Informationen wiederum nur dem Spezialisten klassischer Retrievaltechniken zur Verfügung, der nicht nur die notwendigen Kommandos beherrscht, sondern auch die Qualität der Ergebnisse sei-

⁹<http://eur-op.eu.int>



Abbildung 11: Fehlermeldung des EU-Informationsservers

ner Suchanfragen realistisch einzuschätzen vermag. Gerade dieser Punkt aber ist in vielen Fällen von enormer Bedeutung. Oft geht es ja nicht nur darum, über einige besonders gute Treffer zu einer Suchanfrage einen Einstieg in das thematische Feld zu finden, sondern auch die Gewissheit zu haben, eine relativ zum Datenbestand *vollständige* Antwort erhalten zu haben. Auch wenn der Benutzer noch in der Lage sein sollte, die Qualität der Treffer, die er bekommt, zu beurteilen (was alles andere als selbstverständlich ist), so ist er doch ohne Detailkenntnisse des Datenbestandes nicht in der Lage abzuschätzen, wie vollständig die Systemantwort ist (siehe auch Seite 39). Die Folge ist zum Teil eine trügerische Sicherheit der Benutzer (vgl. 2.2) oder aber ein (berechtigtes) Mißtrauen dem System gegenüber, das sich negativ in Studien zur Nutzerzufriedenheit niederschlägt.

Das aber kann weitreichende und auch bezifferbare Folgen haben. Kunisch (1999) stützt sich auf die Nielsen Norman Studie (siehe auch Seite 19) wenn er für die USA den Kostenvorteil einer Online-Stellenanzeige (verglichen mit einem Zeitungsinserat) auf 8.000 USD pro Bewerber beziffert. Dieser Kostenvorteil werde durch schlecht gestaltete und disfunktionale Online-Stellenanzeigen nicht nur verspielt, sondern es würden Bewerber darüber hinaus auch noch abgeschreckt, auf herkömmliche, gedruckte Stellenangebote desselben Unternehmens zu reagieren.

Shneiderman et.al. schließlich fassen die Bedeutung des einfachen Zugriffs auf online verfügbare Informationen noch weiter, wenn sie schreiben:

„In particular, the future of the Web as a universal tool may depend on interface developers' ability to reduce frustration and confusion for the masses of users while enabling them to reliably find what they need.“(Shneiderman et al., 1998, :98)

2.2 Bibliotheken als exemplarische Anwendung

Elektronisch verfügbare Bibliothekskataloge sind eigentlich nichts weiter als Datenbanken mit einer Rechercheschnittstelle für Endbenutzer. Es handelt sich um eine weit verbreitete Schnittstelle mit Zugang zu großen Mengen derzeit beinahe ausschließlich sprachlicher Daten, die für Forschung und Wirtschaft von großem Wert sind. Dabei spielen Aktualität und schneller Zugriff auf die Daten eine große Rolle. Die besondere Bedeutung der Bibliothek als Informationslieferant unterstreicht neben anderen auch die EU-Kommission. Bibliotheken würden sich für Infrastrukturinvestitionen besonders anbieten, um den online Zugriff auf Informationen für alle Bürger zu ermöglichen (Europäische Kommission, 1998, :11). Als Rechercheschnittstelle betrachtet, läßt sich am Beispiel der Bibliothekskataloge die in 2.1 beschriebene Problematik der Online-Informationszugänge in konzentrierter Form beobachten.

Die Entwicklung dieser Schnittstelle für die Online-Literatursuche beginnt in den 70er Jahren mit der ersten Generation von Schnittstellen für vermittelte Recherchen. Professionelle Rechercheure nahmen offline Suchaufträge entgegen, die sie aufgrund ihrer Fachkenntnisse inhaltlich interpretieren und mit Hilfe ihrer Datenbank- und Systemkenntnisse in Suchanfragen umsetzen konnten. Der Bibliotheksnutzer hatte für die persönlich durchgeführte Suche den sogenannten Zettelkatalog zur Verfügung, in dem jedes Buch auf einer Karteikarte verzeichnet war.

In den 80er Jahren folgte dann die sogenannte zweite Generation des Online Public Access Catalog (OPAC), der technisch zumeist über eine kommandozeilenorientierte VT100-Oberfläche erstmals dem Endnutzer die eigenständige Recherche im Katalog ermöglichte. Bis heute hat in diesem Bereich keine echte Weiterentwicklung stattgefunden. Immer noch stellt diese zweite Generation des OPAC den Standard in den meisten Bibliotheken dar – eine Situation, vor der bereits Hildreth gewarnt hatte: „The danger exists that these commercial suppliers of online catalog systems will become stuck on the plateau of second-generation developments.“(Hildreth, 1987, :649) Dies ist umso bedauerlicher, da der OPAC dieser Generation gravierende Schwächen aufweist.

Wenn im Folgenden nun von „OPAC“ die Rede sein wird, ist immer der OPAC der zweiten Generation gemeint. Dabei geht es um den OPAC in seiner Funktion als Endbenutzerschnittstelle für die Recherche, nicht um die zusätzliche und für den Endbenutzer unsichtbare Funktionalität, die dem Bibliothekar die Verwaltung der Ausleihe,

das Mahnwesen, Beschaffungen, Fernleihe usw. ermöglicht.¹⁰

Ein Benutzer des OPAC wird zumeist bereits durch eine wenig ergonomische Oberfläche abgeschreckt. In der Universitätsbibliothek Osnabrück, mit einem Bestand von ca. 1 Million Büchern eine mittelgroße Bibliothek in Deutschland, kann zur Recherche im Bestand seit 1993 auf einen OPAC zurückgegriffen werden, dessen primäre Benutzerschnittstelle in Abbildung 12 wiedergeben ist:



Abbildung 12: Osnabrücker OPAC mit VT100-Schnittstelle

Dieser OPAC stellt seit 1996 die einzige Möglichkeit dar, in Osnabrück auf aktuellen Daten zu recherchieren, da seitdem der Zettelkatalog nicht mehr gepflegt wird. Dabei steht der gezeigte OPAC stellvertretend für andere OPAC-Systeme. Tatsächlich ist es so, daß das in Osnabrück eingesetzte System im ganzen Norddeutschen Bibliotheksverbund, d.h. in den Bundesländern Schleswig-Holstein, Hamburg, Niedersachsen, Bremen, Sachsen-Anhalt, Mecklenburg-Vorpommern und Thüringen sowie in anderen Verbänden, insgesamt also in mehr als 100 wissenschaftlichen Bibliotheken Deutschlands eingesetzt wird. Zusätzlich ist dieses OPAC-System in zahlreichen Bibliotheken des europäischen Auslands installiert.

Aus verschiedenen Untersuchungen ist bekannt, daß Benutzer eines OPACs große Probleme haben, diesen adäquat zu bedienen (Schulz, 1994; Schmidt et al., 1997; Dreis, 1994; Glöckner-Rist et al., 1989; Glöckner-Rist, 1993b; Gattung, 1991; Hildreth, 1983; Peters, 1989; Hamborg et al., 1998). Die Schwierigkeiten beginnen bereits bei der verwendeten Oberfläche, die für Benutzer ohne Computervorerfahrung

¹⁰Daß auch hier oftmals schwerwiegende Mängel zu beklagen sind, belegt z.B. Müller et al. (1998) und Droste et al. (1998).

so unübersichtlich ist, daß eine erfolgreiche Suche im System in Frage gestellt wird. Aus dieser Einsicht heraus gibt es seit etwa drei Jahren verstärkte Anstrengungen, den OPAC der zweiten Generation mit einer verbesserten, zumeist graphischen Oberfläche auszustatten. Die Vorteile graphischer Oberflächen sind von Arbeitspsychologen bestätigt worden (Moranz, 1998; Hamborg et al., 1998; Rauterberg, 1989, 1992, 1995; Altmann, 1987) und liegen für Nutzer ohne Vorerfahrung vor allem in größerer Übersichtlichkeit und einfacheren Interaktionsmöglichkeiten mit Hilfe der Maus. Dabei ist in vielen Fällen zu erkennen, daß es sich um eine bloße Übertragung handelt, ohne konsequent die Möglichkeiten der neuen Technik zu nutzen.

Abbildung 13 läßt sich vielleicht am ehesten als den Versuch der Nachbildung einer Telnet-Verbindung mit den Mitteln von HTML bezeichnen. Der Vorteil für den Anwender fällt dabei im Vergleich zu dem auf Telnet basierenden OPAC sehr gering aus:

Übertragen mittles DUE1 Refresh Stop

UBM / Recherche

056: zwischen 2 Wörtern muss 1 Operator stehen UBHOPAC /WWWOP567

TITELANFANG ohne einl. Art. : I

oder STICHWÖRTER (max. 5) :

oder SCHLAGWÖRTER (max. 5) :

oder komb. STICH-/SCHLAGW. : data mining

AUTOR/HRSG. (Name, Vorname) :

AUTOR/HRSG. (Name, Vorname) :

KÖRPERSCHAFT (max. 5 Wörter) :

ORT: VERLAG: ISBN:

Eingrenzung der Suche

Jahr: von bis vor nach Band: Zschr.:

SYSTEMATISCHE NOTATION :

SIGNATUR (nicht kombinierbar):

Operatoren: + (UND); / (ODER); - (NICHT) Worttrunkierung: * (Anfang oder Ende)

EINGABE > <

P1---P2---P3---P4---?+P5--P6---P7---P8---x+P9--x+P10---x+P11---P12--

Menü Hilfe Eingabe Lösche Index Suche

Logics ep9750: LTG

Abbildung 13: Web-OPAC der Universitätsbibliothek München

Deutlich elaborierter ist die Oberfläche der WWW-Version des Osnabrücker OPACs (Abbildung 14). Sie ist mit einer echten Navigationsleiste ausgestattet und ermöglicht über per Maus bedienbare Felder die Auswahl der Operatoren sowie der zu durchsuchenden Felder. Verglichen mit der in Abbildung 12 gezeigten VT100-Schnittstelle bietet dieses System aber keine neuen Suchmöglichkeiten. Dies ist wiederum nicht spezifisch für das Osnabrücker System, sondern kennzeichnet den derzeitigen Stand der Technik: Weiterentwicklungen kommerzieller Systeme finden gegenüber der sogenannten zweiten Generation von OPACs nur auf der Oberfläche statt, nicht aber an der eigentlichen Systemfunktionalität.

Dabei ist es die Systemfunktionalität, die nach allem, was aus der Literatur bekannt

Ende Katalogmenü Suchmenü Mail Hilfe

LBS

Sachgebiete
Nutzerdaten

Suchmenü
Suchmaske mit erweiterten Suchmöglichkeiten

suchen Titelstichwörter Semiotistik

und Körperschaften Morgenländische Gesellsch

und Nummern (z.B. ISBN, ..) 02597934

und Materialart f

Abschicken

Erklärung der Suchmaske

Diese Suchmaske kann für komplexe Suchfragen verwendet werden. Klicken sie auf "Hilfe" für weitere Hilfen zum Suchen.
Mit den ersten 9 Eingabefeldern kann eine Suche zusammengestellt werden. Diese kann später eingeschränkt (und) oder erweitert (oder) werden, bzw. können Suchbegriffe ausgeschlossen werden (nicht).

- und die angegebenen Suchbegriffe müssen im Titel vorkommen
- oder eine der angegebenen Suchbegriffe muß im Titel vorkommen
- nicht die angegebenen Suchbegriffe dürfen nicht im Titel vorkommen

Mit den letzten Eingabefeld kann die Suchaktion auf ein Erscheinungsjahr oder eine Materialart eingeschränkt werden.
Pro Feld sollte nur ein Suchbegriff eingegeben werden.

Abbildung 14: Osnabrücker WWW-OPAC

ist, für die eigentlichen Schwierigkeiten der Benutzer verantwortlich ist. Die Bedienung des OPAC orientiert sich immer noch an Systemen, die bislang allein speziell ausgebildeten Rechercheuren vorbehalten waren:

„Die Suche nach Zeitschriftentiteln erfolgt durch die Kombination des Titels mit dem Suchbegriff zs. Die Suchstruktur lautet: s ti [für den Titelanfang] = [gesuchter Titel] & zs = oder: s zs = & ti = [gesuchter Titel], z.B. s ti = Bibliotheksdienst & zs = oder: s zs = & ti = Bibliotheksdienst. Sie erhalten dann die Titelanzeige der gesuchten Zeitschrift. Um eine Übersicht über den vorhandenen Bestand zu erhalten, geben Sie b0001 (für Bestand) ein. Haben Sie bei der Suche nur einen Treffer erhalten, so können Sie auch nur b eingeben.“ (Aus der Online-Hilfe des OPAC der Universitätsbibliothek München)

Aus eigenen Untersuchungen (Hamborg et al., 1998) und der Literatur (vgl. Dreis (1994) und Schulz (1994)) geht aber hervor, daß Bibliotheksbenutzer die Bedeutung der meisten angebotenen Suchmöglichkeiten nicht verstehen. Und auch dort, wo sich zusätzlich zur oben zitierten Eingabesyntax andere Eingabemodi etabliert haben, sieht es nicht besser aus. So kann ein Benutzer, statt beispielsweise *s ti = Bibliotheksdienst* einzugeben, mit der Funktionstaste F1 (Telnet Version) oder einem Mausklick auf das entsprechende Feld (WWW-Version) direkt in der Kategorie *Titelstichwörter* suchen. Der Osnabrücker OPAC stellt dafür sowohl in der Telnet-Version (Abbildung 12), als auch in der WWW-Version (Abbildung 14) dem Benutzer die Kategorien *Titelstichwörter*, *Personennamen*, *Körperschaften*, *Kongreßtitel*, *Serientitel*, *Systematik*, *Num-*

mern (ISBN etc.) und *Signatures* zur Verfügung.¹¹ Jedoch werden von diesen Kategorien lediglich die beiden ersten, *Titelstichwörter* und *Personennamen*, von den Benutzern verstanden und damit auch zur Suche genutzt. So berichtet Gattung (1991), daß 50 % der Anfragen am Düsseldorfer OPAC mit dem Suchaspekt *Titelstichwörter* durchgeführt werden, in beinahe 40 % der Fälle wird die Suchoption *Personennamen* verwendet. Weniger als 10 % der Nutzer führen eine kombinierte Suche durch, d.h. sie verknüpfen Personennamen und Titelstichwort miteinander. Etwa 1 % (!) der Anfragen werden über den Aspekt *Körperschaften* gestartet, der Suchaspekt *Systematik* wird überhaupt nicht verwendet. Diese Zahlen decken sich mit Beobachtungen, die selbst vor Ort in Osnabrück gemacht wurden (Moranz, 1998). Peters (1989) kommt schon Jahre früher zu ganz ähnlichen Ergebnissen.

Dabei kommt es in der Literatur zum OPAC zu der paradoxen Situation, daß einerseits z.B. Schmidt et al. (1997) schreiben, daß die OPAC-Systeme in erster Linie an den Anforderungen der Bibliothekare ausgerichtet seien: „So errechnen Erwerbungsmodule Medienpreise in unterschiedlichen Währungen, werden Vorakzession und Mahnschreiben automatisiert; die Benutzer dagegen müssen mit einem OPAC vorlieb nehmen, der wenig mit ihren eigenen Suchstrategien, ihrem Vorwissen und Verständnis von adäquater Dienstleistung zu tun hat.“ (Schmidt et al., 1997, :215) Andererseits bemerkt Pietzka (1995) (ohne damit in einem Widerspruch zu Schmidt et al. (1997) zu stehen):

„Die Arbeit mit OPACs erfordert vom Bibliothekar nicht nur die Beherrschung seiner eigenen Disziplin, vielmehr werden auch detaillierte Kenntnisse im Bereich der elektronischen Datenverarbeitung impliziert. Qualifikationen, die insbesondere ältere Bibliothekare nicht immer aufweisen können.“ (Pietzka, 1995, :378)

Wirklich einfach und intuitiv ist der OPAC offenbar von keiner Zielgruppe zu bedienen.

Dies könnte zumindest auf Seiten der Endnutzer darin begründet sein, daß nur auf den ersten Blick ein großer Bruch zwischen den Karteikarten des Zettelkataloges einerseits und der OPAC-Recherche andererseits besteht. Bei der Suche im Zettelkatalog geht es konstruktionsbedingt immer um die Suche nach Einzelexemplaren – je Karteikarte betrachtet der Benutzer unmittelbar einen Titel. Bei der Suche im OPAC dagegen tritt eine die Suche vermittelnde Schicht zwischen Benutzer und Titel. Indem der Benutzer eine Suchanfrage an das System stellt, überläßt er die Ausführung der Suche dem OPAC und nimmt später nur das fertige Suchergebnis in Empfang. Dabei suggeriert der OPAC bei vielen Benutzern eine trügerische Vollständigkeit des Ergebnisses.

Das diesem Trugschluß zugrundeliegende Problem scheint in der Übertragung der vom Zettelkatalog her bekannten Arbeits- und Denkweisen auf den Computer zu liegen. Konkret ist dies die naive Vorstellung, durch Überführen der Texte der Karteikarten

¹¹Einige Bibliotheken bieten darüber hinaus noch einen Suchaspekt *Schlagwort* an. Auf dessen besondere Problematik wird auf Seite 37 eingegangen.

in ein maschinenlesbares Format einen schneller und besser bedienbaren Zettelkatalog geschaffen zu haben. Bei dieser Umsetzung werden aber die zuvor dem Betrachter abverlangten kognitive Prozesse beim Suchen und Bewerten der Karteikarten ignoriert. Indem ein menschlicher Benutzer eine Karteikarte betrachtet, ist er in der Lage, mit Hilfe seiner gesamten sprachlichen Kompetenz und ggfs. seiner Erfahrung, den Inhalt der Karteikarte zu erfassen und für seine Zwecke zu bewerten. Demgegenüber arbeitet der OPAC zwar unvergleichlich viel schneller als der blätternde Mensch, jedoch geht er dabei natürlich vollständig mechanisch vor und ist insbesondere nicht in der Lage, Inferenzen zu ziehen, Assoziationen zu knüpfen oder ein Suchergebnis in Bezug zur eigenen Erfahrung zu setzen.

Die Konsequenz der stattdessen eingesetzten (derzeit primitiven) maschinellen Suchverfahren sind Systemschwächen, wie sie sich bspw. in mangelnder Robustheit gegenüber fehlerhaften Sucheingaben oder dem alleinigen Verlaß auf Trunkierung bei der Suche im Wortfeld äußern.

Das Ergebnis der Vorstellung der Benutzer, der OPAC sei ein automatisch funktionierender Zettelkatalog, ist eine, wie in Benutzerstudien (z.B. Dreis (1994)) nachgewiesen, abnehmende kritische Distanz zum gefundenen Material. Die Ergebnisse einer manuellen Suche im Zettelkatalog wurden durchweg subjektiv schlechter beurteilt als die Ergebnisse einer Suche im OPAC. Diese Ansicht, gepaart mit dem ebenfalls nachweislich fehlenden Wissen über bibliothekarische Erschließung sowie Funktionen eines OPAC, führen zu einem gefährlichen und angesichts der im Folgenden genannten objektiven Unzulänglichkeiten auch meist unangebrachten Vertrauen des Benutzers in die Ergebnisse seiner Recherche am OPAC.

Da sich die allermeisten Bibliotheksbenutzer nicht bewußt sind, welche Diskrepanz zwischen ihrer kognitiven Leistung bei der Betrachtung und Bewertung einer Karteikarte und der mechanischen Suche des OPAC besteht, sind sie auch nicht in der Lage, ein Bewußtsein für die notwendigen Techniken im Umgang mit Systemen wie dem OPAC zu entwickeln.

Die Folge ist, daß OPAC-Benutzer, wie im Anschluß stellvertretend am Beispiel des OPAC der Universitätsbibliothek Osnabrück dargestellt, sehr oft von den Ergebnissen ihrer Rechercheversuche enttäuscht sind.

Um mit Hilfe des OPAC Literatur zu suchen, stehen dem Nutzer die bereits erwähnten Aspekte *Titelstichwörter*, *Personennamen*, *Körperschaften*, *Kongreßtitel*, *Serientitel*, *Systematik*, *Nummern* (ISBN etc.) und *Signaturen* zur Verfügung, von denen in der Praxis aber nur die beiden ersten verwendet werden (siehe S. 29). Mit Hilfe der Optionen *Titelstichwörter* und *Personennamen* ist es einfach möglich, das Vorhandensein eines Werkes, dessen Titel und Verfasser sowie deren genaue Schreibweise bekannt sind, zu überprüfen. Schwieriger wird es, wenn die Schreibweise des Namens oder die genaue Titelformulierung nicht bekannt sind. Stichwortsuche findet im OPAC über reinen Strukturvergleich (pattern matching) statt, d.h. nur, wer mit dem verwendeten Stichwort die vom Autor des Buches gewählte Titelformulierung trifft, hat Erfolg: Die Suchanfrage *italienischer film* findet weder das Werk *Lexikon des italienischen Films* noch die *Geschichte italienischer Filme*, vom Buch *Das italienische*

Nachkriegskino einmal ganz zu schweigen. Mit diesen Mitteln eine Suche nach Literatur zu einem Thema, d.h. einer Reihe von Büchern, die sich einem bestimmten Thema widmen, durchzuführen, wird für die Benutzer frustrierend. Die Ergebnisse solcher Suchanfragen sind oft gekennzeichnet durch übergroße Treffermengen oder leere Treffermengen.

Übergroße Treffermengen kommen bei der Stichwortsuche in einem OPAC leicht vor, wenn bestimmte Suchbegriffe wie *Deutschland* (an einem deutschen Katalog), *Geschichte*, *Chemie* oder Kombinationen wie *Deutschland AND Geschichte* verwendet werden. Solche Eingaben führen zum Abbruch der Suche, denn das System kann nur maximal 1.000 gefundene Titel darstellen. Aber auch an ganz unerwarteter Stelle kann pattern matching zu übergroßen Ergebnissen führen: Sucht ein Benutzer mit dem Stichwort *Rom* nach Literatur zur italienischen Hauptstadt, so findet er in Osnabrück fast 500 Titel. Davon sind die allermeisten aber gar nicht relevant, sie werden nämlich nur deshalb angezeigt, weil ihnen eine *CD-Rom* beiliegt und dies in der Titelaufnahme vermerkt ist. Solche Eingaben, die weniger als tausend Treffer liefern, führen dann zu dem in Bibliotheken oft zu beobachtenden Effekt, daß Benutzer über einen sehr langen Zeitraum an einem Bildschirm verweilen und Listen von mehreren hundert Suchergebnissen durchblättern. Schulz schreibt hierzu:

„Benutzer haben grundsätzlich Schwierigkeiten, unhandliche Treffermengen auf eine Menge wirklich relevanter Treffer einzugrenzen. Zwar enden viele Recherchen mit 0 Treffern; die, die es nicht tun, finden aber mitunter so viel Material, daß die Recherche abgebrochen werden muß. 29 % der Recherchen nach Library of Congress Subject Headings produzierten mehr als 100 Treffer; die Benutzer brachen entweder ihre Suche ab oder sahen sich jeden Datensatz an – wobei sie das Terminal für andere Benutzer blockierten [...]“ (Schulz, 1994, :300)

Treffermengen sind also oft entweder nicht handhabbar groß oder werden erst gar nicht angezeigt.

Leere Treffermengen sind ebenfalls sehr häufig, und zwar aus drei Gründen: Zum einen, weil wie oben bereits erwähnt, die Suche im OPAC nur aus reinem pattern matching besteht. Wer das *Sprachwissenschaftliche Handbuch* sucht, wird weder das *Linguistische Wörterbuch* von Lewandowski noch das *Lexikon der Sprachwissenschaft* von Bußmann finden. Hinzu kommen Eigenheiten des Systems, die z.B. Benutzung von Sonderzeichen betreffen. *C^{*}-Algebren* sind im OPAC nicht auffindbar, weil der * im Titelstichwort vom OPAC als Wildcard identifiziert wird und somit nach *C* mit beliebigem rechten Rand gesucht wird. Um zu große Treffermengen zu verhindern, hat der OPAC aber eine Stopwortliste eingebaut, die neben häufig vorkommenden Wörtern wie Konjunktionen und Artikeln auch alle einzelnen Buchstaben des Alphabets umfaßt. Eine Suche nach *C^{*}-Algebren* wird mit der Begründung, *C* sei ein Stopwort umgehend abgelehnt. Selbst wenn der Suchende diese Meldung noch versteht, so hilft ihm das nicht viel, denn der Versuch mit *Algebra* oder *Algebren* zu suchen scheitert,

weil weit mehr als 1.000 Treffer erzielt werden. Weil *C* ein Stopwort ist, ist es am OPAC auch kaum möglich, Literatur zur Programmiersprache *C* zu finden. Zweitens ist es so, daß neben diesem durch den Autor eingeführten Vokabular sich kaum weiteres sprachliches Material, von einem Bibliothekar den Datensätzen angefügt, in einem OPAC findet:

„Sie [die Benutzer] erzielen 0 Treffer bei jeder zweiten bis dritten Recherche. Dies ist keineswegs erstaunlich, da Datensätze in OPACs inhaltlich äußerst dürftig deskribiert sind; es ist schon ein glücklicher Zufall, wenn ein Suchterminus mit einem Indexierungsterminus übereinstimmt.“(Schulz, 1994, :299)

Gattung entwirft ein ähnlich frustrierendes Bild des Suchenden:

„[Der Benutzer] verbringt ca. 15 Minuten vor dem Bildschirm. Innerhalb dieser Zeit startet er drei Anfragen, wobei eine mit '0 Treffer' beantwortet wird. Da er dem Null-Treffer-Ergebnis mißtraut, wiederholt er exakt diese Anfrage noch einmal und erhält – natürlich – wiederum '0 Treffer'.“(Gattung, 1991, :104)

Diese Beobachtungen, insbesondere die Wiederholung¹² der Eingabe in der Null-Treffer-Situation, bestätigen ein weiteres Mal die Untersuchungsergebnisse von Taylor (1984) und Peters (1989). Auch Borgman (1986) kommt zu ähnlichen Ergebnissen hinsichtlich der Null-Treffer-Situationen.

Und schließlich führen die mangelnde Robustheit und Flexibilität sowie generelle Probleme der Handhabbarkeit von OPAC-Systemen dazu, daß viele Sucheingaben scheitern. So muß der Benutzer, um am OPAC eine Zeitschrift zu suchen, zunächst eine Suche nach einem allgemeinen Merkmal wie *Titelstichwort* durchführen. Die Ergebnismenge dieser Anfrage kann er dann durch nachfolgende Eingabe von *UND* um ein weiteres Merkmal einschränken. An dieser Stelle erweitert sich die Liste der ihm angebotenen Kategorien u.a. um das Merkmal *Materialart*. Wählt er diese Kategorie aus, so kann er aus den in Abbildung 15 gezeigten Materialarten die Zeitschriften wählen. Dadurch wird in der Ergebnismenge der ersten Anfrage die Teilmenge der Zeitschriften isoliert.

Auffällig ist an dieser Stelle zunächst der seltsame Bezeichner für *Zeitschrift*. Dies wird erst verständlich, wenn man weiß, daß *Zeitschrift* im Niederländischen *Tijdschrift* heißt und daß die Herstellerfirma des beschriebenen OPAC-Systems aus den Niederlanden stammt. In Abbildung 14 ist übrigens zu sehen, daß sich diese Merkwürdigkeit auch in der überarbeiteten WWW-Version des Programmes wiederfindet – hier allerdings ohne die erklärende Beschreibung *Zeitschrift*. Dies ist nicht nur eine Kuriosität, sondern ein Beispiel für die vielfach kritisierte Unverständlichkeit der in OPAC-Systemen verwendeten Bezeichnungen (vgl. z.B. Janosky et al. (1986)). Viele

¹²Obermeier (1999) berichtet, daß manche Benutzer bis zu sieben Mal genau dieselbe Suchanfrage wiederholen.

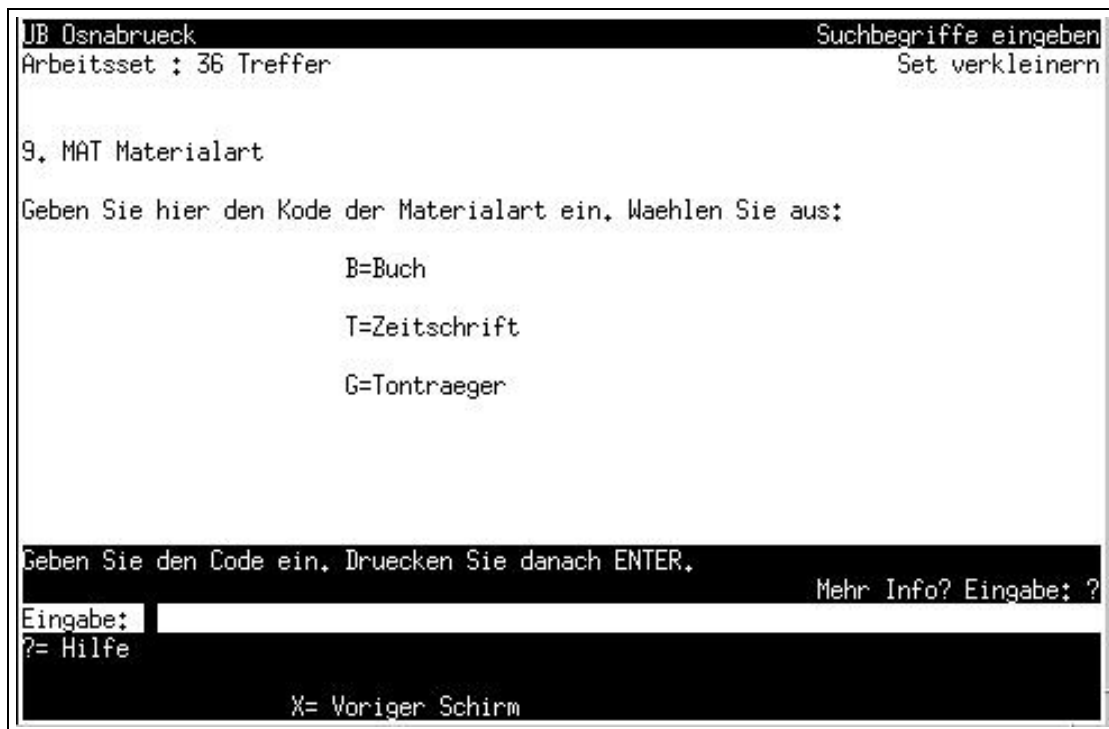


Abbildung 15: Schwer nachvollziehbare Bezeichnungen

Benutzer kommen allerdings nicht einmal zu dem Punkt, an dem sie einfach wissen müssen, daß *t* für *Zeitschrift* steht. Nach Gattung (1991) sind nicht einmal 10 % der Benutzer in der Lage, eine verknüpfte Suche durchzuführen. Der Anteil der Nutzer, die Suchmengen nachträglich einschränken können, dürfte nur unwesentlich höher liegen. Der langwierige Weg und die unintuitive Bezeichnung *Materialart* machen nach eigenen Untersuchungen (vgl. Moranz et al. (1999)) die Suche nach Zeitschriften am OPAC für viele Benutzer unmöglich. Schulz (1994) bestätigt dies, wenn sie schreibt, daß Endnutzer häufig schon dann die Orientierung verlieren würden, wenn sich der Dialog auch nur einen Schritt von der unmittelbaren Abfolge von Eingabe und Ergebnispräsentation entfernt. Die Rückmeldungen des Systems sind für den Benutzer zudem oftmals vollkommen unverständlich (siehe Abbildung 16).

Auch der Suchaspekt *Personennamen* ist nicht ohne Schwierigkeiten. Um am OPAC Erfolg zu haben, muß der Suchende sich an die bibliothekarischen Ansetzungsregeln halten und die Eingabe in der Form Nachname, Vorname vornehmen. Dies ist laut Schmidt et al. (1997) nicht von allen Benutzern zu erwarten und bereits Dickson (1984) berichtet, daß nach ihren Untersuchungen 18 % der Autorensuchen *nur* deshalb scheitern, weil Vorname und Nachname nicht in der invertierten Form eingegeben wurden. Borgman (1986) berichtet immerhin von fast 13 % in diesem Zusammenhang. In einer Untersuchung von Taylor (1984) scheiterten 16,7 % der Suchanfragen nach Autorennamen, weil Vor- und Nachname in natürlicher und nicht in invertierter Form eingegeben wurden. Dieser Punkt, auf den seit über 15 Jahren in Untersuchungen zu



Abbildung 16: Unverständliche Fehlermeldung

Bibliothekskatalogen hingewiesen wird, unterstreicht sehr deutlich die auf Seite 24 zitierte Aussage von Hildreth, die Hersteller von OPAC-Systemen würden die zweite Generation dieser Systeme nicht weiterentwickeln. Daß es sich dabei nicht um ein technisch unlösbares Problem handelt, wird in 3.1 gezeigt.

Gattung schreibt, daß Nutzer des OPAC bei allen Suchaspekten, nicht nur bei Personennamen, den angebotenen Index nicht nutzen. „Eher versucht [der Nutzer] immer wieder einen Namen in verschiedenen Schreibweisen einzugeben, anstatt sich die korrekte Form aus dem Index herauszuselektieren“ (Gattung, 1991, :106).

Und auch, wenn man sich über diese Konvention im Klaren ist, ist es nicht immer trivial, den Nachnamen eindeutig zu identifizieren. Während *Hoffmann von Fallersleben* als *Fallersleben*, *Hoffmann von* zu suchen ist, muß *Wernher von Braun*, weil er als Amerikaner betrachtet wird, als *VonBraun*, *Wernher* dem System genannt werden.

Auch in einer WWW-Version des OPAC muß sich die Bedienbarkeit nicht notwendig verbessern, wie Abbildung 13 beweist. Hier ist es einfach die Tatsache, daß zwei Suchbegriffe in ein Suchfeld geschrieben wurden, die die Anfrage scheitern läßt. Die Systemmeldung „056: Zwischen 2 Wörtern muß 1 Operator stehen“ trägt nicht zur Verständlichkeit bei, weil einer der wenigen Vorteile zumindest dieser WWW-Oberfläche ja gerade darin besteht, daß die Operatoren implizit die Eingabefelder verknüpfen.

Die Kontinuität der Systemfunktionalität beim Übergang vom Telnet-OPAC zum WWW-OPAC wird auch bei der Systemreaktion auf erfolglose Sucheingaben deutlich: Immer noch wird in dieser Situation der alphabetische Index der vorhandenen

Suchbegriffe eingeblendet. Hierzu bemerkt Schulz (1994):

„Ist der Suchterminus nicht im Index enthalten, ist das automatische Aufblättern des entsprechenden alphabetischen Abschnitts unverständlich und geradezu zynisch.“(Schulz, 1994, :299)

Neben diesen Unzulänglichkeiten des Systems sind es aber hauptsächlich fehlerhafte Sucheingaben (Tippfehler, orthographische Fehler etc.), die in vielen Fällen eine Suche scheitern lassen. Dieses Phänomen ist aus zahlreichen Benutzeruntersuchungen bekannt (vgl. z.B. Taylor (1984); Peters (1989)): so geht Schulz (1994) davon aus, daß jede zehnte Suchanfrage allein aufgrund von durch den Benutzer nicht erkannten(!) Eingabefehlern zu einer leeren Ergebnismenge führt. Borgman (1986) berichtet, daß 15 % der Titelsuchen aufgrund von Fehlschreibungen scheitern. Dickson (1984) spricht von ebenfalls knapp 15 % in diesem Zusammenhang.

Als probates Mittel für die Suche in Bibliothekskatalogen wird immer wieder die Verwendung von Wildcards (*Trunkierung*) angeführt. Trunkierung, auch Maskierung genannt, bezeichnet das Abschneiden eines Wortes und die Verwendung eines Platzhalters oder Jokerzeichens, um beliebige Wortendungen zu treffen. Streicht man in *Autradio* das *radio* und ersetzt es durch einen Platzhalter, z.B. das Symbol \$, dann würde nach dem Prinzip der Trunkierung jedes Wort gefunden, das mit *Auto* beginnt, neben *Autoradio* also auch *Autodach*, *Autotiür*, *Autoscheibe* usw. Dieser abstrakte Umgang mit Suchtermini kann von den Benutzern aber nicht erwartet werden, denn Benutzeruntersuchungen (vgl. u.a. Borgman (1986); Glöckner-Rist et al. (1989); Glöckner-Rist (1993b); Obermeier (1999)) haben gezeigt:

„Der Zweck von Trunkierungsmöglichkeiten wird offensichtlich nicht verstanden, die Funktion folglich nicht genutzt.“
(Schulz, 1994, :300)

Gattung beschreibt den durchschnittlichen OPAC-Nutzer, den sie aus Log-Dateien des Düsseldorfer OPACs ermittelt hat so:

„Als Titel- oder Schlagwort gibt der Benutzer alles ein, was ihm einfällt. [...] Er benutzt kaum mehr als zwei Suchbegriffe, er trunkiert nicht und er setzt auch nicht bewußt Boolesche Operatoren ein.“(Gattung, 1991, :105)

Außerdem führt Trunkierung in der Stichwortsuche noch eher zu den erwähnten übergroßen Treffermengen: Angenommen, ein Benutzer sucht nach Literatur zum Thema *genetische Manipulation* und verwendet in seiner Suchanfrage Jokerzeichen: *genetisch\$ UND Manipulation\$*. Dann trifft eine simple Endtrunkierung wie in *Manipulation\$* bereits sehr viele in diesem Zusammenhang gar nicht erwünschte Wörter:

Manipulation \$

Manipulation skunststücke

Manipulation sbefürchtung

Manipulation staktiken

Ein Adjektiv wie *manipulatorisch*, das zu finden sicherlich wünschenswert wäre, angesichts möglicher Titel wie *Kritik einer manipulatorischen Genetik* wird durch Manipulation\$ aber nicht gefunden. Das Jokerzeichen müßte also noch weiter am Wortanfang stehen (manipulat\$), wodurch die Suchergebnisse aber auch noch stärker streuen. Und auch dies ist noch nicht genug. Das Deutsche ist gekennzeichnet durch seine Produktivität im Bereich der Komposita, wie z.B. *Genmanipulation*. Um dieses Kompositum zu finden, müßte das Jokerzeichen auch vorn stehen (\$manipulat\$), was auch *Spendenmanipulationskritik* träfe. Endgültig ad absurdum getrieben wird diese Strategie angesichts des unregelmäßigen Plurals im Deutschen. Ein Suchstring wie \$haus\$ trifft den Plural *Häuser* eben nicht. Die Idee, mit \$h\$us\$ zu suchen, verbietet sich angesichts der zu erwartenden Treffermenge natürlich von selbst.

Suchergebnisse werden durch Trunkierung also sehr viel unpräziser, weil die spezifischen Regeln der Variation des sprachlichen Materials nicht berücksichtigt werden und stattdessen *beliebige* Wortfortsetzungen getroffen werden. Für das Beispiel von Seite 35 bedeutet das, daß neben den erwünschten und genannten Treffern *Autodach*, *Autotür* auch die in diesem Zusammenhang sicherlich unerwünschten Wörter *Autodidakt*, *Automat*, *Autopsie*, *Autor* und viele andere gefunden werden.

Zudem stellt Trunkierung für semantische Phänomene (bspw. Synonyme und Ambiguitäten) ebenfalls keine Lösung dar: *Geige* kann nicht durch Trunkierung auf *Violine* abgebildet werden und *Morphologie* wird mit oder ohne Trunkierung in einem OPAC quer durch die Disziplinen *Geologie*, *Medizin*, *Biologie* und *Linguistik* treffen. *Tierwelt* ist ebensowenig auf *Fauna* abbildbar wie *Meere* auf *Ozeane*.

Doch wie steht es nun um die eingangs bereits erwähnten Schlagwörter? Schlagwörter sollen das Auffinden von Büchern erleichtern, die aufgrund ihrer Titelformulierung sonst schwer auffindbar wären. So bekommt ein Buch mit dem Titel „Italiens schöne Hauptstadt“ das Schlagwort *Rom*, damit die Chancen, es zu finden, steigen. Ein anderes Beispiel ist die Vergabe eines Schlagwortes wie *Kernenergie* für das Buch „Heller als tausend Sonnen“.

Mit Schlagwörtern erscheint also möglich, was am Beispiel der Stichwortsuche noch als praktisch unmöglich geschildert wurde: Die Suche nach Büchern, die die italienische Hauptstadt *Rom* zum Thema haben. Dabei würde es keine Überschneidungen mit *CD-Rom* geben, weil ja nicht auf dem Titelmateral, sondern nur auf den Schlagwörtern gesucht wird.

Dennoch werden Schlagwörter von den Bibliotheksnutzern beinahe überhaupt nicht genutzt. Das liegt zum einen daran, daß der vorhandene Buchbestand in Osnabrück nicht durchgängig mit Schlagwörtern versehen ist. Schlagwörter werden intellektuell durch ausgebildete Fachkräfte einem Datensatz angefügt (den intellektuellen und personellen Aufwand beschreibt z.B. Maaßen (1981)). Die hohen Kosten dieser personalintensiven Erschließung haben dazu geführt, daß nur etwa 50 % der Bestände der Universitätsbibliothek Osnabrück mit Schlagwörtern versehen sind. In dieser Hinsicht unterscheidet sich die Universitätsbibliothek Osnabrück nicht von den meisten anderen deutschen Bibliotheken. Aus diesem Grunde wird im Osnabrücker OPAC auch keine Suchoption *Schlagwort* angeboten – mit ihrer Hilfe würden Benutzer von vornherein nur auf einem Bruchteil des eigentlichen Bestandes suchen.

Angesichts der Kosten der Schlagwortvergabe erscheint es illusorisch zu hoffen, daß durch aufwendige Investitionen der Bestand der Bibliotheken flächendeckend verschlagwortet werden wird. Und auch durch verstärkte Anstrengungen bei der Katalogisierung neuer Bücher hätte man immer noch keine Perspektive für die bereits vorhandenen Titel.

Aber auch ein vollständig mit Schlagwörtern versehener Bestand wäre entgegen der landläufigen Annahme kaum besser für die Literaturrecherche von Endbenutzern geeignet: Aufgrund der enormen Komplexität der deutschen Ansetzungsregeln für Schlagwörter¹³ erscheint eine Benutzung der Schlagwörter durch einen bibliothekarischen Laien kaum möglich. Nur eine regelkonforme Anwendung garantiert aber den Erfolg: Die Suche mit dem Schlagwort *Hund* trifft auch in einem vollständig verschlagworteten Bestand keine inhaltlich relevanten Einträge, denn die Regel vom „engen Schlagwort“ sorgt dafür, daß die entsprechenden Datensätze mit dem Namen der

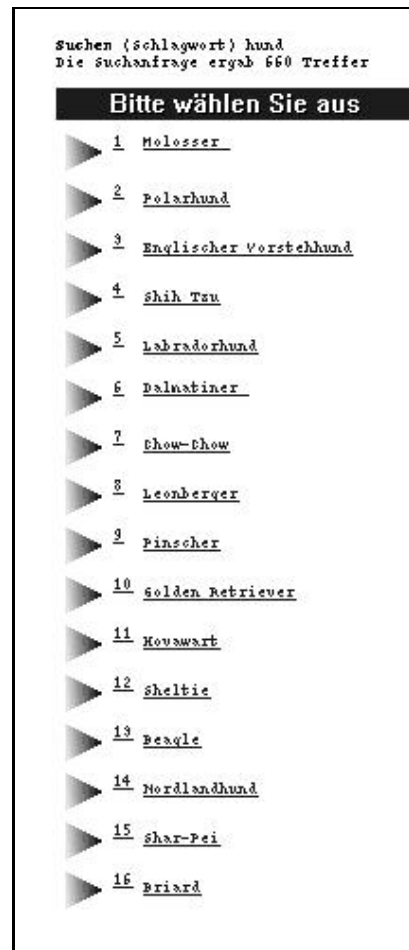


Abbildung 17:
Schlagwort *Hund*

¹³RSWK – Regeln für den Schlagwortkatalog.

Hunderasse versehen werden. Abbildung 17 zeigt das Ergebnis einer Suche mit dem Schlagwort *Hund* im Gesamtkatalog des Norddeutschen Bibliotheksverbundes (GBV). Da aber auch auf Schlagwortmaterial nur mit Strukturvergleich d.h. pattern matching gesucht wird, bleibt auch in einem zu hundert Prozent verschlagworteten Bestand die Suche nach Literatur zu *Hund* ohne Erfolg. Erst eine Suche auf den Schlagwortketten würde zum Erfolg führen (z.B. **Hund ; Magen ; Dilatation ; Rezidivprophylaxe*). Jedoch wird nach Schmidt et al. (1997) von Bibliotheksbenutzern der Sinn der Schlagwortketten (von denen es zu einem Schlagwort Hunderte geben kann) nicht verstanden. Folglich kann, wie auch schon in anderen Fällen gezeigt, eine angemessene Verwendung durch die Suchenden nicht erwartet werden (siehe auch Obermeier (1999)).

Problematisch sind die Schlagwortketten auch deshalb, weil die Relation zwischen den einzelnen Bestandteilen der Kette unspezifiziert ist. So kann eine Schlagwortkette *Goethe; Schiller; Weimar* bedeuten, daß das so bezeichnete Werk sich mit *Goethe* und *Schiller* und ihrem gemeinsamen Wirken in *Weimar* beschäftigt. Es kann aber auch der Einfluß *Goethes* auf *Schiller* in *Weimar* oder die Darstellung *Weimars* in den Werken *Goethes* und *Schillers* gemeint sein. Um die Relationen zwischen den Bestandteilen von Schlagwortketten präzise fassen zu können, wurde schon Anfang der siebziger Jahre mit PRECIS ein Verfahren entwickelt, das halbautomatisch qualitativ hochwertige Einträge für Schlagwortkataloge erstellt (Detemple, 1982; Austin, 1984). PRECIS greift auf Ideen der (Computer-)Linguistik zurück, indem es (wenn auch nur manuell) syntaktische und semantische Analyse, Dekomposition und Abbildung auf thematische Rollen vornimmt. Dazu wird in einem mehrstufigen Prozeß das Thema des zu indexierenden Werkes zunächst in Form einer komplexen Nominalphrase gefaßt ("Zerstörung von Holz durch Insekten in Norwegen") und dann oberflächlich syntaktisch und semantisch analysiert. Die Bezeichnung "oberflächlich" ist dabei nur aus heutiger Sicht gerechtfertigt - bedenkt man, daß PRECIS schon zu Beginn der siebziger Jahre entwickelt wurde und für einen Bereich gedacht ist, in dem noch Anfang der achtziger Jahre ein wesentlicher Vorbehalt gegenüber PRECIS dessen starke EDV-Abhängigkeit war (Maaßen (1983, :299)), muß die Beurteilung der Analyse relativiert werden. Eine wie oben genannte Nominalphrase als Beschreibung des Themas eines Werkes wird nun manuell in mehrere funktionelle Bestandteile wie "Ordinator", "Qualifikator" und "Modifikator" zerlegt und die Rollen der einzelnen Einträge in den Beziehungen untereinander durch verschiedene Operatoren ("Objekt einer transitiven Handlung") klargestellt. Auf der Grundlage dieser (manuellen) Analyse setzt dann die automatische Produktion der Schlagwortkatalogeinträge ein. Dieses Verfahren wurde mit Erfolg über mehrere Jahre im Alltagsbetrieb der British Library eingesetzt und auch in Deutschland zumindest erprobt. Daß PRECIS sich nicht als Indexierungsstandard durchgesetzt hat, ist aus heutiger Sicht überaus bedauerlich und vermutlich nicht (nur) auf systeminhärente Schwächen zurückzuführen (vgl. Maaßen (1983); Junginger (1984)). Mit Hilfe detaillierter und theoretisch fundierter Indexeinträge wie sie PRECIS zur Verfügung stellt, wären aktuelle Retrievalsysteme in einer sehr viel besseren Ausgangssituation.

Eng verwandt ist ein anderes Problem, das den Erfolg von Anfragen am OPAC in Frage

stellt: In deutschen Bibliotheken ist die Verwendung deutscher Begriffe vorgeschrieben. Wissenschaftliche Fachtermini, die gerade in den Naturwissenschaften und der Mathematik oftmals aus dem Englischen stammen (häufig ohne geläufiges deutsches Äquivalent), werden dazu durch einen passenden oder einen dem Bibliothekar passend erscheinenden deutschen Ausdruck ersetzt. Der Zwang zur Eindeutschung führt dazu, daß, wenn überhaupt ein Schlagwort vergeben wird, es in diesen Fällen eine den Fachwissenschaftlern selbst völlig unbekannte Übersetzung darstellt. *String theory* in der theoretischen Physik oder *loop spaces* in der Mathematik würden durch *Saitenbeschreibungsmodell* bzw. *iterierte Schleifenräume* ersetzt. D.h. selbst wenn Schlagwörter vorhanden sind, müssen sie also noch lange nicht dem Suchvorgang förderlich sein.

Eine weitere Möglichkeit zu suchen besteht in der klassifikatorischen Erschließung der Titel. Bei der Katalogisierung wird einem Titel aufgrund seines Inhalts vom zuständigen Fachreferenten eine Notation des Klassifikationssystems zugeordnet. So bekommt z.B. in Osnabrück ein Buch, das sich mit dem Thema *Computerlinguistik* beschäftigt, die Notation BFL der Gesamthochschulbibliotheks-Systematik (GHBS). Im Gegensatz z.B. zu den USA verwendet in Deutschland beinahe jede wissenschaftliche Bibliothek ihr eigenes Klassifikationssystem. Der Name „Gesamthochschulbibliotheks-Systematik (GHBS)“ täuscht darüber hinweg, daß beinahe alle Bibliotheken, die die GHBS einsetzen, dieses Klassifikationssystem nach ihren eigenen Vorstellungen zu einer individuellen Lokalversion ausgebaut haben. Daß dieses Klassifikationssystem von den Suchenden nicht genutzt wird, hat verschiedene Gründe. Zum einen liegt es in Osnabrück daran, daß die Erschließungsmerkmale (Notationen), unter denen die Bücher nach inhaltlichen Kriterien erfaßt sind, nur über eine gedruckte Ausgabe der Systematik ermittelt werden können (auch hier unterscheidet sich die Universitätsbibliothek Osnabrück nicht von den meisten anderen Bibliotheken). Aber auch unabhängig von der Tatsache, ob die Systematik nun maschinenlesbar oder nur auf Papier dem Benutzer zugänglich ist, stellt Gattung (1991) fest, daß Recherchen mit Hilfe klassifikatorischer Merkmale von den Suchenden vollständig gemieden würden.

Vor diesem Hintergrund bleiben als Mittel zur Literaturrecherche im OPAC nur die Stichwort- und die Autorensuche. Wie oben ausgeführt ist dies insbesondere für die thematische Recherche, d.h. für die Suche nach Literatur zu einem bestimmten Thema von Bedeutung. Gelingt es dem Suchenden, durch geschickte Formulierung seiner Anfrage sowohl leere als auch übergroße Ergebnismengen zu vermeiden, gerät er in die eigentlich bedrohlichste Lage: Wie möchte er sichergehen, daß die gefundenen Titel wirklich alle für seine Themenstellung relevanten Titel im Bestand sind?

Ein Beispiel: Sucht ein Benutzer im OPAC der Bibliothek Osnabrück nach Literatur zum Thema *Sport im Alter*, so ist er gezwungen, einen Ausdruck mit Hilfe Boolescher Operatoren zu konstruieren, etwa wie folgt: `sport UND alter`. Das Ergebnis dieser Anfrage sind fünf Bücher. Benutzer, die erwartungsgemäß dem System vertrauen, ärgern sich, daß ihre Interessen nicht in der Beschaffungspolitik der Bibliothek berücksichtigt werden. Benutzer, die (zu Recht) an der Zuverlässigkeit des Systems zweifeln,

beginnen nun, Ausweichstrategien zu entwickeln, um die Bücher zu finden, die der OPAC ihnen vorenthält. Eine Suche im semantischen Umfeld mit *Senioren-sport* oder *Alterssport* fördert schon mehr zu Tage, aber auch mit den geschicktesten Strategien werden Benutzer nie sicher sein können, alle für ihre Suchanfrage inhaltlich relevanten Werke der Bibliothek mit Hilfe von Titelstichwörtern gefunden zu haben. Schulungen und Hilfetexte, die in diesem Zusammenhang immer wieder (gerade von bibliothekarischer Seite) gefordert werden, sind keine Lösung des Problems. Schulz bringt es wiederum treffend auf den Punkt, wenn sie schreibt:

„ ‚Gutgemeinte‘ schriftliche Hilfen nützen dem guten Gewissen der Bibliothekare, nicht aber den OPAC-Nutzern. [...] Außerdem nehmen OPAC-Nutzer nur zu geringem Anteil an Schulungen teil, bzw. vergessen das Gelernte schnell wieder.“ (Schulz, 1994, :300)

Diese Zusammenfassung deckt sich mit den Aussagen von Wright (1983) und Rauterberg (1992). Schmidt et.al. schreiben, daß oft „Probanden sich aufmerksam lange Hilfetexte durchlasen, aber nur selten half ihnen die Information weiter“ (Schmidt et al., 1997, :222). Janosky bemerkt dagegen:

„[The users] do not even read the entire contents of one section of the help. They appear to make judgements about how much information is needed and stop reading as soon as they think they know enough.“ (Janosky et al., 1986, :587)

Aber auch wenn Hilfetexte zukünftig ansprechender und verständlicher formuliert (vgl. McDonald und Searing (1983)) und tatsächlich von den Nutzern gelesen werden, so wäre dies keine echte Lösung. Vielmehr sollte ein Nutzer von einer modernen Schnittstelle erwarten dürfen, daß diese zu bedienen nicht mehr als Aufmerksamkeit, normales Alltagswissen und etwas guten Willen erfordert.

2.3 Fazit

Bei der Betrachtung von Online-Schnittstellen in 2.1 und speziell von Bibliothekskatalogen in 2.2 war der erhebliche intellektuelle Aufwand für die Konstruktion geeigneter Anfragen aufgefallen. In 2.1 betraf dies vor allem die großen Unterschiede zwischen den verschiedenen Oberflächen: Der Nutzer von Online-Schnittstellen muß sich immer wieder auf andere, ihm neue Arten der Bedienung einstellen – auch auf verschiedenen Seiten einer Organisation (z.B. der Europäischen Union). Die Benutzung von Online-Schnittstellen sollte aber einheitlich sein, ohne daß dazu unbedingt ein Standard vonnöten wäre. Intelligente Schnittstellen, die es dem Nutzer ermöglichen, sein Informationsbedürfnis beispielsweise natürlichsprachlich mitzuteilen, könnten aus Sicht des Benutzers eine gewisse Vereinheitlichung bringen, indem sie zu einer Entlastung des Suchenden führen, ohne auf Systemseite eine Standardisierung zu erzwingen. Die Vereinheitlichung der Systembedienung würde dabei über die gleichartigen Nutzerkenntnisse der natürlichen Sprache und ihres Gebrauchs erreicht. In 2.2 zeigte sich für den deutlich kleineren Markt der Bibliotheksrecherchesysteme aufgrund historischer Entwicklungen („second generation OPAC“, s. S. 24) und Marktdominanz einzelner Unternehmen (s. S. 25) eine geringere Variationsbreite der Oberflächen, aber auch eine nicht mehr zeitgemäße Gestaltung dieser Schnittstellen (s. z.B. S. 27) zusammen mit einer lediglich kosmetischen Überarbeitung der Oberflächen (s. S. 28).

Weiterhin wurde die ergonomische Gestaltung der Oberfläche bemängelt, die einer einfachen Benutzung der Schnittstelle oft im Wege steht (vergleiche z.B. die Oberfläche der durch animierte Werbebanner und zahlreiche Optionen überladenen Suchmaschine Altavista). Eine ideale Online-Schnittstelle sollte so intuitiv, übersichtlich und verständlich gestaltet sein, daß das dahinterliegende System für Anwender ohne spezielles Training zu benutzen ist. Zur einfachen Bedienbarkeit gehört aber nicht nur eine übersichtliche Gestaltung, sondern auch die Robustheit des Systems gegenüber Fehlern oder Ambiguitäten in der Eingabe. Fehler können in der Schreibweise von Namen sowie generell in der Orthographie begründet sein, aber auch versehentlich entstehen (Tippfehler).

Natürlich lassen sich nicht alle Arten der unbeabsichtigten Fehlbedienung erkennen und behandeln, denn sogar zwischen kompetenten menschlichen Sprechern kommt es gelegentlich zum unbeabsichtigten Zusammenbruch der Kommunikation. Jedoch gibt es an Schnittstellen zahlreiche Eingabefehler, die sich leicht erkennen lassen, und auf die in ganz unterschiedlicher Art und Weise – von der stillschweigenden Korrektur bis hin zur expliziten Rückfrage – konstruktiv reagiert werden kann (s. S. 35). Dies setzt aber natürlich stärkere Fähigkeiten zur Eingabeanalyse voraus, als sie bislang vorhanden sind oder eingesetzt werden.

Daneben wurde die vielfach zu bemerkende Notwendigkeit kritisiert, über die Struktur des Datenbestandes informiert zu sein, um erfolgreich Anfragen gestalten zu können (s. Abbildung 1 auf S. 14). Diese und andere starke Voraussetzungen über die Fachkenntnisse des Nutzers z.B. bei der Wortwahl auf der Oberfläche (s. Abbildung 15 auf S. 33) müssen bei einfach zu bedienenden Schnittstellen vermieden werden. Bei der

Vermeidung von Voraussetzungen innerhalb der Oberfläche selbst kann die stärkere Nutzung graphischer Mittel zur Repräsentation zumindest teilweise helfen. Die derzeit häufig bestehende Notwendigkeit, Wissen über die Struktur des Datenbestandes zu haben, kann dagegen nur durch systeminhärente Intelligenz behoben werden, die die Interessen und Ziele des Benutzers erkennt und selbständig auf die Besonderheiten des Datenbestandes abbilden kann.

Von den bekannten Ansätzen, den Umgang mit Online-Schnittstellen einfacher zu gestalten, wurde die Vorgabe von Schemata, in denen Nutzer sich selbst oder ihr Anliegen einordnen müssen (s. Abbildung 3 auf S. 16) kritisiert. Dieses Vorgehen wurde als unzureichend bezeichnet, weil es in der Regel nicht flexibel und dynamisch genug ist und immer in der Gefahr steht, entweder zu grob und damit unbrauchbar oder zu fein und damit unhandlich zu sein. Eine andere Strategie zur Vereinfachung des Zugriffs ist die Antizipation der Benutzerbedürfnisse in der Struktur der Eingabefelder (s. Abbildung 2 auf S. 15 und Abbildung 7 auf S. 18). Dies ist von Vorteil, weil das System dem Benutzer bei der Konstruktion der Eingabe entgegenkommt. Der Nachteil liegt natürlich darin, daß dieses Entgegenkommen bereits bei der Erstellung der Schnittstelle fest in deren Struktur verankert wird. Entscheidend ist also, daß bei diesem Prozeß die Bedürfnisse des Benutzers richtig erkannt werden, daß die erkannten und in der Struktur der Oberfläche verankerten Bedürfnisse dem Benutzer auch vermittelt werden und daß die Bedürfnisse des Benutzers nicht variieren. Demgegenüber bietet die Möglichkeit, Boolesche Operatoren und reguläre Ausdrücke zu verwenden, dem Benutzer sehr viel größere Freiheit. Kritisiert wurde dies jedoch aufgrund der Tatsache, daß Boolesche Operatoren und reguläre Ausdrücke einerseits für die Suche auf *sprachlichen* Daten ungeeignet erscheinen und andererseits von den Suchenden offenbar nicht beherrscht werden (s. S. 21 sowie S. 35). Wie schon für die Vorgabe von Schemata und die Nutzerantizipation in der Oberflächengestaltung gilt also auch hier, daß die Interaktionsformen, derer sich Online-Schnittstellen bedienen, für den Benutzer einfacher und zugleich mit mehr Freiheitsgraden versehen sein müssen. Das heißt aber auch, wie schon mehrfach festgestellt, daß die Schnittstelle intelligenter werden muß: indem sie z.B. den Suchenden bei der Umsetzung des zu lösenden Problems in die Formulierung eines Informationsbedürfnisses unterstützt. Dabei sollte sich die Schnittstelle dynamisch den sich ändernden Bedürfnissen des Nutzers anpassen, eigenständig die für das aktuelle Informationsbedürfnis adäquate Informationsquelle auswählen und unter Berücksichtigung ihrer spezifischen Eigenschaften befragen.

Schließlich wurden Systemausgaben, insbesondere die Fehlermeldungen (s. Abbildung 11 auf S. 23 und Abbildung 16 auf S. 34), aber auch die Vertrauenswürdigkeit der Rechercheergebnisse an sich als aus Sicht des Suchenden ungenügend kritisiert. So ist es derzeit für einen Nutzer unmöglich, ohne detailliertes Wissen über den Datenbestand und das verwendete System abzuschätzen, von welcher Qualität seine Suchergebnisse sind und wieviele vergleichbare Treffer es eventuell noch im Datenbestand geben mag (vgl. S. 39). Auch dieser Punkt wird nur durch eine bessere Auswertung der Daten seitens der Rechercheprogramme möglich, d.h. es ist wiederum das System, das den Suchenden entlastet und ihm (mit mehr Aufwand erzeugte) hochwertigere

Suchergebnisse liefern soll. Ein anderer Punkt ist die Einbettung des Systems in ein Gesamtkonzept. Ein wirklich gut zu benutzendes Recherchesystem endet nicht bei der Ausgabe der Suchergebnisse in einer ansprechenden und übersichtlichen Form, sondern sollte darüber hinaus dem Suchenden die Rechercheergebnisse in einem für die eventuelle Weiterverarbeitung optimierten Format anbieten. So geht es bei der Suche in Bibliothekskatalogen nicht nur um die Bestandsfeststellung, sondern oftmals auch um die Erstellung oder Ergänzung einer Literaturliste, z.B. für einen wissenschaftlichen Aufsatz. Eine Übernahme der Rechercheergebnisse in einem Standardformat (z.B. BIBTEX) würde hier den Rechercheprozeß in den gesamten Arbeitsablauf des Suchenden sinnvoll einbetten.

Wie eine solche intelligente Schnittstelle aussehen und wie sie umgesetzt werden könnte, darum soll es in den weiteren Teilen dieser Arbeit gehen. Denen, die an dieser Stelle einwenden, daß es eine solche Schnittstelle nie geben wird, hält der Autor frei nach den Bremer Stadtmusikanten entgegen: „Etwas Besseres als den OPAC finden wir allemal!“.

3 Das Bibliothekssystem OSIRIS

3.1 Konzeption des OSIRIS-Systems

Das OSIRIS-Projekt hatte zum Ziel, die in 2.2 geschilderte Situation der Bibliotheksnutzer zu verbessern. Seit Herbst 1997 steht OSIRIS allen Angehörigen der Universität Osnabrück im Internet zur Recherche im Osnabrücker Bibliothekskatalog zur Verfügung. Seit der Präsentation auf der 50. Frankfurter Buchmesse 1998 ist OSIRIS weltweit zugänglich. Mittlerweile steht OSIRIS in der Version 3.0 öffentlich zur Verfügung – auf diese Version bezieht sich die folgende Systembeschreibung.

Eine der grundlegenden Voraussetzungen des OSIRIS-Projektes war es, ohne manuelle Bearbeitung der Daten auszukommen. Angesichts von ca. 600.000 Datensätzen in der Osnabrücker Universitätsbibliothek wäre die Anreicherung der Datensätze per Hand um z.B. zusätzliche Schlagwörter nicht bezahlbar. Da darüber hinaus der Nutzen dieser Schlagwörter für die Recherche (vgl. S. 37) nur sehr begrenzt ist, stand fest, daß OSIRIS ausschließlich auf der Grundlage des vorhandenen Datenbestandes und mit automatischen Mitteln eine qualitative Verbesserung der Literatursuche des Endbenutzers erreichen muß.

OSIRIS trägt konkret mit zwei Dingen zur Verbesserung der geschilderten Lage bei: einerseits mit einem robusten Endbenutzerzugang zur Datenbank, der die in 2.1 gezeigten Schwächen bestehender Schnittstellen zu vermeiden sucht. Dazu ist es möglich, mit bestimmten Einschränkungen natürlichsprachliche Eingaben zu tätigen. Der zweite Aspekt mit dem OSIRIS dem Suchenden hilft ist eine intelligente, automatische Aufbereitung des verfügbaren Datenbestandes. Auf diese Weise kann jemand der mit OSIRIS sucht sicher sein, daß das Ergebnis seiner Anfrage vollständig die Menge der im Bestand enthaltenen relevanten Treffer umfaßt.

In OSIRIS kann ein Benutzer Suchanfragen in einem gewissen Rahmen so eingeben, wie er sie als Aufforderung gegenüber einem Menschen ausdrücken würde: Zu diesem Zweck wird ein Satzanfang auf der Oberfläche vorgegeben („*Suche Literatur zum Thema ...*“), den der Suchende durch das Thema seiner Suche ergänzen kann.

Die Möglichkeit, natürlichsprachliche Eingaben zu tätigen, betrifft ausschließlich die Literaturrecherche: OSIRIS ist kein Expertensystem im Sinne eines allgemeinen Bibliotheksauskunftssystems. Die Eingaben, die OSIRIS zu verarbeiten hat, werden also stets Suchanfragen nach bestimmten Büchern oder Büchern zu einem Thema sein. Als Vervollständigung des vorgegebenen Satzanfanges sind daher Nominalphrasen zu erwarten, wie sie als Titel wissenschaftlicher Arbeiten oft verwendet werden: *Mathematikunterricht in der gymnasialen Oberstufe, Subventionspolitik und Steuerreform* oder *Automatische Datenerfassung in deutschen Bibliotheken*.

Die Möglichkeit einer völlig unbeschränkten Texteingabe durch den Benutzer erscheint angesichts der Aufgabe von OSIRIS, Buchtitel zu recherchieren, unangemessen. Solange der Aufgabenbereich von OSIRIS auf die Literaturrecherche eingeschränkt bleibt, würde sich der Freiraum des Benutzers im Falle unbeschränkter Ein-

Abbildung 18: Thematische Recherche mit OSIRIS

gabe auf die Variation der Suchintention beschränken: „Suche Literatur zum Thema Lachmöwenpopulationen im Binnenland“, „Suche Literatur zu Lachmöwenpopulationen im Binnenland“, „Wo finde ich Literatur zum Thema Lachmöwenpopulationen im Binnenland“, „Gibt es hier Literatur zu Lachmöwenpopulationen im Binnenland“, „Haben Sie etwas zu Lachmöwenpopulationen im Binnenland“ usw. Beide, der Suchende wie das System, wären einer unnötig höheren Belastung ausgesetzt. Der Nutzer hätte es mit einem Interface zu tun, das ohne für ihn erkennbaren Mehrwert dem oft zitierten Vorurteil, natürlichsprachliche Schnittstellen seien langatmig und aufwendig zu bedienen (Androutsopoulos et al., 1995), gerecht würde. Umgekehrt muß eingeräumt werden, daß die adäquate Verarbeitung beliebigen natürlichsprachlichen Inputs derzeit technisch nicht möglich erscheint – zumindest nicht unter den Rahmenbedingungen (Verarbeitungszeit, Robustheit usw.), die für ein System wie OSIRIS gelten (vgl. 4.1).

Aus diesen Gründen erscheint die Vorgabe eines Satzanfanges, den der Benutzer zu vervollständigen hat, als eine sinnvolle und akzeptable Möglichkeit, die Benutzerinteraktion zu beschränken. Ohne daß der Benutzer gezwungen wird, sich an eine künstliche Eingabesyntax zu gewöhnen, wird er unbewußt zur Formulierung seiner Eingabe mit bestimmten syntaktischen Strukturen geleitet, weil ein vorgegebener Satzanfang wie „Suche Literatur zum Thema ...“ syntaktisch korrekt nur durch eine Nominalphrase zu vervollständigen ist. Nominalphrasen, zu denen auch Eigennamen gezählt werden („China“), können natürlich intern eine komplexere Struktur aufweisen. Insbesondere sind Ergänzungen mit Hilfe von Präpositionalphrasen oder adverbialen Bestimmungen des Ortes oder der Zeit möglich („China zur Zeit der Kulturrevolution“, „Pädagogik in Frankreich“, „Abfallwirtschaft mit dem neuen Gemeindesteuerecht“). Solange jedoch eingebettete Nebensätze ausgeschlossen werden, stellen Nominalphrasen an der Eingabeschnittstelle eine deutliche Reduktion der Komplexität und Ambiguität dar – die Analyse wird so mit Hilfe eines relativ kleinen, effizienten Parsers, der

auf die Verarbeitung komplexer Nominalphrasen hin optimiert wurde, möglich. Die Systembenutzung wird auf diese Weise nicht nur sehr viel einfacher und angenehmer, sondern durch die computerlinguistische Analyse der Benutzereingabe werden die Treffer im Bestand auch sehr viel genauer: OSIRIS behandelt die Eingabe nicht als arbiträre Zeichenkette, sondern analysiert sie mit Hilfe einer Grammatik als Teil eines definierten Sprachausschnittes (vgl. 3.2.4). Deshalb ist es ein Unterschied, ob ein Benutzer nach Literatur zum Thema *Fernsehen in Afrika* oder nach Literatur zum Thema *Afrika im Fernsehen* fragt. Der OPAC würde aus beiden Anfragen unter Verwendung einer Stopwortliste, die u.a. Präpositionen und Artikel enthält, schlicht `Fernsehen AND Afrika` machen.

Die in OSIRIS eingesetzte Endbenutzerschnittstelle hat also die Aufgabe, natürlich-sprachliche Anfragen des Benutzers syntaktisch und semantisch so zu analysieren, daß mit dem Ergebnis der Analyse eine Suche auf der Datenbank erfolgreich durchgeführt werden kann.

Dazu müssen Phrasen erkannt und die sie konstituierenden Teile semantisch zueinander in Relation gesetzt werden. Komposita müssen in ihre Bestandteile zerlegt werden, um sie analysieren zu können (s. 3.2.6). Flektierte Formen müssen normalisiert werden, d.h. je nach Ansetzungspraxis muß Plural auf Singular (für das Deutsche) oder Singular auf Plural zurückgeführt werden, um Suchergebnisse untereinander und mit dem kontrollierten Vokabular bibliothekarischer Datensätze vergleichen zu können. Weiterhin können, als Vorbereitung einer eventuell mehrstufigen Datenbankanfrage, Synonyme, Hypo- und Hyperonyme sowie eventuell einfache Kategorieninformationen bereitgestellt werden (s. S. 162).

Da mit dem Gegenstand der Eingabe, der Literatursuche, keine sicherheitsrelevanten Aspekte einhergehen, muß nicht jede Benutzereingabe auch tatsächlich erfolgreich analysiert werden können. Natürlich ist es für die Benutzerakzeptanz unabdingbar, daß jede auf eine Anfrage gegebene Antwort korrekt ist, aber es ist umgekehrt nicht zwingend notwendig, auch wirklich zu jeder dem System gestellten Anfrage eine Antwort anbieten zu können.

Um die im Alltagseinsatz der Bibliothek nötige Robustheit gegenüber fehlerhaften Eingaben zu erlangen, muß OSIRIS in der Lage sein, einfache Buchstabenvertauschungen, Verdoppelung einzelner Buchstaben sowie Benutzung von auf der Tastatur benachbarten Tasten zu erkennen (vgl. S. 66). Dies wird durch der Eingabeverarbeitung nachgeschaltete Heuristiken erreicht (s. Abbildung 19, vgl. auch 3.2.7).

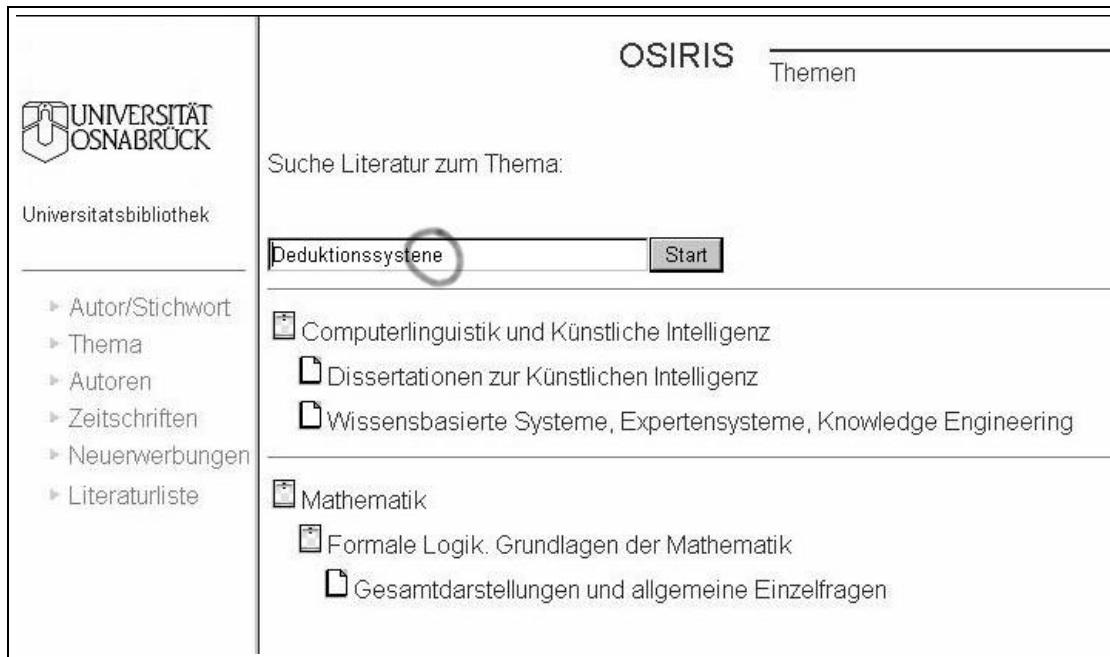


Abbildung 19: Tippfehlerkorrektur in OSIRIS

Insbesondere wird OSIRIS auf diese Weise tolerant gegenüber Rechtschreibreformen. Durch die regelbasierte Eingabeanalyse (vgl. 3.2.4) ist OSIRIS ohne Änderung der Daten in der Lage, beispielsweise *Schiffahrt* (mit drei *f*, neue Rechtschreibung) auf *Schiffahrt* (bislang mit zwei *f*) abzubilden.

Weiterhin hilft OSIRIS bei der Suche nach Personennamen. Hier erwartet OSIRIS nicht vom Benutzer, daß er sich an bibliothekarische Eingabekonventionen gewöhnt, wie sie für den OPAC gezeigt wurden (vgl. S. 33). Im Gegensatz zum OPAC ist es in OSIRIS möglich, Personennamen sowohl in der natürlichen Form Vorname Nachname, als auch „bibliothekarisch korrekt“ in der invertierten Form Nachname, Vorname zu suchen.

Insbesondere hilft OSIRIS dem Suchenden aber auch im Bereich der notorisch schwierigen Schreibung von Eigennamen. So kann in OSIRIS erfolgreich eine Autorin Namens *Isabelle Ebers* gesucht werden, obwohl die betreffende Person *Isabella Schneider-Eberz* heißt. Dies ermöglichen phonetische Algorithmen (vgl. 3.2.9), die in einer Zwischenrepräsentation ähnlich klingende Namensbestandteile aufeinander abbilden (s. Abbildung 20).



Abbildung 20: Phonetische Suche in OSIRIS

Die angesprochene intelligente Aufbereitung des vorhandenen Datenbestandes geschieht, indem einmalig aus dem im OPAC vorhandenen Datenbestand eine OSIRIS-Wissensbasis generiert wird, auf der im laufenden Betrieb die Suchanfragen prozessiert werden. Dieser Wissensbasis werden regelmäßig Informationen über neue oder geänderte Datensätze hinzugefügt. Die Informationen, die in dieser Wissensbasis zur Verfügung stehen, stammen im wesentlichen aus dem OPAC. In diesem sind mehr Informationen enthalten, als sie der Endbenutzer normalerweise zu sehen bekommt: Neben den üblichen Angaben zu Verfasser und Titel finden sich im Idealfall noch deutsche wie fremdsprachige Schlagwörter, Notationen nicht lokaler Klassifikationssysteme und anderes mehr. Daß diese verborgen bleiben, liegt einerseits daran, daß der Endbenutzer bereits mit den jetzt im OPAC verfügbaren Möglichkeiten überfordert ist (vgl. S. 29). Andererseits, und dies ist ein noch viel wichtigerer Grund, liegen die genannten Informationen zumeist nicht flächendeckend vor, wie am Beispiel der Schlagwörter bereits erörtert wurde (vgl. S. 37). Diese Informationen werden von OSIRIS *vollständig* gesammelt und wie im Folgenden detailliert erläutert ausgewertet und angereichert.

Die Auswertung der Daten beruht darauf, daß ein Fachreferent der Bibliothek ein neu erworbenes Buch einer Klasse der lokalen Systematik zuordnet. Beispielsweise kann dies für ein Buch mit dem Titel *Lie groups and quantum mechanics* von *Peter Cunningham* die Notation *TEN* sein. Dieselbe Arbeit aber hat in einer anderen Bibliothek eventuell bereits ein anderer Fachreferent getan: So ist es gut möglich, daß z.B. in der Deutschen Bibliothek in Frankfurt dem Datensatz die Information angehängt wurde, daß das Buch inhaltlich mit den Themen *Quantenmechanik* und *Lie'sche Gruppe* zu tun hat. Eventuell wurde das Buch auch schon bei Erscheinen klassifiziert. Dann wird in der Library of Congress ein Mitarbeiter festgestellt haben, daß der Autor sich in diesem Werk mit den Gebieten *quantum mechanics* und *lie groups* beschäftigt. OSIRIS wertet diese Informationen systematisch aus und stellt fest, daß von den Bü-

chern, die in Osnabrück in der Systemstelle *TEN* stehen und Schlagwörter tragen, signifikant viele mit dem Schlagwort *Quantenmechanik* versehen sind. Deutlich weniger Exemplare tragen fremdsprachige Schlagwörter, unter diesen aber sind wiederum signifikant viele mit dem Schlagwort *quantum mechanics* versehen. Offenbar werden in der Systemstelle *TEN* die Bücher klassifiziert, die sich inhaltlich mit dem beschäftigen, was im Deutschen *Quantenmechanik* und im Englischen *quantum mechanics* genannt wird. Das gilt dann natürlich auch für die Bücher, für die keine Schlagwörter vergeben wurden, die aber ebenfalls unter *TEN* klassifiziert wurden. Damit OSIRIS auf dem gesamten Buchbestand suchen kann, ist es also nicht notwendig, daß alle Bücher tatsächlich mit Schlagwörtern versehen sind. Vermutlich reicht ein Anteil von etwa 30% verschlagworteten Titeln aus, damit OSIRIS erfolgreich eine Wissensbasis aufbauen kann (Recker et al., 1996).

Finden sich weitere Bücher von *Peter Cunningham* im Bestand, von denen einige ebenfalls das Thema *Quantenmechanik* oder auch *quantum mechanics* behandeln, so kann daraus geschlossen werden, daß es sich offenbar um ein Arbeitsgebiet des Autors handelt.

Mit Hilfe der geschilderten Analyse der Benutzeranfrage gelingt es OSIRIS, einerseits die Benutzeranfrage in eine Anfrage zu überführen, die über die eigentliche Eingabe hinaus sehr viel bessere Chancen hat, Treffer zu erzielen („query expansion“, vgl. 3.3). Andererseits sorgt die Aufbereitung des Datenbestandes dafür, daß systematisch mehr Material zur Verfügung steht, auf das die Eingabe abgebildet werden kann.

Von besonderer Wichtigkeit ist, daß OSIRIS die klassifikatorische Erschließung der Titel (vgl. 2.2, S. 39) sowohl zur Suche, als auch zur Präsentation der Suchergebnisse nutzt. Das Ergebnis einer Suche in OSIRIS zum Thema *Deduktionssysteme* ist in Abbildung 21 wiedergegeben.

Die Präsentation der Suchergebnisse in OSIRIS ist dreigeteilt. Zuerst werden die Treffer angezeigt, die aufgrund von pattern matching im Titelmateriale entdeckt wurden. Dieses Vorgehen hat zwei Gründe: Zum einen trägt dies der (nachweisbaren) Erwartung der Nutzer Rechnung, Bücher, die einen mit dem Thema identischen Titel haben, unabhängig von deren Inhalt auch bei einer thematischen Recherche zu finden. Ohne diese Suche kann es zu Diskrepanzen kommen, wenn ein Buch mit einem der thematischen Recherche entsprechenden Titel aufgrund einer besonderen inhaltlichen Ausprägung in eine thematisch nicht verwandte Klasse der Systematik eingeordnet wurde. Wenn der Benutzer beispielsweise nach Literatur zum Thema *Alterssport* sucht, dann wird ihm ein Buch mit diesem Titel, das sich aber mit den kriminalistischen Neigungen alter englischer Damen beschäftigt und deshalb in die Klasse *Belletristik/Kriminalroman* eingeordnet wurde, „vorenthalten“. Diese Strategie, die OSIRIS in früheren Versionen verfolgte, führte zu regelmäßigem Protest der Suchenden, die sich offenbar anhand solcher Titel rückversichern wollten, daß das System auch tatsächlich nach ihren Vorgaben gesucht habe. Daß OSIRIS auf der Klassifikation und nicht auf dem Titelmateriale sucht (d.h. die Unterscheidung in Formal- und Sachrecherche), war nur in wenigen Fällen zu vermitteln. Der zweite Grund, weshalb OSIRIS die mit Hilfe von pattern matching erzielten Treffer anzeigt, ist der Vergleich mit dem OPAC. Die

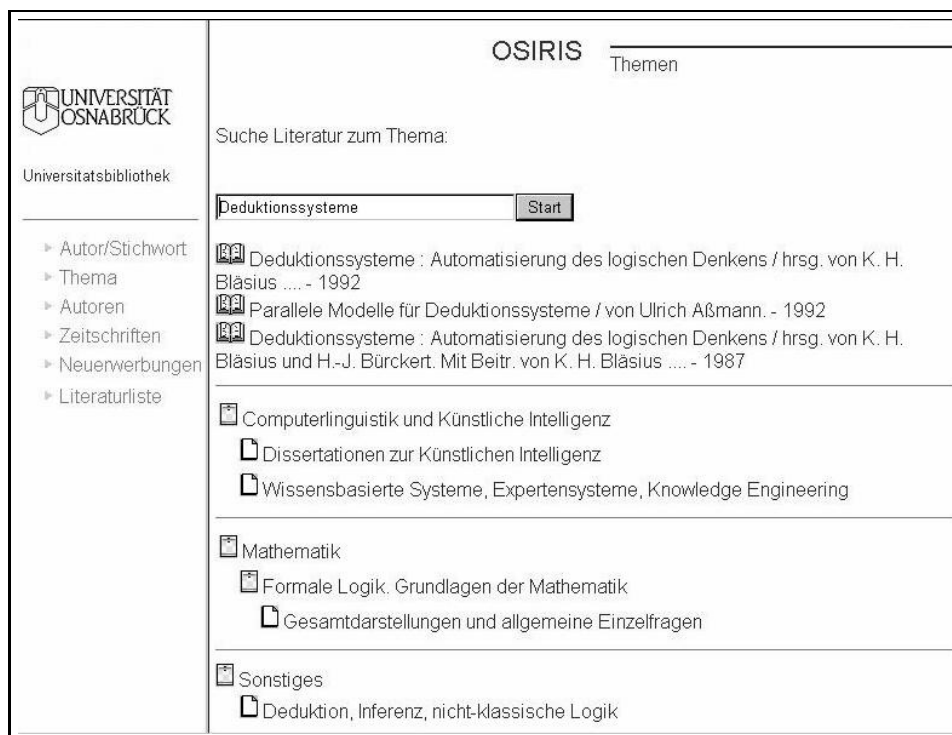


Abbildung 21: Die Darstellung des Suchergebnisses

in Abbildung 21 am Beginn der Ergebnisliste gezeigten 3 Titel sind exakt die Treffer, die der OPAC (weil ebenfalls mit pattern matching operierend) erzielt hätte. Neben der Befriedigung eines expliziten Nutzerbedürfnisses dient dieser erste Teil der Ergebnispräsentation also auch zugleich der Demonstration des Fortschrittes, den OSIRIS gegenüber der Suche mit dem OPAC darstellt.

Der mittlere Teil des Suchergebnisses präsentiert *Klassen* der Systematik, in denen für die Benutzereingabe relevante Literatur vermutet werden kann. Diesen Klassen der Systematik sind Werke zugeordnet, die sich inhaltlich mit dem beschäftigen, was als Klassenbezeichnung angegeben ist, genauer: Der Klassenname ist die natürlichsprachliche Bezeichnung der Notation, die das inhaltliche Erschließungsmerkmal darstellt, auf der die klassifikatorische Erschließung des Bestandes beruht (vgl. S. 39). Aufgrund der Suche über der Klassifikation und der Ergebnispräsentation in Form von Klassen derselben wird die Benutzeranfrage bereits frühzeitig disambiguiert. Eine Anfrage nach Literatur zum Thema „Altersvorsorge“ führt so zu verschiedenen Klassen, nämlich u.a. zu „Rentenversicherung, Alterssicherung, Altenhilfe“ in den Wirtschaftswissenschaften, „Sozialpolitik und Geschlechterrolle. Frauen und Armut“ in der Frauenforschung, in den Geschichtswissenschaften zur „Sozialfürsorge. Sozialpolitik“ sowie in den Rechtswissenschaften zum „Familienrecht“.

Der untere, dritte Teil der Ergebnispräsentation besteht aus einer virtuellen Klasse. Hinter dieser verbergen sich die realen Klassen der Systematik, die aufgrund der gefundenen Titel im ersten Teil der Ergebnispräsentation als relevant erkannt wurden

(und noch nicht im mittleren Teil aufgeführt sind).

Ein Beispiel: Sucht ein Benutzer nach Literatur zum Thema *Semitistik*, so wird er im ersten Teil die Titel sehen, die das Wort *Semitistik* enthalten (klassische Titelstichwortsuche des OPAC).

Im mittleren Teil werden die Klassen der lokalen Systematik präsentiert, die sich inhaltlich mit *Semitistik* beschäftigen (*Orientalistik*). Nun wird überprüft, welchen Klassen die aufgrund der Titelstichwortsuche gefundenen Werke im ersten Teil der Ergebnispräsentation angehören. Dabei werden zum Teil genau die Klassen gefunden werden, die aufgrund der Suche in der lokalen Klassifikation bereits getroffen und dem Benutzer präsentiert wurden. Zum Teil werden Unterklassen identifiziert, die indirekt, nämlich hinter einer bereits präsentierten Klasse dem Benutzer gezeigt werden. Es werden aber auch Titel dabei sein, die aus Klassen stammen, die OSIRIS nicht aufgrund der Systematik als relevant erkannt hat. Für die Suche nach *Semitistik* wird die Klasse „Umwelt. Bildwerke. Atlanten. Geographie“ gezeigt. Aufgrund der Notation hat diese Klasse auf den ersten Blick nichts mit *Semitistik* zu tun, jedoch finden sich hier zahlreiche einschlägige Werke wie „Orte und Landschaften der Bibel“, „Atlas zur Bibel“ oder auch „Beiträge zur Arabistik, Semitistik und Islamwissenschaft“.

Zusammenfassend geht OSIRIS also drei Wege bei der Suche im Datenbestand: Zum einen liefert die klassische Titelstichwortsuche einzelne Titel als Treffer, zum anderen wird über der lokalen Systematik nach relevanten Klassen gesucht. Und schließlich wird die Trefferlage der Titelstichwortsuche in Abhängigkeit von der Suche in der Systematik für einen Verweis auf weitere Klassen der Systematik genutzt.

Die Präsentation in Form von Klassen der Systematik bringt es mit sich, daß erst ein Mausklick auf das Symbol neben einem Klassennamen den Benutzer zu konkreten Titeln führt, nämlich zu den Titeln, die dieser Klasse vom Bibliothekar zugeordnet wurden (s. Abbildung 22).

The screenshot shows the OSIRIS search interface. On the left is the University of Osnabrück logo and a navigation menu with options like 'Autor/Stichwort', 'Thema', 'Autoren', 'Zeitschriften', 'Neuerwerbungen', and 'Literaturliste'. The main area displays the search results for the query 'deduktionssysteme'. The results are listed under the heading 'Wissensbasierte Systeme, Expertensysteme, Knowledge Engineering' and include titles such as 'The essence of artificial intelligence', 'Expert systems and probabilistic network models', 'Data warehouse', 'Maschinelles Lernen und wissensbasierte Systeme', 'Applications of logic databases', 'Datenbankeinsatz', 'Diagrammatic reasoning', and 'Wiederverwendung von Commonsense Ontologien im Knowledge Engineering'.

Abbildung 22: Titel in der Klasse *Wissensbasierte Systeme*

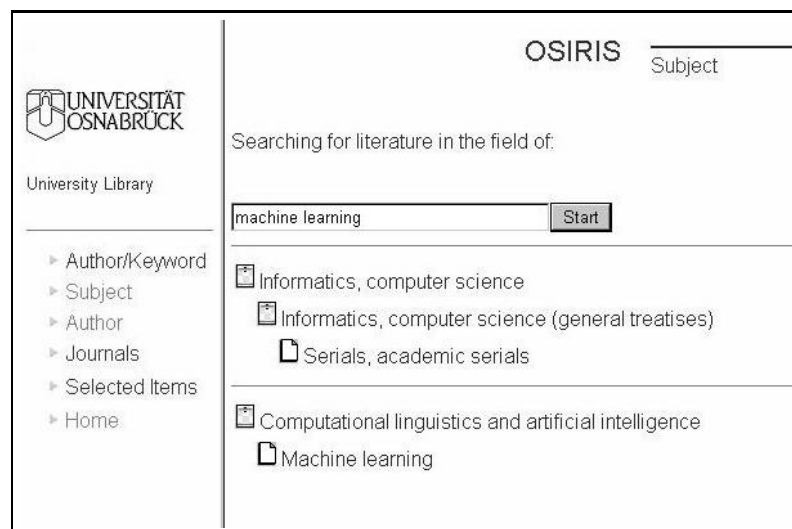
Ein großer Vorteil für den Nutzer besteht darin, daß durch die Suche auf der OSIRIS Wissensbasis deutsche wie fremdsprachige Fachliteratur gefunden wird. Da die klassifikatorische Erschließung nicht auf das Titelmateriale, sondern auf den Inhalt des Buches referiert, werden unabhängig davon, ob mit einem englischen oder einem deutschen Begriff gesucht wird, deutsch- und englischsprachige Titel gemeinsam gefunden. Durch die Aufbereitung der Daten in OSIRIS ist dies flächendeckend möglich. Ein Mausklick auf das Symbol neben der Titelangabe führt zur Vollanzeige des Titels (vgl. Abbildung 23). Hier besteht die Möglichkeit zu überprüfen, ob das entsprechende Buch bereits ausgeliehen ist. Die angezeigten Titel können auch in eine persönliche Literaturliste übernommen werden, z.B. um sich selbst ein Literaturverzeichnis anzulegen.

Die Verknüpfung von Elementen klassifikatorischer und verbaler Erschließung sowie deren Aufbereitung in der OSIRIS Wissensbasis macht es möglich, daß für eine deutsche Anfrage auch relevante englischsprachige Literatur gefunden wird. Umgekehrt bedeutet dies, daß OSIRIS aufgrund der Struktur der Wissensbasis alle Voraussetzungen für eine englische Version mit sich bringt. Aus diesem Grunde wurden die Bezeichnungen der Klassen der lokalen Systematik ins Englische übersetzt, um eine adäquate Präsentation der Suchergebnisse zu ermöglichen. Weiterhin notwendig sind natürlich die entsprechenden sprachspezifischen Teile der Eingabeverarbeitung, nämlich ein englisches Lexikon und eine englische Grammatik (vgl. 3.2.4). Zusammen mit entsprechend aufbereiteten Eingabemasken öffnet OSIRIS so den Osnabrücker Bibliothekskatalog für englischsprachige Benutzer, ohne daß aufwendige (und feh-



Abbildung 23: Titelvollanzeige

leranfällige) Übersetzungswörterbücher gepflegt werden müßten.

Abbildung 24: Titel in der Klasse *Wissensbasierte Systeme*

Die Anfrage nach „machine learning“ wird dabei ausschließlich auf den englischsprachigen Schlagwörtern der OSIRIS-Wissensbasis verarbeitet und führt den englischsprachigen Benutzer zur Klasse „Maschinelles Lernen“ in der Osnabrücker Bibliothek. Dorthin wird natürlich auch der deutschsprachige Nutzer geleitet, der auf der deutschen Oberfläche von OSIRIS den aus dem Englischen übernommenen Fachterminus eingibt.

Neben der thematischen Recherche ist in OSIRIS selbstverständlich auch die Suche nach bestimmten Werken möglich. Um die für den OPAC gezeigten Schwierigkeiten (vgl. 2.2) der Benutzer bei der Durchführung einer kombinierten Suchanfrage zu

umgehen, reagiert OSIRIS auch hier flexibler und robuster als herkömmliche OPAC-Systeme. So ist es in OSIRIS möglich, auf verschiedene Arten nach dem Werk eines Autors zu suchen. Vorbild war dabei die Art und Weise, wie man in einer Buchhandlung nach diesem Werk fragen würde (s. Abbildung 25).

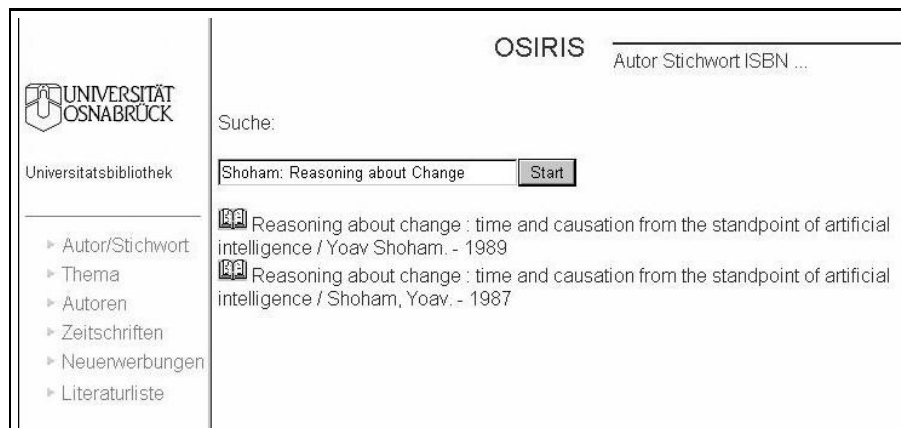


Abbildung 25: Autor-Stichwort-Suche

Alternativ kann nach demselben Werk auch mit *Reasoning about Change* von Shoham sowie *Shoham, Reasoning about Change* oder einfach *Shoham Reasoning about Change* gesucht werden. Selbstverständlich kann auch nur der Titel des Werkes (ohne den Verfasser) eingegeben werden. OSIRIS schaltet dabei automatisch in die Themensuche um, wenn das Trefferbild dies nahelegt d.h. sehr viele Treffer aus relativ wenig unterschiedlichen Klassen der Systematik gefunden werden. Wie bereits am Beispiel der Präsentation des Ergebnisses der Themenrecherche ausgeführt (siehe 3.1), ist die bibliothekarisch übliche Unterscheidung in Formal- und Themenrecherche den meisten Nutzern nicht verständlich. Der automatische Wechsel aus der Titel- in die Literatursuche sowie (in der entgegengesetzten Richtung) die Präsentation von Titeln innerhalb der Ergebnisse der thematischen Recherche entspricht den Erwartungen der Benutzer.

Insbesondere verbessert OSIRIS die Formalrecherche in hierarchischen Strukturen wie mehrbändigen Werken. Mehrbändige Werke stellen eine weitere Schwierigkeit für den bibliothekarischen Laien dar, indem sie in deutschen Bibliotheken (und nur dort) als hierarchische Strukturen behandelt werden und so für Laien schwer auffindbar sind. So ist im OPAC ein Werk in fünf Bänden wie das *Grünbuch Ökologie*, herausgegeben von Jürgen Grumbach nur unter seinem übergeordneten Titel zu finden. Eine Titelstichwortsuche nach Grünbuch AND Ökologie wird also genauso Erfolg haben wie eine kombinierte Suche nach dem Autor und den Stichwörtern Grünbuch oder Ökologie. Was dagegen nicht möglich ist, ist die kombinierte Suche nach Autor und Titel eines einzelnen Bandes oder die Kombination aus Stichwörtern des Gesamttitels und des Einzelbandtitels. Der fünfte Band des genannten mehrbändigen Werkes trägt den Titel *Tschernobyl und die Konsequenzen: Arbeit und Umwelt; Waldsterben;*

Landwirtschaft und Umweltschutz; vor Ort: Wackersdorf. Eine Suche nach Grumbach: Waldsterben führt nicht zum Erfolg, weil es sich um die Verknüpfung eines Elementes des übergeordneten Titels (*Grumbach*) mit einem Stichwort des Einzelbandes handelt. Aus dem gleichen Grunde bleibt eine Anfrage nach *Grünbuch Waldsterben* ohne Erfolg.

Angesichts dieses Beispiels mag man geneigt sein, das Problem der hierarchischen Strukturen mehrbändiger Werke niedrig einzuschätzen, schließlich ist ein Gesamttitel wie *Grünbuch Ökologie* an sich bereits aussagekräftig und darüber hinaus auch einprägsamer als der längliche Titel eines Einzelbandes. Kritisch wird die Situation jedoch dadurch, daß (wie für viele mehrbändige Werke üblich) Einzelwerke bekannter Autoren unter dem Titel *Gesammelte Werke* o.ä. zusammengefaßt werden. In diesen Fällen ist es gerade der Titel des Einzelbandes, der bekannt ist und der nun in Kombination mit dem Autorennamen nicht mehr gefunden werden kann.

In OSIRIS wurde zur Lösung dieses Problems ein Mandi¹⁴-Suchalgorithmus implementiert, mit dessen Hilfe Suchbegriffe aus übergeordneten und untergeordneten Ebenen zusammengeführt werden. So kann in OSIRIS erfolgreich nach *Grünbuch Waldsterben* gesucht werden.¹⁵ Der Benutzer profitiert von stark verbesserten Suchmöglichkeiten in redundanzfrei gehaltenen bibliographisch hierarchischen Strukturen, ohne daß der (wenig erfolgversprechende) Versuch unternommen worden wäre, die Aufnahme mehrbändiger Werke in deutschen Bibliotheken an sich ändern zu wollen.

¹⁴Mandi=Multivolume **And** Instruction

¹⁵Allerdings befindet ist diese Funktionalität in OSIRIS Version 3.0 noch in der Testphase und ist aus Performanzgründen eingeschränkt.

3.2 Aufbau und Funktion der Dialogschnittstelle

3.2.1 Überblick über die Systemkomponenten

Das OSIRIS-System ist modular aufgebaut, d.h. alle Komponenten kommunizieren über Internet-Sockets miteinander, so daß die gesamte Anwendung verteilt in einem Netz unter Einbeziehung verschiedener Betriebssysteme arbeitet. Dabei zerfällt OSIRIS in zwei wesentliche Teile, nämlich einerseits die Dialogschnittstelle mit den computerlinguistischen Analysekomponenten für die Verarbeitung der Benutzereingabe und andererseits die OSIRIS-Wissensbasis, die die gesamte Retrievalfunktionalität für die Datenbank sowie die Bewertung der Suchergebnisse enthält. Die Wissensbasis präsentiert sich den computerlinguistischen Komponenten logisch als ein Index aus Schlagwörtern und Ansetzungsformen der Systematik der Bibliothek.

In diesem Abschnitt wird es ausschließlich um die Dialogschnittstelle gehen. Sie ist seit der Version 2.0 mit Hilfe einer Blackboardarchitektur realisiert, deren Kernkomponenten der Task Manager, der Parser und das Lexikon sind. Diese Komponenten erstellen aus dem natürlichsprachlichen Input eine geeignete semantische Analyse für eine Anfrage an die Wissensbasis, deren Ziel die gewichtete Abbildung auf eben jenen Index ist.

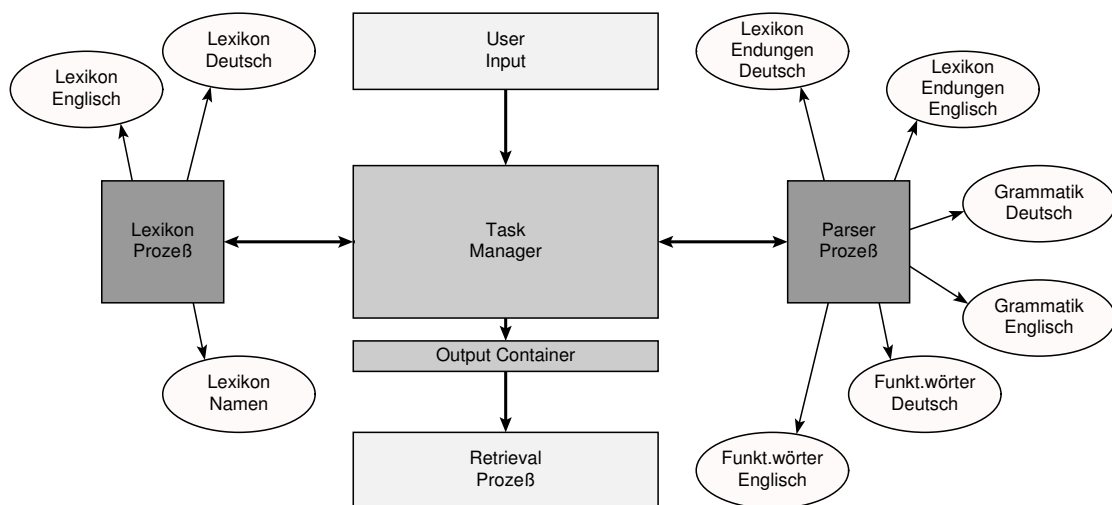


Abbildung 26: Schematische Darstellung der Eingabeverarbeitung

Im Folgenden sollen diese Komponenten im Detail vorgestellt und ihr Zusammenspiel verdeutlicht werden.

3.2.2 Der Task Manager

Der Task Manager ist für die syntaktische Vorverarbeitung verantwortlich. Er bekommt von der Oberfläche die Benutzereingaben an einen Port geschickt und startet daraufhin

einen separaten Prozeß. Der Port wird währenddessen weiterhin überwacht, um ggfs. für eingehende Anfragen anderer Benutzer weitere Prozesse des Task Managers starten zu können.

Die OSIRIS-Komponenten arbeiten auf einem dedizierten Rechner, der mit dem eigentlichen Datenbankserver über HTTP¹⁶ d.h. stateless kommuniziert. Da HTTP ein zustandsloses Protokoll ist, ist das Verhalten des OSIRIS-Systems nur stringent und an sich nicht unbedingt bemerkenswert. Jedoch steht OSIRIS dabei im Gegensatz zum bereits gezeigten OPAC-System, daß nicht-stateless operiert: D.h. der OPAC nimmt auf dem nächsten freien Port eine Verbindung entgegen und hält diese offen, bis der Benutzer sie explizit beendet oder aber eine gewisse Zeit ohne Interaktion vergangen ist („time-out“). Nimmt der Suchende länger als z.B. fünf Minuten keine Eingaben vor, so wird dies von Seiten des OPAC als Endesignal interpretiert und die Verbindung wird geschlossen. Einem Benutzer, der während seiner Literatursuche andere Arbeiten einschiebt, bietet sich deshalb regelmäßig das in Abbildung 27 gezeigte Bild.

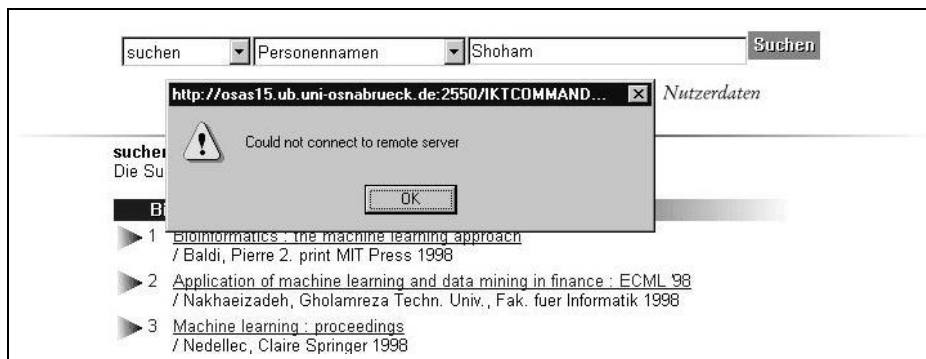


Abbildung 27: Timeout des nicht-stateless operierenden WWW-OPAC

Für eine erneute Anfrage muß erst eine neue Verbindung (auf einem neuen Port) zum OPAC aufgebaut werden.

Zu den Aufgaben des Task Managers in OSIRIS gehört die wortweise Befragung des Lexikons sowie die Reduktion und Wiedervorlage von als unbekannt zurückgewiesenen Wörtern. Allgemein finden im Task Manager die Manipulationen statt, die nicht direkt während des Lexikonzugriffs realisiert werden können.

Der für eine Benutzereingabe wie *die mittelalterliche Burg* gestartete Task Manager Prozeß wertet die übergebene Datenstruktur aus und kommuniziert in einem ersten Schritt mit dem Lexikon. Die Ergebnisse des Lexikonprozesses werden durch den Task Manager sortiert und eine Auswahl dem Parserserver übergeben. Die Ergebnisse des Parsingprozesses und der lexikalischen Analyse werden nach verschiedenen, parametrisierbaren Heuristiken durch den Task Manager bewertet (vgl. 3.2.10) und der OSIRIS-Wissensbasis in Form einer Attribut-Wert-Struktur als Grundlage für die Suchanfrage übergeben (vgl. Abbildung 30, S. 72).

¹⁶Hypertext Transfer Protocol

3.2.3 Das Lexikon

Um komplexe Nominalphrasen bestehend aus Eigennamen, Einschränkungen und Ergänzungen verarbeiten zu können, müssen grammatische Informationen zu einzelnen Wörtern der Eingabe verfügbar sein. Solche Informationen sind im Lexikon abgelegt, wo sich Angaben über z.B. Wortart, Numerus und Genus befinden. Im Lexikon befindet sich auch eine rudimentäre Semantik der Präpositionen wie *in*, *mit* etc., die eine Einschränkung oder Modifikation ausdrücken. Das Lexikon enthält Stammformeneinträge, wobei Endungen und ggf. erforderliche Umlautungen Bestandteil der lexikalischen Information im Eintrag sind. Genaugenommen verfügt OSIRIS über mehrere Lexika, da einerseits je Sprache ein eigenes Lexikon benötigt wird, und andererseits je Sprache bestimmte Informationen (z.B. Wortendungen) in eigenen Lexika ausgelagert werden (vgl. Abbildung 26, S. 56).

Beim Zugriff auf das Lexikon wird eine morphologische Reduktion vorgenommen, deren Ergebnis die Zerlegung in Endung, Lemma, erkannte Affixe sowie Kongruenzinformationen ist. Kompositazerlegungen, Endungsanalyse und Korrektur einfacher Fehlschreibungen erfolgen ebenfalls in diesem Schritt. Das Lexikon wurde von Dr. habil. Helmar Gust entwickelt und wird in überarbeiteter Form inzwischen kommerziell eingesetzt.¹⁷

Zusätzlich verfügt OSIRIS über ein spezielles Lexikon für Eigennamen. Wird ein Teil der Benutzereingabe als Eigenname erkannt, so wird dessen Analyseergebnis gesondert behandelt. Im Detail wird dies in 3.2.4 und 3.2.10 beschrieben.

Anforderungen an das Lexikon Das Einsatzgebiet „wissenschaftliche Bibliothek“ stellt hinsichtlich des Sprachumfangs der zu erwartenden Eingaben in struktureller Hinsicht eine echte Einschränkung dar, in lexikalischer Hinsicht aber nicht (zur allgemeinen Problematik von Subsprachen siehe 4.1). Die an OSIRIS gerichteten Anfragen werden stets Suchanfragen nach bestimmten Büchern oder Büchern zu einem Thema sein, syntaktisch gesehen wird es sich um Nominalphrasen handeln, die als Ergänzung des auf der Oberfläche vorgegebenen Satzanfanges entstehen.

Trotz Ausbildung von Bestandsschwerpunkten muß in einer wissenschaftlichen Bibliothek aber mit Literatur zu *allen* Wissensgebieten und das in einem Spektrum vom einführenden Werk bis hin zu Tagungsbänden und Dissertationen gerechnet werden. Deshalb muß OSIRIS mit dem gesamten aktuellen wissenschaftlichen Vokabular vertraut sein.

Die Schwierigkeiten an dieser Stelle liegen dabei nicht nur in der bloßen Menge des Vokabulars, sondern auch in der zu gewährleistenden Aktualität. Manuell zu wartende Wörterbücher sind hierfür ungeeignet, weil aufwendig zu pflegen. OSIRIS setzt dagegen auf regelbasierte Ansätze, inkrementelle und automatische Lexikonerweiterung, sowie in vielen Bereichen (wie z.B. der Kompositazerlegung) auf Heuristiken.

¹⁷Die Lexikontechnik wird in einigen der Recherche Tools der Firma ICDM (www.icdm.de) eingesetzt.

Während zu Beginn des Projektes noch mit ca. 200.000 Lexikoneinträgen für eine voll ausgebaute OSIRIS-Version gerechnet wurde, hat das deutsche Lexikon am Ende des Projektes einen Umfang von ca. 400.000 Stammformeneinträgen (Adjektive und Nomen) erreicht. Für das Englische handelt es sich derzeit um ca. 50.000 Stammformen der genannten Wortarten. Ein Eintrag meint die Stammform (Lemma) eines Wortes, die durch Deklination oder Flexion in verschiedenen anderen Wortformen auftreten kann. Endungen und ggf. erforderliche Umlautungen sind Bestandteil der lexikalischen Information. Die Menge der im OSIRIS-Lexikon enthaltenen Wortformen wird auf etwa 1.5 Millionen geschätzt. Um mit diesen Datenmengen effizient arbeiten zu können, werden die Lemmata als Suchschlüssel in einem Buchstabenbaum kodiert und die Inhalte der Einträge separat in einer Datenbank gespeichert, so daß sehr schnelle Zugriffe möglich sind.

Suffixanalyse Teil des Lexikonzugriffs bei der Inputerkennung ist die Suffixanalyse. Im Lexikon sind regelmäßig gebildete Suffixe als eigenständige Einträge abgelegt, so daß ein Wort wie *mittelalterliche* genauso wie ein Kompositum zerlegt werden kann in *mittelalter* und *liche*. Natürlich ist im Lexikon vermerkt, daß Einträge wie *liche* wohl Suffix eines Wortes sein können, aber nicht eigenständig ein wohlgeformtes Wort bilden können (zur Behandlung von Affixen als eigenständige Lexikoneinträge vrgl. z.B. Höhle (1982), Kanngießner (1985) und Wunderlich (1986)). Darüber hinaus ist dort vermerkt, welche Schlüsse sich aus dem Auftreten des Suffixes auf die Kategorie des ganzen Wortes ziehen lassen. Im Fall von *liche* kann z.B. geschlossen werden, daß es sich bei *mittelalterliche* um ein Adjektiv handelt. Weiterhin können aus der Form des abgetrennten Suffixes erste Schlüsse auf Kasus, Numerus und Genus gezogen werden. Die Annahme, ein Wort, das auf *liche* endet, sei ein Adjektiv, ist aber nur dann sicher, wenn der Wortstamm (*mittelalter*) dem Lexikon bekannt ist. Ist der Wortstamm nicht bekannt, so kann nur mit einer gewissen Wahrscheinlichkeit vermutet werden, es handle sich um ein Adjektiv. Beispielsweise geht diese Vermutung für das Nomen „Schliche“ in die Irre. Obwohl für „Schliche“ wiederum eine Heuristik möglich ist, die erkennt, daß „Sch“ kein im Deutschen wohlgeformter Wortstamm sein kann, handelt es sich dennoch stets um heuristische d.h. unsichere Schlüsse.

Details der Lexikonstruktur Die Struktur des Lexikons sieht für den Eintrag *mittelalter* wie folgt aus (Dr. habil. Helmar Gust: personal communications):

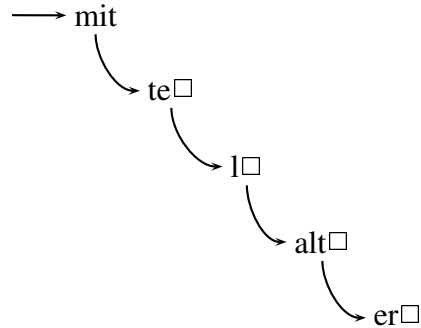


Abbildung 28: Lexikonausschnitt

mit ist das erste wohlgeformte Präfix in *mittelalter* und deshalb der erste Eintrag im Baum. Da die Präposition *mit* weder konjugiert noch dekliniert werden kann, müssen nicht eventuell vorhandene Suffixe berücksichtigt werden. Das zweite Präfix *mitte* kann dagegen ein pluralbildendes Suffix *-n* mit sich führen, so daß im Baum durch □ dem Auftreten von Suffixen Rechnung getragen wird. Konkret wird *mitte* im Lexikon durch *te*□ realisiert, da *mitte* durch das Präfix *mit* und die Fortsetzung *te* eindeutig bestimmt wird.

Natürlich stellt *mit* nicht nur ein Präfix zu *mitte* dar, sondern auch zu Worten wie *Mitella*, *Mittag*, *Mitternacht*, *Mittler* oder *Mitra*. Das Lexikon enthält also alternative Fortsetzungen und sieht dabei eher wie folgt aus:

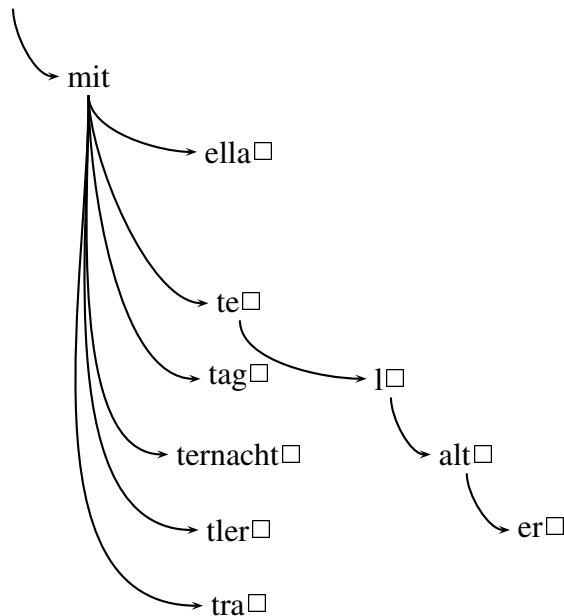


Abbildung 29: Erweiterter Lexikonausschnitt

So wie *mitverantwortliche* durch Suffixtilgung und -analyse als adjektivische Form des Wortstamms *mitverantwort-* erkannt wird, ist durch Kennzeichnung des Lexikoneintrages *mit* als Derivationspräfix *mitverantwort* auf *mit* und *verantwort* zurückzuführen.

Wird ein Eintrag wie *mittelalter* im Lexikon gefunden, so stellt sich die Frage, welche Informationen diesem Eintrag angefügt sind. Hier müssen die Wortart, sowie gegebenenfalls eine rudimentäre Semantik und weitere Informationen zu Besonderheiten der Wortbildung stehen. Die Wortart ist an dieser Stelle deshalb wichtig, weil wie bereits gezeigt, aus dem Suffix *liche* allein nur mit einer gewissen Unsicherheit auf die Wortart geschlossen werden kann.

Wird ein Eintrag wie *mittelalter* im Lexikon nicht gefunden, so werden Heuristiken eingesetzt, um dennoch zu einem Ergebnis zu kommen. Dazu wird zunächst durch rekursives Löschen von Zeichen am linken Rand versucht, einen bekannten rechten Rand zu erzeugen. Weiterhin wird die heuristische Fehlerkorrektur für die Eingabe aktiviert (siehe 3.2.7), und schließlich wird das Ergebnis der Endungsanalyse (siehe 3.2.3) verwendet, um eine Hypothese über die Wortart aufzustellen. Jedes mögliche Ergebnis dieser Schritte ist mit einer gewissen Unsicherheit behaftet, deshalb werden die so entstandenen Analyseergebnisse mit einem entsprechenden Vermerk versehen, der bei der Bewertung im Task Manager berücksichtigt wird (vgl. 3.2.10).

Allgemein gilt, daß der Index, auf dem in OSIRIS die Suchanfragen des Benutzers verarbeitet werden, eine echte Teilmenge des Lexikons darstellt. Das heißt, daß ein Suchwort, das *nicht* im Lexikon gefunden wird (auch wenn vielleicht Teile des Wortes, z.B. ein Suffix gefunden wurden), auch nicht im Index verzeichnet ist. Umgekehrt gilt aber nicht, daß ein Lexikoneintrag notwendigerweise Teil des Index sein muß.

Kommuniziert der Task Manager mit dem Lexikon, so wird vom Lexikon eine Kopie zur Bearbeitung der Anfrage des Task Managers gestartet, während der ursprüngliche Prozeß für weitere Anfragen des Task Managers zur Verfügung steht.

Ergebnis der Analyse des Lexikonprozesses ist ein temporäres Lexikon mit Einträgen für Adjektive und Nomen in einem für den Parser optimierten Format. Der Parser verfügt selbst bereits über eine Reihe eigener Lexika für Wortendungen sowie Konjunktionen, Präpositionen und Determiner (vgl. Abbildung 26, S. 56). Zusammen mit dem gelieferten temporären Lexikon verfügt der Parser für die Analyse über vollständige lexikalische Informationen zum eingegebenen Satz.

Das Ergebnis des Lexikonprozesses wird an den Task Manager Prozeß zurückgereicht und der Lexikonprozeß stirbt. Der Task Manager kommuniziert im Anschluß mit dem Parserprozeß.

3.2.4 Der Parser

Wie bereits mehrfach angesichts des eingeschränkten Aufgabenbereiches und der syntaktischen Reduktion der Benutzereingabe durch Vorgabe eines Satzanfanges festgestellt, benötigt das OSIRIS-System für die syntaktische Analyse der Benutzereingabe einen kleinen, für Nominalphrasen optimierten, robusten und effizienten Parser. Dieser

muß in der Lage sein, die wesentlichen Fälle zu verarbeiten, wobei einzelne Eingaben durchaus zurückgewiesen werden dürfen.

Der Parserprozeß analysiert auf der Grundlage einer kontextfreien Grammatik mit Merkmalsannotationen und unter Berücksichtigung der durch den Task Manager gefilterten und sortierten Ergebnisse des Lexikonprozesses die syntaktische Struktur der Eingabe.

Der Parser besteht aus zwei Teilen, einerseits den Grammatikregeln, die festlegen, wie ein korrekter deutscher bzw. englischer Satz aufgebaut ist und andererseits dem Parsing-Programm, das für eine konkrete Eingabe einer Nominalphrase überprüft, ob diese sich mit Hilfe der Grammatikregeln erzeugen läßt. Als Formalismus wird eine kontextfreie Grammatik mit Merkmalsannotationen in einer PATR II ähnlichen Notation verwendet (Shieber, 1986). Ein Beispiel für eine Grammatikregel in OSIRIS ist:

$$NP \rightarrow N' [PP]^*$$

Das heißt, eine Nominalphrase wie z.B. „Geometrie in der Oberstufe“ kann aus einem Nomen, gefolgt von einer oder mehreren Präpositionalphrasen bestehen. Diese rein syntaktische Regel wird im internen Verarbeitungsprozeß um einige andere, z.B. semantische Informationen, angereichert:

```
NP : N_1 PP [
    $0 : [
        SEM = [ PRED = { | SYN | HEAD, \, xrestricted_by, | SYN | RCOMP }; ]
        SYN = [ HEAD = [ SYN = $1:SYN; SEM = $1:SEM; ]
                RCOMP = [ SYN = $2:SYN; SEM = $2:SEM; ]
            ]
    ]
]
```

Als Analyseergebnis wird bei Verwendung dieser Regel eine sogenannte F-Struktur konstruiert:

```
( $0 : [ SEM = [ PRED = { | SYN | HEAD, xrestricted_by, | SYN | RCOMP }; ]
  SYN = [ HEAD = [ SYN = $1:SYN; SEM = $1:SEM; ] RCOMP = [ SYN = $2:SYN; SEM = $2:SEM; ] ] ),
```

In diese werden die Analyseergebnisse der Konstituenten (\$1 für N_1 und \$2 für die PP) in Anlehnung an Grammatikformalismen wie LFG oder HPSG integriert.

Die Anforderungen, die OSIRIS an den Parser stellt, bestehen vor allen Dingen in einer optimierten Behandlung von Nominalphrasen sowie geringen Ressourcenansprüchen. Prinzipiell ist OSIRIS nicht auf die Verwendung eines bestimmten Parsers festgelegt. Derzeit wird eine Reimplementation des bekannten GEPARD-Parsers eingesetzt, der von Dr. habil. Hagen Langer an der Universität Osnabrück entwickelt wurde (Langer, 1996). Der GEPARD-Parser beruht auf dem Earley-Algorithmus mit Vorwärtsverkettung und zeichnet sich insbesondere dadurch aus, daß er auch auf weniger leistungsfähigen Plattformen wie z.B. einem PC mit umfangreichen Lexika und Grammatiken

eine hohe Performanz erzielt. Dazu wird eine deklarativ formulierte, hinsichtlich der formalen Mächtigkeit aber stark eingeschränkte Grammatik durch einen mehrstufigen parametrisierbaren Compiler in ein C-Programm übersetzt, so daß im Ergebnis ein sehr effizienter, für die jeweils gegebene Eingabegrammatik optimierter Chartparser zur Verfügung steht. Ein finite-state-basiertes stochastisches Modell für Linksableitungen läßt dabei Bewertungen der Parsing-Ergebnisse zu. Dieser Parser wurde durch Wilfried Teiken mit dem Ziel überarbeitet, auch für unvollständige oder fehlerhafte Eingaben noch sinnvolle Parsing-Ergebnisse zu erhalten (Teiken, 1998). Diese robuste Variante wird mit einigen weiteren Modifikationen in OSIRIS eingesetzt: Sie ist speziell für Nominalphrasen optimiert und arbeitet deshalb im OSIRIS-Kontext besonders effizient.

Der Umfang des Sprachausschnittes, der durch die im Parser integrierte Grammatik abgedeckt wird, wurde aufgrund einer Analyse des möglichen Benutzerinputs abgesteckt. Der aktuell verarbeitbare Sprachausschnitt beinhaltet Konstruktionen folgender Bauart:

(Artikel) (Adjektiv) Nomen (Artikel) (Nomen) (Präposition) (Artikel) (Nomen)

(Artikel) (Adjektiv) Nomen (Präposition) (Artikel) (Nomen) (Artikel) (Nomen)

Dabei soll die Klammerung andeuten, daß es sich bei den entsprechenden Elementen um optionale Satzteile handelt. Beispiele für Konstruktionen, die hierdurch beschrieben werden, sind

- (1) *(das) Bibliothekswesen*
- (2) *Kants Werke*
- (3) *modernes Bibliothekswesen*
- (4) *virtuelle Bibliotheken im Internet*
- (5) *das Gute in der Philosophie (der Neuzeit)*

Neben der deutschen Grammatik befindet sich auch eine englische Nominalphrasengrammatik im Parser. Dadurch, daß diese Grammatiken kompatibel gehalten sind, können nicht nur englische Eingaben unter der englischen Oberfläche verarbeitet werden, sondern auch englische Ausdrücke in deutschen Konstruktionen:

- (6) *Global players an der deutschen Börse*

Das angegebene Beispiel zeigt, daß auch für aus dem Englischen stammende Fach- oder Modewörter eine syntaktische Analyse notwendig ist („Small and Medium Sized Enterprises“) und nicht einfach nur ein Verweis auf ein um englische Begriffe angereichertes Lexikon ausreicht.

Der Parser wird in bestimmten Fällen durch Annotationen der Ergebnisse des Lexikonzugriffs gesteuert. Ist eine Eingabe vom Lexikon als Länder- oder Personennamen (zumindest in einer Analysevariante) erkannt worden, so ist dies im Analyseergebnis vermerkt und bewirkt, daß der Parser besondere Regeln, die nur für Namen in Frage kommen, anwendet. So ist es beispielsweise nur dann möglich, daß eine im Genitiv auftretende Nominalphrase keinen Artikel vorangestellt bekommt, wenn der Kopf der Nominalphrase ein Name ist:

- (7) * Bürgermeisters Erfolg
- (8) des Bürgermeisters Erfolg
- (9) Kants Werke
- (10) Europas Eroberung

3.2.5 Morphologie

Es ist notwendig, Eingaben des Benutzers wie *Datenbankinformation* und *Datenbankinformationen* oder auch *Datenbank* und *Datenbanken* nicht als voneinander unabhängig (d.h. insbesondere mit einem eigenen Lexikoneintrag) zu betrachten. Um die einzelnen morphologisch markierten Formen eines Wortes zueinander in Beziehung zu setzen, ist regelbasiertes Wissen über z.B. Flexion nötig. Mit Hilfe einer Morphologiekomponente könnten auch Eingaben wie *Marktwirtschaft und China* und *chinesische Marktwirtschaft* aufeinander bezogen werden.

OSIRIS nutzt zu diesem Zweck einerseits in das Lexikon integrierte morphologische Informationen aus Korpora¹⁸ und andererseits eigenständige, kommerzielle Morphologieprogramme der Firma Lingsoft, Helsinki (Haapalainen und Majorin, 1995). Die kommerziellen Programme sind GERTWOL für die morphologische Behandlung des Deutschen und ENGTWOL für die des Englischen. Da diese Komponenten meist mehr als eine Analysemöglichkeit für eine Eingabe finden, ist eine nachträgliche Bewertung und Sortierung der Ergebnisse durch den Task Manager unerlässlich (vgl. 3.2.10).

Aber auch wenn die morphologische Analyse keinen Erfolg hat, stehen immer noch zwei Informationsquellen für die Analyse des Wortes zur Verfügung. Die erste Informationsquelle zu einem unbekanntem Wort stellt ein möglicherweise vom Wort abtrennbares Suffix dar, aus dessen Lexikoneintrag Rückschlüsse auf die morphologische Beschaffenheit des gesamten Ausdrucks gezogen werden können. Auch eine exotische Spezialwissenschaft ist hier, wenn es sich um eine *XYZlogie* handelt, als ein Nomen zu identifizieren, weil bei der Suche nach einem Lexikoneintrag für *XYZlogie* ein Eintrag für *logie* gefunden wird. Die zweite Informationsquelle stellt der Kontext des unbekanntem Wortes dar: Sucht der Benutzer Literatur zum Thema *die mittelalterliche XYZlogie*, dann ist durch den aus Artikel und Adjektiv bestehenden linken

¹⁸Hier ist insbesondere die Zusammenarbeit mit dem Projekt *Computerbasierte Modellierung orthographischer Prozesse (CMP)* im Institut für Semantische Informationsverarbeitung zu nennen.

Kontext sowie dem nur aus dem leeren String bestehenden rechten Kontext festgelegt, daß es sich bei *XYZlogie* um ein Nomen handelt, damit eine syntaktisch wohlgeformte Phrase entstehen kann. Diese Aufgabe wird vom Task Manager bei der Sortierung und Bewertung erledigt (vgl. 3.2.10).

3.2.6 Kompositazerlegung

Eingaben wie *Pädagogikstudium* können angesichts der im Deutschen ungeheuer produktiven Kompositabildung nicht konsequent als eigenständige Lexikoneinträge behandelt werden, sondern sollten zurückgeführt werden auf *Studium der Pädagogik* oder *Studieren von Pädagogik*.

Die Analyse von Komposita wirft im allgemeinen Probleme auf, die im Grenzbereich von Morphosyntax und Semantik anzusiedeln sind. Die Beziehung, in der der linke Teil zum rechten Teil steht, ist in einigen Fällen in Abhängigkeit von der Struktur des rechten Teils zu bestimmen. Im allgemeinen Fall kann aber nicht mehr über die Beziehung zwischen linkem Teil und rechtem Teil gesagt werden, als daß das Kompositum eine Spezifizierung des durch den rechten Teil ausgedrückten Konzepts ist.

Ein Beispiel möge das Kompositum „Steinlaus“ sein. Eine „Steinlaus“ kann z.B. eine Laus aus Stein sein, eine Laus, die auf Steinen lebt, oder die Steinen äußerlich ähnelt. Vielleicht hieß ihr Entdecker aber auch Stein, oder es gibt einen alten Volksglauben, der besagt, daß dieses Tier, wenn man es auf Steinmauern findet, dem Bauwerk ein langes Leben verspricht. Isoliert betrachtet, ohne zusätzliches Wissen, ist die Beziehung zwischen den beiden Komponenten „Stein“ und „Laus“ nicht näher zu spezifizieren.

Für eine „Mädchenschule“ gilt, daß es sich um eine spezielle Art von Schule handelt. Dies gilt genauso für „Handelsschule“, „Oberschule“, „Sprachenschule“, „Sommer-schule“, „Schwimmschule“ oder „Puppenschule“. Mehr als dies kann im allgemeinen Fall aber kaum gesagt werden, ohne daß Arbeit in eine vollständige Ontologie investiert wird, um über Inferenzprozesse zu erschließen, daß „Mädchen“ eine Schule besuchen können, wohingegen „Sprachen“ etwas sind, was man in einer Schule lernen kann. Damit wäre aber immer noch nicht das Problem gelöst, den Unterschied zwischen „Mädchenschule“ und „Puppenschule“ zu erkennen oder zu erklären, daß eine „Schwimmschule“ eine andere Art Schule ist, als eine „Handelsschule“.

Auch vermeintliche Genitive als Indiz für eine Teil-Ganzes Beziehung zwischen den einzelnen Komponenten erweisen sich als irreführend. Eine „Kalbsleber“ ist die Leber eines Kalbes, aber eine „Kalbsleberwurst“ ist keine Wurst aus Kalbsleber, sondern eine aus Schweineleber mit einem geringen Anteil an Kalbfleisch.

Für zusätzliche Probleme sorgen phonologische Phänomene an der Schnittstelle zwischen beiden Komponenten. So ist die „Hausfront“ die Front eines Hauses, wohingegen die „Häuserfront“ die Front mehrerer Häuser ist. Ein „Katzenhaar“ ist allerdings niemals das Haar mehrerer Katzen. Die vermeintliche Pluralendung *-en* erscheint in „Katzenhaar“ nur deshalb, weil das eigentlich stringent aus „Katze“ und „Haar“ gebildete „Katzehaar“ im Deutschen phonologisch markiert ist (siehe auch Gallmann

(1999)).

Die Strategie in OSIRIS zur Behandlung der Semantik von Komposita stützt sich auf das einzige greifbare Faktum: Der größtmögliche im Lexikon vorhandene rechte Rand eines Wortes, als das meist spezifische, darin enthaltene Konzept, ist das geeignete Objekt für die Suche in der Wissensbasis. Die Art der Beziehung zu den weiter links stehenden Teilen, ob ebenfalls bekannt oder auch nicht, ist nicht weiter einschränkbar, weil sie nach dem Stand der Kunst auch mit aufwendigen Mitteln formal syntaktisch und semantisch nicht zu greifen ist. Das heißt, daß „Kalbsleberwurst“ in jedem Fall eine Wurst bezeichnet; ihre genaue Beziehung zu „Kalbsleber“, „Kalb“ oder „Leber“ bleibt aber offen.

Die Zerlegung der Komposita ist in den Lexikonzugriff eingebettet. Das Ergebnis der Suche nach „Jugendarbeit“ ist zunächst „Jugend“ und „Arbeit“, sofern die Bestandteile dem Lexikon bekannt sind. Sonst wird ggf. in einem zweiten Schritt versucht, zumindest den rechten Rand zu identifizieren.

3.2.7 Fehlerkorrektur

Aufgrund des stetig zunehmenden Fachvokabulars und der Fruchtbarkeit der deutschen Kompositabildung ist auch das OSIRIS-Vokabular offen. Dem System unbekannte Wörter müssen deshalb nicht zwangsläufig (Tipp)Fehler sein. Die Entwicklung von Regeln und Heuristiken zur Erkennung von Fehlern ist sehr aufwendig und befindet sich noch im Forschungsstadium. Typische Bedienerfehler wie das Verwechselln zweier benachbarter Tasten („Deduktionssysteme“, vrgl. Abbildung 19, S. 47) sind in eingeschränktem Maße behandelbar. In OSIRIS wird dabei auf Wissen über mögliche Buchstabenkombinationen zurückgegriffen. Enthält ein unbekanntes Wort eine Folge von Buchstaben, die in dieser Form im Deutschen nicht auftritt, kann von einem Fehler ausgegangen werden. Um Fehler erkennen zu können, bedarf es auch bestimmter Annahmen über das Verhalten des Benutzers und die Natur des Fehlers. Dieses sogenannte Benutzermodell beschreibt in OSIRIS einen „unsicheren 10-Finger-Schreiber“. Dabei werden Tastengruppen auf der Tastatur angenommen, wie sie von je einem Finger bedient werden können. Vor diesem Hintergrund werden mittels Heuristiken beim Lexikonzugriff folgende mögliche Fehler erkannt:

- In einem Wort sind zwei benachbarte Buchstaben vertauscht, weil ein Finger schneller als der andere war.
- In einem Wort ist ein Buchstabe zuviel vorhanden, weil ein Finger nicht nur die gewünschte Taste, sondern auch eine benachbarte Taste, die vom gleichen Finger bedient wird, traf.
- In einem Wort ist ein Buchstabe falsch, weil seine Taste von einem Finger getroffen wurde, obwohl die benachbarte Taste, die vom gleichen Finger bedient wird, gemeint war.

In OSIRIS wird in unbekanntenen Wörtern maximal ein Tippfehler erkannt, da die Fehlererkennung ein sehr rechenaufwendiger Prozeß ist. Zudem ist sie sprachabhängig und zur Zeit in OSIRIS auf das hin Deutsche optimiert. Die Fehlererkennung befindet sich noch im Prototypenstadium.

3.2.8 Semantik

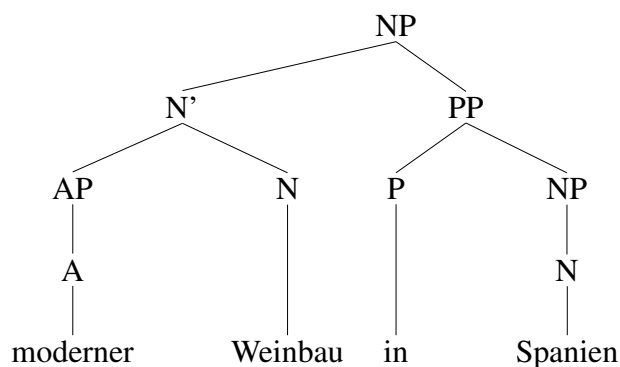
Zur Annäherung an die notorisch schwierige Semantik von Präpositionen wird ein sehr einfaches Modell verwendet. Wenn Präpositionen mit einem zeitlichen oder räumlichen Aspekt annotiert sind, „restringieren“ sie das vom Nomen vertretene zentrale Konzept in Raum oder Zeit, andernfalls „modifizieren“ sie es. D.h. daß für die Präpositionalphrasen eine unterschiedlich enge Bindung an den Kopf der Nominalphrase angenommen wird.

$$(11) NP \rightarrow N' [PP]^*$$

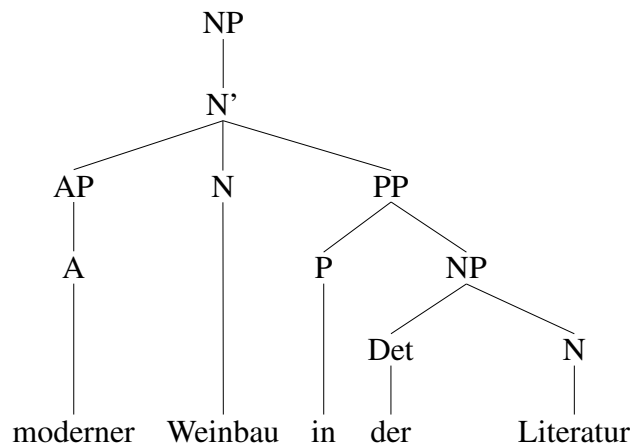
$$(12) N' \rightarrow [DET] [AP] N [NP] [PP]^*$$

Je nachdem, ob die Präposition *in* mit einem zeitlichen oder räumlichen Aspekt annotiert ist oder nicht, trifft (11) beziehungsweise (12) zu, d.h. „restringiert“ *in* das vom Nomen vertretene zentrale Konzept in Raum oder Zeit oder „modifiziert“ es. Damit wird die Idee umgesetzt, daß eine räumliche Spezifikation mit Hilfe einer Präpositionalphrase wie „in Spanien“ eine eigenständige Dimension in die Suche einführt, die entsprechend in der Analyse weiter entfernt von der NP repräsentiert werden muß. Hingegen wird eine Präpositionalphrase wie „in der Literatur“ als eine direkte Modifikation des Kopfes der NP angesehen, was durch eine enge syntaktische Bindung ausgedrückt werden soll.

Eine Benutzereingabe wie *moderner Weinbau in Spanien* führt daher entsprechend (11) zu folgender syntaktischen Analyse:



Eine Eingabe wie *moderner Weinbau in der Literatur* wird hingegen wie folgt analysiert:



Natürlich kann aus der Präposition allein nicht darauf geschlossen werden, ob sie eine PP einleitet, die einen lokalen oder einen temporalen Aspekt ausdrückt. Darüber entscheidet das in der PP auftretende Nomen, für das im Lexikon ein entsprechender Eintrag vorhanden sein muß. Dies ist aber auch der Grund, weshalb die unterschiedliche syntaktische Repräsentation zur Zeit in OSIRIS zwar angelegt, aber nicht konsequent genutzt wird. Um diese Information für die nachfolgende Suche nutzen zu können, müssen die Informationen zum Nomen flächendeckend im Lexikon vorhanden sein. Dies ist momentan aber nicht der Fall.

Solange die Voraussetzungen für eine differenziertere Behandlung von Präpositionalphrasen nicht gegeben sind, werden Präpositionen einfach generell als eine unbestimmte Modifikation betrachtet. Anfragen wie „Datenverarbeitung in Bibliotheken“ werden daher semantisch genauso interpretiert wie „Datenverarbeitung in der Bibliothek“, nämlich als „Datenverarbeitung modifiziert-durch Bibliothek“. Im Vergleich mit dem OPAC, der Präpositionen generell mittels einer Stopwortliste ignoriert, ist dies aber immer noch ein beachtlicher Fortschritt, der sich auch an den Retrievalergebnissen ablesen läßt (vgl. 3.3).

3.2.9 Phonetik

Da die Beziehung zwischen Zeichen und Laut mehrdeutig ist, d.h. orthographisch verschiedenen Zeichenketten („Meyer“, „Maier“) dieselbe phonetische Repräsentation zugeordnet werden kann, gestaltet sich die Suche nach Personennamen oftmals schwierig (Hitzenberger, 1986). Neben der Variabilität der Schreibweisen, bei denen sich alte orthographische Standards mit neuen mischen („Thaler“ vs. „Taler“), ist auch die Unsicherheit der Benutzer über die korrekte Schreibung komplexer Namen („Meyer zu Broxten“) ein Nutzungshindernis.

In OSIRIS sind Personennamen phonetisch kodiert und nach Namensbestandteilen sortiert abgelegt. Der Zugriff auf Personennamen ist somit weder an die lineare Ordnung der Eingabe noch (in gewissen Grenzen) an die korrekte Schreibweise gebunden. Auch

ist es möglich, mit Teilen des Namens, z.B. bei Doppelnamen, erfolgreich zu suchen: die Suche nach „Westfalen“ trifft auch „Westphalen-Dörenberg“.

Bei der Erstellung der phonetischen Zwischenrepräsentation werden Buchstaben oder Kombinationen von Buchstaben aufgrund ihrer phonetischen Eigenschaften substituiert. Beispielsweise werden Labiodentale (*f*, *ph*, *v* und *w*) und Bilabiale (*b* und *p*) durch ein gemeinsames Zeichen (F bzw. B) ersetzt und Doppelkonsonanten werden zu einem Zeichen reduziert. So wird aus der Eingabe *PHILLIP* in einem ersten Schritt *FILIB*. In einem zweiten Schritt werden die Vokale entfernt, so daß das konsonantische Gerüst *FLB* als phonetischer Code übrig bleibt. Der Verzicht auf die Vokale liegt in der größeren Bedeutung der Konsonanten als distinktive Merkmale für Eigennamen begründet (Postel, 1969; Hitzenberger, 1986).

Von dieser Idee ausgehend wurden verschiedene Reduktionsverfahren entwickelt, um zu phonetischen Repräsentationen zu gelangen, die einerseits sprachspezifische Ähnlichkeiten adäquat abbilden und andererseits die resultierenden Ähnlichkeitsklassen in einer für das Retrieval förderlichen Größenordnung halten.

Das älteste Verfahren zu diesem Zweck ist der Russel-Soundex-Code, der bereits Anfang des 20. Jahrhunderts in den USA für die Personenerfassung im Rahmen der Volkszählungen entwickelt wurde. Soundex wird in OSIRIS zur Behandlung der englischen Personennamen eingesetzt. Dieser Algorithmus reduziert Namen auf vier Zeichen, wobei das erste Zeichen des Soundex-Codes der erste Buchstabe des ursprünglichen Namens ist. Die drei folgenden Zeichen entstehen, indem:

1. Vokale und Umlaute entfernt werden.
2. Doppelkonsonanten zu einfachem Auftreten reduziert werden.
3. die verbliebenen Zeichen je nach Zugehörigkeit zu einer aus insgesamt sechs verschiedenen Äquivalenzklassen in eine diese Klasse repräsentierende Ziffer (1 bis 6) transformiert werden. Dabei werden *w*, *y* und *h* ignoriert, d.h. nicht transformiert.
4. doppelt und benachbart auftretende Ziffern zu einem Vorkommen reduziert werden.
5. die ursprünglich ignorierten (aber im vorangegangenen Schritt wichtigen) Zeichen *w*, *y* und *h* eliminiert werden.
6. vom so entstandenen Code nur die ersten drei Ziffern berücksichtigt werden. Im Fall, daß der entstandene Code weniger als drei Ziffern enthält, wird mit der 0 aufgefüllt.

Mit Hilfe dieses Algorithmus werden die Eingaben „Smith“ und „Smythe“ durch denselben Soundex Code (S530) repräsentiert. Allerdings werden auch „Hilbert“ und „Heilbronn“ identisch repräsentiert (H416).

Soundex ist sehr weit verbreitet (u.a. liegt es großen Datenbankmanagementsystemen wie Oracle und DBase bei, ist aber auch Bestandteil des Emacs) und für nordamerikanische Namen optimiert. Dies zeigt sich beispielsweise in der starken Betonung des Anlautes sowie im Zusammenfall von *b* und *v* (aufgrund der hohen Anzahl spanischer Namen). Eine weitere Einschränkung für den Einsatz außerhalb des anglo-amerikanischen Sprachraumes ist neben der fehlenden Behandlung von Umlauten die Kontextfreiheit: Zeichen werden grundsätzlich kontextfrei transformiert, lediglich bei der Löschung wird der Kontext berücksichtigt, d.h. Digraphen werden nicht berücksichtigt.

Als Überarbeitung des Soundex-Algorithmus und als speziell auf das Deutsche hin zugeschnittenes Reduktionsverfahren wurde die „Kölner Phonetik“ (Postel, 1969) entwickelt. Ausgehend von einer Untersuchung von über 1 Million Personennamen wurde dabei die durchschnittliche Länge eines Namens mit 9 Buchstaben bestimmt, wobei die Initialen zu 80 % Konsonanten waren und in den anderen Teilen des Namens der Anteil der Konsonanten bei 63 % lag. Wie auch Soundex setzt die Kölner Phonetik also auf die besondere Aussagekraft von Konsonanten zur Beschreibung von Eigennamen und ignoriert Vokale bis auf solche am Wortanfang. Dort werden Vokale berücksichtigt, weil Zeichen am Anfang des Wortes generell als aussagekräftiger angesehen werden als im Rest des Wortes. Konkret werden so 26 Buchstaben kontextsensitiv in 11 Klassen eingeteilt und durch 9 verschiedene Ziffern substituiert. So werden *d* und *t* beispielsweise durch 8 substituiert, wenn *c*, *s* oder *z* folgen, andernfalls durch 2.

Weitere Besonderheiten der Kölner Phonetik betreffen phonetische Sonderbedingungen in Silben und Buchstabenkombinationen. So werden durch eine Vorverarbeitung Kombinationen wie *own* in *Brown* auf *aun* abgebildet und *wsk* wie in *Orlowski* auf *rsk*. Die Untersuchung der Eigennamen ergab ferner, daß ca. 5 % der Namen aus mehr als einer Komponente bestehen. Um mit diesen Fällen umgehen zu können, werden z.B. Doppelnamen sowohl in der kombinierten Version *Leutheuser-Schnarrenberger* als auch in ihre Bestandteile zerlegt verarbeitet.

Die Kölner Phonetik wurde in einigen wesentlichen Punkten für den Einsatz in Namensdatenbanken wie z.B. der Verwaltung eines Krankenhauses durch PD Dr. Martin Haase und Dr. Kai Heitmann (Haase und Heitmann, 2000) überarbeitet. Auf diese „Erweiterte Kölner Phonetik“ stützt sich die Implementation, die zur phonetischen Analyse von Eigennamen in OSIRIS verwendet wird.

Dabei kommt der phonetischen Analyse entgegen, daß in OSIRIS die einzelnen Namensbestandteile wie Vorname, Namenszusätze, Adelstitel und Nachnamen getrennt abgelegt sind. Außerdem existieren diese Bestände jeweils in der originalen Form sowie bereits phonetisch reduziert. In OSIRIS findet also zunächst auf funktional unterschiedenen Indizes eine Suche nach Namensteilen in der Form der Eingabe statt, und erst, wenn dies scheitert, werden die phonetischen Codes herangezogen. Dieser Teil der computerlinguistischen Funktionalität ist aus technischen Gründen in die Datenbank ausgelagert. Prinzipiell wäre es natürlich möglich, wie z.B. bei der Komposita-zerlegung die Analyse in den Lexikonzugriff zu integrieren.

3.2.10 Bewertung und Ergebnisübergabe

Das zuvor beschriebene Ensemble von Komponenten dient also der Analyse natürlicher Benutzereingaben. Ziel dieser Analyse ist es, den vom Benutzer formulierten Themenkomplex für die Literatursuche zu bestimmen. Aufgabe der computerlinguistischen Komponenten ist es, dieses Thema aus der Eingabe des Benutzers möglichst exakt zu bestimmen und durch zusätzliches Wortmaterial (query expansion) der nachfolgenden Suche auf der Wissensbasis eine möglichst günstige Ausgangssituation zu bieten. Bei der Suche auf der Wissensbasis geht es dann darum, das Thema des Benutzers auf die Terminologie der Klassifikation der Bibliothek abzubilden.

Haben das Lexikon und der Parser ihre Analyse beendet, muß der Task Manager zunächst die Ergebnisse dieser Komponenten bewerten. Im Regelfall wird es so sein, daß z.B. das Lexikon für ein Wort mehrere verschiedene Analysen liefert. So würde für das erwähnte Beispiel „Schliche“ (wenn das Wort denn im Lexikon verzeichnet ist) sowohl die Analyse *Nomen (Plural, kein Singular möglich)* als auch *Adjektiv (unbekannter Wortstamm „Sch-“)* geliefert. Für eine Eingabe wie „Kampfsport“ liefert das Lexikon alternative Zerlegungen in *Kampf-sport* und *Kampf-s-port*. Der Task Manager bewertet diese Analyseergebnisse aufgrund ihrer Komplexität. Generell wird die einfachere Analyse der komplizierteren vorgezogen, d.h. daß die Zerlegung in *Kampf-sport* präferiert wird, weil sie ohne die Annahme eines Fugenmorphems wie in *Kampf-s-port* auskommt. Für Komposita wird generell die Lösung bevorzugt, die mit möglichst wenig Zerlegungen und innerhalb der einzelnen Bestandteile mit möglichst wenig Analyseaufwand (Derivation, Genitive, Flexion) auskommt. Auch die Analyse von „Schliche“ als Nomen wird gegenüber der als Adjektiv mit unbekanntem Wortstamm „Sch-“ präferiert. Auch bei einfachen Wörtern wie „Haus“ werden mehrere Analysemöglichkeiten geliefert: dabei wird *Haus* als Nomen im Singular (das Gebäude) der Analyse als Genitiv von *Hau* vorgezogen.

Andere Bewertungsschritte betreffen Analysen, die mit dem Zusatz „Hypothese“ gekennzeichnet sind. Solche Analysen entstehen aufgrund von Suffixanalysen (siehe 3.2.3) bei unbekanntem Wortanfang. Auch als Eigennamen erkannte (d.h. im speziellen Namenlexikon gefundene) Nomen werden gesondert bewertet, vor allen Dingen wird eine Zerlegung für solche Nomen ausgeschlossen. So kann verhindert werden, daß Eigennamen die ursprünglich einmal Komposita waren („Beckenbauer“) zerlegt werden.

Ist dieser Bewertungsprozeß abgeschlossen, so wird die beste Lösung in einer Attribut-Wert-Struktur, wie sie in Abbildung 30 für die Benutzeranfrage *ernährungsphysiologische Bewertung der Lebensmittel aus Biolandbetrieben* dargestellt ist, der nachfolgenden Suche auf der OSIRIS-Wissensbasis übergeben (die unveränderte Benutzereingabe findet sich immer vorweg in der ersten Zeile):

ernährungsphysiologische Bewertung der Lebensmittel aus Biolandbetrieben

§T: lebensmittel
 §T1: mittel
 §TA1: leben
 modified_by
 §T§M§T: biolandbetrieb
 §T§M§T1: landbetrieb
 §T§M§T2: betrieb
 §T§M§TA1: bio
 §T§M§TA2: bioland
 modified_by
 §T§M§TS: ernährungsphysiologische bewertung
 §T§M§T: bewertung
 modified_by
 §T§M§T§M: ernährungsphysiologisch

Abbildung 30: Struktur des Outputs der computerlinguistischen Komponenten

In dieser Containerstruktur ist der thematische Aufbau der Benutzereingabe kodiert. Nach der originalen Benutzeranfrage folgt das von der Analyse erkannte Thema der Anfrage, in diesem Falle *Lebensmittel*. Dabei handelt es sich um ein Kompositum, dessen Bestandteile *Mittel* und *Leben* mit angegeben werden. Modifiziert wird dieses Thema durch *Biolandbetrieb*, wiederum ein Kompositum, dessen einzelne Bestandteile mit angegeben werden. Die gesamte bislang repräsentierte Struktur, nämlich das Thema *Lebensmittel* mit seiner Modifikation *Biolandbetrieb* wird nun wiederum als in einer Modifikationsrelation stehend erkannt, nämlich zu einem sogenannten Superthema *ernährungsphysiologische Bewertung*. Zusammen mit dem Superthema werden dessen Bestandteile *ernährungsphysiologisch* und *Bewertung* geliefert.

Im Falle eines Mißerfolges der computerlinguistischen Analyse wird die Benutzereingabe unverändert den nachfolgenden Retrievalkomponenten übergeben.

Für die nachfolgende Suche auf der Wissensbasis werden unter Berücksichtigung von Teilergebnissen der Suche weitere Verfahren zur query expansion wie z.B. Kompositagenerierung angewandt. Führt eine Suche mit den Begriffen „Sport“ und „Alter“ nicht zum gewünschten Erfolg, dann wird mit dem online gebildeten Kompositum „Alterssport“ weiter gesucht. Daß hier auch unbrauchbare Komposita wie „Sportalter“ gebildet werden, ist unerheblich, weil die Suche mit einem solchen Begriff in der Datenbank eine leere Ergebnismenge hat und Suchvorgänge ohne Ergebnis keine nennenswerte Rechnerbelastung darstellen. Das Ergebnis der Suche wird in Abhängigkeit von der Größe der Ergebnismenge bewertet und ggfs. die Darstellung einzelner Klassen unterdrückt (Recker et al., 1996). Entspricht die Eingabe einem Buchtitel, so wird dieser Titel auch dann präsentiert, wenn die entsprechende Klasse insgesamt nicht als zur Anfrage passend betrachtet wird.

Spätestens an diesem Punkt wird die Relevanz eines guten Parsers im Vergleich mit herkömmlichen Recherchesystemen noch einmal deutlich: Würden Benutzereingaben generell nur als Stichwortreihung verstanden, würde also die syntaktische Struktur der natürlich-sprachlichen Eingabe ignoriert, dann könnte die vom Benutzer mit Ausdrücken wie *zur Zeit der Stauffer, unter Friedrich dem Großen, im Nordosten Frankreichs, mit den Mitteln der Spektralanalyse, für das Lehramt* usw. explizit gemachte Modifikation eines Ausdrucks nicht als solche erkannt werden, sondern müßte als eigenständiges Stichwort bei der Suche in der Klassifikation behandelt werden.

Andererseits erscheint die Aufgabe des Parsers im Vergleich zu den anderen computerlinguistischen Komponenten vor dem Hintergrund der ausführlichen Vorverarbeitung und der heuristischen Bewertung weniger zentral. Der Parser hat lediglich aus den bereits für die einzelnen Inputteile vorliegenden Informationen eine insgesamt wohlgeformte Analyse zu erstellen.

3.3 Verbesserung der Recherche durch OSIRIS

Die deutlichen Verbesserungen der Recherchemöglichkeiten durch OSIRIS sollen anhand einiger Praxisbeispiele veranschaulicht werden.

OSIRIS liefert für die bereits erwähnte Themenanfrage *Sport im Alter* die Klasse *Senioren-sport* mit 49 Titeln, sowie die Klassen *Sport und Heilpädagogik* und *Motorisches Lernen*, deren Bestand durchweg relevant für die Frage ist. Im OPAC trifft die Frage nach *Sport im Alter* gerade 5 Titel.

Das ebenfalls bereits angeführte Beispiel *Datenverarbeitung in Bibliotheken* führt im OPAC auf 5 Titel, OSIRIS präsentiert verschiedene Unterklassen von *Buch- und Bibliothekswesen*, darunter auch *Automatisierte Datenerfassung und Datenverarbeitung in der Bibliothek*. Ein Klick auf diese Klasse liefert bereits 38 Titel, die aufgrund der Klassenzugehörigkeit natürlich alle relevant sind.

Eine Anfrage wie *Chemie* wird am OPAC nicht beantwortet: Die Zahl der Treffer liegt im Bereich von mehreren Tausend, und solche Treffersets werden nicht mehr dargestellt. OSIRIS liefert für die Frage nach *Chemie* genau eine Antwort: Ein Symbol für die Klasse *Chemie*. Klickt der Benutzer auf das die Klasse repräsentierende Symbol, so findet er natürlich nicht alle Titel der chemischen Literatur vor, sondern die Unterklassen des Faches Chemie, *Organische Chemie*, *Anorganische Chemie* usw.

Sucht ein Benutzer nach *machine learning with neural networks* so findet der OPAC nur ein Buch, in dessen Titel genau die gesuchten Stichworte vorkommen:

- (13) Hojjat Adeli; Shih-Lin Hung: Machine learning : neural networks, genetic algorithms, and fuzzy systems

Alle anderen Bücher zu diesem Thema, insbesondere die deutschsprachigen, werden nicht gefunden. Das ist in der Praxis durchaus ein Problem, denn die deutschen Entsprechungen zu Fachtermini wie bspw. *loop space* oder *string theory* sind ungebräuchlich und deshalb dem Suchenden meist unbekannt (vgl. S. 39). OSIRIS liefert auf die oben genannte Anfrage neben besagtem Werk, dessen Titel mit der Suchanfrage übereinstimmt, weitere Klassen aus Informatik, Künstlicher Intelligenz und Wirtschaftswissenschaften (Statistik). Diese Anfrage wird in OSIRIS nicht mit Hilfe einer Übersetzung bearbeitet, sondern mit den englischsprachigen Schlagwörtern, die in der OSIRIS-Wissensbasis vorhanden sind (vgl. 3.1).

Vielen Benutzern ist zunächst unklar, daß ihre Suchbegriffe in anderen Disziplinen mit ganz anderer Bedeutung verwendet werden können (*Morphologie*, *Gruppentheorie* ...). Sucht ein Benutzer im OPAC nach *Morphologie*, so findet er in Osnabrück 289 Titel. Anhand der unstrukturierten Titelliste, die der OPAC als Ergebnis präsentiert, ist es für den Benutzer schwer zu erkennen, aus welchen Bereichen (Medizin, Biologie, Linguistik usw.) die Treffer stammen. Versucht er dann die hinsichtlich seiner Interessen inhaltlich relevante Teilmenge zu identifizieren, muß er jeden einzelnen Titel betrachten. Dieselbe Anfrage an OSIRIS liefert als Ergebnis die entsprechenden einschlägigen Unterklassen, u.a. für die *Computerlinguistik* die Klasse *Lexikalisches*

Wissen, Morphologie, Phonetik, für die *Biologie* in der *Zoologie* und in der *Botanik* die entsprechenden Klassen, für die Einzelsprachen die jeweiligen Unterklassen *Morphologie* usw. usf.

Sucht ein Bibliotheksbenutzer nach Literatur zur Programmiersprache *C*, so wird er im OPAC nicht fündig: *C* ist ein sog. Stopwort und wird von vornherein nicht zur Suche zugelassen. Die Suche in OSIRIS führt u.a. in der *Informatik* in die Unterklasse *Problemorientierte Programmiersprachen*.

Eine Suche nach Literatur zur Stadt *Rom* endet im OPAC mit einer Liste von 470 Titeln. Da die Suche auf pattern matching basiert, haben etwa 20% der gefundenen Titel nichts mit *Rom* zu tun: Ihnen liegt eine *CD-Rom* bei oder der Begriff wird im Titel genannt. OSIRIS liefert eine Klasse *Rom* als Ergebnis.

Sucht ein Bibliotheksbenutzer nach dem Namen eines Autors, so führt bereits die bibliothekarische Eingabekonvention *Nachname, Vorname* am OPAC zu Problemen. Autorennamen wie *Isabella Ebers*, *Tillmann Westfalen* oder *Roland Schwenzel* können weder in der dargestellten natürlichen Form von *Vorname Nachname*, noch aufgrund ihrer Schreibweise am OPAC gefunden werden. Die Suche in OSIRIS führt (egal in welcher Form die Namensbestandteile angegeben werden) unter Verwendung der phonetischen Komponenten (vgl. 3.2.9) zu folgenden Ergebnissen:

(14) *Isabella Schneider-Eberz* für die Eingabe *Isabella Ebers*

(15) *Tilman Westfalen* für die Eingabe *Tillmann Westfalen*

(16) *Roland Schwänzl* für die Eingabe *Roland Schwenzel*

Neben der Themensuche und der Autorensuche können in OSIRIS natürlich auch bestimmte Werke gesucht werden. Im Gegensatz zum OPAC reicht es hier allerdings, wenn man seine Anfrage wie in einer Buchhandlung notiert: *Shoham: Reasoning about Change* findet das entsprechende Werk, *Chomsky Syntactic Structures* findet neben dem englischen Original auch die deutsche Übersetzung.

3.3.1 Die Retrievalmöglichkeiten im direkten Vergleich zum OPAC

Die Effizienzsteigerung, die das OSIRIS-System im direkten Vergleich zum Osnabrücker OPAC möglich macht, ist auch auf der Grundlage unabhängiger Beispielfragen untersucht worden (Ronthaler und Zillmann, 1998). Dazu wurden für die thematische Recherche Anfragen eines Retrievaltests des MILOS-Systems (Lepsky et al. (1996), vgl. auch 3.4.2) ergänzt um weitere, eigene Anfragen verwendet.

Die Ergebnisse dieses Vergleichs werden nachfolgend in Form einer Tabelle dargestellt, zu deren Verständnis es wichtig ist zu wissen, daß die für die Suche in OSIRIS angeführten Phrasen ausschließlich in der wiedergegebenen Form (also ohne Trunkierung etc.) verwendet wurden – in OSIRIS fand also nur eine einzige Anfrage statt. Hingegen bringt das Superskript *k* bei den OPAC-Trefferzahlen zum Ausdruck, daß

hinter diesen sich mehrere Anfragen verbergen, mit denen auch Komposita bzw. deren Zerlegung, Trunkierungsmöglichkeiten sowie Synonyme abgefragt wurden. Dabei wurde die vom OPAC ermittelte Trefferzahl direkt übernommen, ohne daß auf Rechercheballast geprüft worden wäre. Das Beispiel *Java* zeigt jedoch die offenkundige Existenz solchen Ballastes in manchen Anfragen.

Gesucht wurde auf dem gesamten maschinenlesbaren katalogisierten Bestand der Universitätsbibliothek Osnabrück, das sind 600.000 Titeldatensätze.

Die Ergebnisse der Recherche mit OSIRIS sind zum einen als Anzahl der jeweils relevanten Dokumente in den getroffenen Klassen angegeben, zum anderen ist der Ausschnitt der Bibliothekssystematik, den das OSIRIS-System für eine Suchanfrage findet, in ihrer hierarchischen Struktur wiedergegeben:

Suchanfrage	Trefferzahl wie vom OPAC angegeben	Relevante Treffer in OSIRIS	• Oberklasse ▷ Unterklasse
-------------	------------------------------------	-----------------------------	-------------------------------

Die Frage, was denn einen relevanten Treffer im Ergebnis des OSIRIS-Systems ausmacht, wird durch die Zugehörigkeit des Titels zu einer relevanten Klasse der Systematik entschieden. Die Relevanz der gefundenen Klassen für die jeweilige Suchanfrage wurde durch den stellvertretenden Direktor der Universitätsbibliothek Osnabrück, Dr. Hartmut Zillmann beurteilt (vgl. Ronthaler und Zillmann (1998)). Die Zugehörigkeit eines Titels zu einer Klasse wird während des Katalogisierungsvorgangs vom Fachreferenten bestimmt (siehe auch 2.2) – dessen Arbeit wurde nicht in Frage gestellt.

Suchanfrage	OPAC	OSIRIS	OSIRIS-Trefferbild
psychologische Diagnostik	22	333	<ul style="list-style-type: none"> • Psychologische Diagnostik <ul style="list-style-type: none"> ▷ Allgemeines. Gesamtdarstellungen ▷ Testkompendien ▷ Testtheorie. Testkonstruktion ▷ Intelligenzdiagnostik. <li style="padding-left: 20px;">Begabungsdiagnostik ▷ Entwicklungsdiagnostik. ▷ Leistungsdiagnostik ▷ Exploration. Anamnese ▷ Persönlichkeitsdiagnostik ▷ Sonstiges. Einzelfragen

Suchanfrage	OPAC	OSIRIS	OSIRIS-Trefferbild
Sport im Alter	18 ^k	50	<ul style="list-style-type: none"> • Allgemeine Trainingslehre • Seniorensport • Sportpsychologie <ul style="list-style-type: none"> ▷ Motorisches Lernen • Sportpädagogik <ul style="list-style-type: none"> ▷ Sport- und Heilpädagogik. Behindertensport. Eurhythmie
Erkrankungen der Lunge	2 ^k	32	<ul style="list-style-type: none"> • Innere Medizin <ul style="list-style-type: none"> ▷ Erkrankungen der Atmungsorgane
Datenverarbeitung in Bibliotheken	5	47	<ul style="list-style-type: none"> • Bibliotheksverwaltung <ul style="list-style-type: none"> ▷ Automatisierte Datenerfassung und Datenverarbeitung in der Bibliothek ▷ ... ▷ Organisation und Betrieb. Rationalisierung • Anwendungen der Informatik <ul style="list-style-type: none"> ▷ Anwendungen in ... der Öffentlichen Verwaltung
Gerontologie	47	104	<ul style="list-style-type: none"> • Fachtheorie/Fachpraxis Krankenpflege <ul style="list-style-type: none"> ▷ Gerontologie/ Geriatriische Krankenpflege • Sonstige klinische Fächer <ul style="list-style-type: none"> ▷ Geriatrie. Gerontologie
Theatergeschichte	28	64	<ul style="list-style-type: none"> • Theatergeschichte <ul style="list-style-type: none"> ▷ Allgemeine Darstellungen ▷ Einzelne Epochen ▷ Einzelne Länder ▷ Einzelne Orte und Bühnen
Festkörperphysik	159	1309	<ul style="list-style-type: none"> • Festkörperphysik <ul style="list-style-type: none"> ▷ Grundlagen, Allgemeines ▷ Spezielle Themen ▷ Theoretische Festkörperphysik ▷ Mechanische Festkörpereigenschaften. Phononen ▷ Optische Festkörpereigenschaften. Festkörperspektroskopie ▷ Kristallphysik ▷ ... ▷ Festkörperphysik nach Stoffgruppen ▷ Sonstiges zur Festkörperphysik

Suchanfrage	OPAC	OSIRIS	OSIRIS-Trefferbild
algebraische Geometrie	15	283	<ul style="list-style-type: none"> • Algebra <ul style="list-style-type: none"> ▷ Algebraische Geometrie ▷ ...
algebraic geometry	131	283	<ul style="list-style-type: none"> • Algebra <ul style="list-style-type: none"> ▷ Algebraische Geometrie ▷ ...
Ökologie der Gewässer	3^k	63	<ul style="list-style-type: none"> • Sicherheitstechnik. Umwelttechnik <ul style="list-style-type: none"> ▷ Gewässerverunreinigung. Gewässerschutz • Allgemeine Ökologie <ul style="list-style-type: none"> ▷ Einzeldarstellungen der Ökologie
Chemie	Treffermenge nicht darstellbar	nicht nachgezählt	<ul style="list-style-type: none"> • Chemie <ul style="list-style-type: none"> ▷ ... ▷ Naturstoffchemie
Der Ledenhof in Osnabrück	1	1	1 Buch: „Der Ledenhof in Osnabrück“
Java	58	nicht nachgezählt	<ul style="list-style-type: none"> • Computerlinguistik und Künstliche Intelligenz <ul style="list-style-type: none"> ▷ KI-Programmierung ▷ ... • Programmiersprachen <ul style="list-style-type: none"> ▷ Problemorientierte Programmiersprachen • Regionale Völkerkunde <ul style="list-style-type: none"> ▷ Asien ▷ ...
Homotopy of loop spaces	0	4	<ul style="list-style-type: none"> • Topologie. Mannigfaltigkeiten <ul style="list-style-type: none"> ▷ Algebraische Topologie. Homologie. Homotopie
C*-Algebras	0	19	<ul style="list-style-type: none"> • Funktionalanalysis <ul style="list-style-type: none"> ▷ Topologische Algebren. Banach-Algebren

Die obenstehende Tabelle zeigt, daß das OSIRIS-System im Vergleich zum OPAC der Universitätsbibliothek Osnabrück für thematische Recherchen auf demselben Datenbestand im Durchschnitt 11mal mehr Treffer liefert.

3.4 OSIRIS im Kontext anderer Bibliothekssysteme

Wie in 3.1 und 3.2 erläutert, trägt OSIRIS im wesentlichen mit zwei Aspekten zur Verbesserung der Recheresituation in der Bibliothek bei: Einerseits mit einer robusten, natürlichsprachlichen Endbenutzerschnittstelle zur Datenbank und andererseits mit einer intelligenten, automatischen Aufbereitung des verfügbaren Datenbestandes. In dieser Kombination und mit der Betonung automatischer, regelbasierter Verfahren, hebt sich das OSIRIS-System deutlich von allen dem Autor bekannten Ansätzen zur Rechercheverbesserung ab. Dennoch gibt es Projekte im bibliothekarischen Bereich, denen OSIRIS in verschiedener Hinsicht ähnlich ist und die daher hier vorgestellt werden sollen.

3.4.1 MILOS

Ausgangspunkt des MILOS¹⁹-Projektes an der Universitätsbibliothek Düsseldorf war die Erkenntnis, daß einer stetig steigenden Zahl maschinenlesbarer Titeldatensätze in den Bibliotheksverbänden keine ausreichend geeigneten Verfahren zu ihrer Erschließung gegenüberstehen. Wie bereits in 2.2 ausgeführt, bleiben aufgrund der hohen Kosten der personalintensiven Erschließung nach RSWK große Mengen auch neuer Daten unbearbeitet, was die Recherche in diesen Daten enorm erschwert. Für die Universitätsbibliothek Düsseldorf wird für einen Gesamtbestand von ca. 800.000 Titeln mit einem Anteil von 35 % verbal erschlossener Titel gerechnet. In der Praxis bedeutet das für den Benutzer, daß er mit Titelstichwörtern arbeiten muß, die oftmals nur in flektierter Form oder als Bestandteile von Komposita vorliegen:

„Die Suche über Titelstichwörter ist ohne begleitende Maßnahmen ein untaugliches Mittel zur thematischen Recherche im OPAC. Dies ist zwar seit langem bekannt [...], ändert jedoch nichts an der Tatsache, daß in der täglichen Praxis dennoch ein großer Prozentsatz von sachlichen Suchen über Titelstichwörter erfolgt.“(Lepsky, 1996b)

Die Ausgangslage des MILOS-Projektes ist also in der hier geschilderten Form durchaus vergleichbar mit der des OSIRIS-Projektes.

Allerdings unterscheiden sich die beiden Projekte darin, wie die gemeinsam als unzulänglich erkannte Situation verbessert werden soll. Im Rahmen von MILOS wird einzig die Retrievalfunktionalität des Systems betrachtet, eine Vorverarbeitung der Benutzereingabe oder eine Erweiterung der Eingabemöglichkeiten, wie sie OSIRIS umsetzt, findet nicht statt. Ausgangspunkt der Verbesserungen in MILOS ist einzig der Basic Index, in dem sich die Titelstichwörter und die vorhandenen Schlagwörter befinden. Zu seiner Anreicherung wurde mit Methoden der linguistischen Datenverarbeitung eine wörterbuchbasierte Analyse und Erweiterung des im Index vorhandenen Vokabulars vorgenommen (Lepsky, 1994; Zimmermann, 1996). Dazu wurde in der MILOS I genannten Projektphase 1993/94 das Index-Vokabular durch sprachspezifische

¹⁹Maschinelle Indexierung zur erweiterten Literaturschließung in Online-Publikumskatalogen.

Grundformenermittlung (Deutsch, Englisch, Französisch) und systematische Vollformenintegration vereinheitlicht und zugleich mit Hilfe von Stopwortlisten reduziert. Ein Dekompositionsschritt sorgte für eine weitere Anreicherung des Wortmaterials. Bei Bedarf wurden auch manuelle Verarbeitungsschritte durchgeführt. Erste Versuche einer wortbasierten Übersetzung zeigten eine unerwartete Komplexität und führten zur Ausgliederung eines eigenständigen Projektes (CANAL/LS, siehe 3.4.2), das sich im Rahmen des „Libraries Program“ der EU um die Entwicklung eines stichwortbasierten Übersetzungssystems bemüht.

Zu Beginn des in MILOS durchgeführten Anreicherungsverfahrens steht der Abgleich der Wörterbücher mit dem tatsächlich im Index vorhandenen Material. Je nachdem wie hoch der Anteil des den Wörterbüchern unbekanntes Vokabulars ist, kommt eine manuelle Überprüfung oder Korrektur sowie eine Anreicherung der Wörterbücher in Betracht. Die Entscheidung für oder gegen eine manuelle Behandlung ist dem Bearbeiter des Index freigestellt. Eine Entscheidung gegen eine manuelle Behandlung bewirkt, daß die als unbekannt im Index markierten Wörter im weiteren Verlauf der Behandlung ignoriert aber nicht eliminiert werden: Sie stehen als Stichwörter für die Suche später weiter zur Verfügung. Im nächsten Schritt wird die Grundform ermittelt, von der ausgehend alle im Wörterbuch verzeichneten Vollformen in den Index geschrieben werden. Wird dabei der in Frage stehende Indexeintrag durch das Wörterbuch als Kompositum erkannt, so werden auch seine Bestandteile in den Index geschrieben. Zur Vorgehensweise bei der Dekomposition gibt es leider keine näheren Angaben: Während einerseits versichert wird, „Komposita werden wörterbuchgestützt in ihre sinnvollen Wortbestandteile zerlegt“ (Lepsky, 1994, :1238), wird an anderer Stelle davon gesprochen, es gebe „algorithmische Teilschritte“ (Lepsky, 1996b). Zur Überprüfung der Anreicherungen des Index kann sich nun wiederum eine manuelle Bearbeitung des Wortmaterials anschließen. Ergebnis dieses Vorgehens ist ein stark angereicherter Index, der sehr viel mehr Wortmaterial bietet und so die Wahrscheinlichkeit erhöht, daß eine Anfrage des Benutzers trifft.

Entscheidend sind die Qualität und Vollständigkeit der verwendeten Wörterbücher: Sind diese zu niedrig, so werden zwangsläufig ein oder mehrere manuelle Bearbeitungsschritte nötig. Die Wörterbücher haben nach Abschluß der Projektphase MILOS I, innerhalb derer 200.000 Wörterbucheinträge (teil)automatisch generiert wurden, einen Umfang von 210.000 Stämmen für das Deutsche und je 80.000 Stämmen für Englisch und Französisch (Lepsky, 1996b). Aber auch bei vollständigen und qualitativ hochwertigen Wörterbüchern wird, wenn nicht im ersten Indexierungslauf, so doch später, regelmäßig manuelle Pflege und Wartung nötig sein, da das Vokabular sich durch neu hinzukommende Titel stetig um bislang unbekannte Wörter erweitern wird. „Nachteilig wirkt sich der erhöhte Aufwand für die notwendige Pflege der Wörterbücher aus. Dies gilt umsomehr für den Einsatz an einer wissenschaftlichen Universalbibliothek, deren fachlich breit gestreutes Textmaterial enorm hohe Ansprüche an das in den Wörterbüchern gespeicherte Vokabular stellt.“ (Lepsky, 1994, :1237)

In einem abschließenden Retrievaltest (Lepsky et al., 1996) auf 40.000 Titeldatensätzen der Universitätsbibliothek Düsseldorf wurden für 50 Beispielanfragen (vgl. 3.3.1) hinsichtlich Recall und Precision sehr gute Ergebnisse erzielt. So konnte der durchschnittliche Recall durch den erweiterten Index auf 51 % gegenüber 14 % für die Suche auf dem reinen Titelstichwortindex gesteigert werden. Allerdings wird der Ausgangspunkt von 14 % von den Projektteilnehmern selbst als skandalös niedrig angesehen (Lepsky, 1996b), ein Teil des guten Abschneidens von MILOS ist also einfach auch darauf zurückzuführen, daß *jeder* Verbesserungsversuch des bestehenden OPAC-Systems bereits einen Erfolg mit sich bringt. Wie schon zuvor in den Abschnitten zu OSIRIS muß auch an dieser Stelle darauf hingewiesen werden, daß hier nicht ein spezifisches lokales OPAC-System kritisiert wird, sondern von den Eigenschaften der OPAC-Kataloge der zweiten Generation, wie sie z.B. im ganzen norddeutschen Bibliotheksverbund eingesetzt werden, gesprochen wird. Dennoch sind die Ergebnisse des MILOS-Retrievaltests insgesamt so ermutigend, daß die Universitätsbibliothek Düsseldorf die in MILOS I entwickelten Indexierungsverfahren als Ergänzung in den Routinebetrieb übernommen hat (Lepsky, 1998).

Im Nachfolgeprojekt MILOS II (1995/96) wurden weitere Anreicherungen des Basic Index angestrebt, vornehmlich Wortrelationen auf Basis der Schlagwortnormdatei (SWD) und Mehrwortgruppenerkennung (Sachse et al., 1998). Da sich im Basic Index auch Schlagwörter befinden, die die Form elliptischer Mehrwortgruppen haben können („Buch- und Bibliothekswesen“), wurde in einem weiteren Schritt (nach der Dekomposition) die Auflösung und Rekonstruktion dieser Einträge vorgenommen („Buchwesen, Bibliothekswesen“). Weiterhin werden aus der Schlagwortnormdatei bekannte Relationen zwischen zwei Schlagwörtern (z.B. Hyponymie) oder einem Stichwort und einem Schlagwort in den Index integriert (Lepsky, 1996b). Das Relationenwörterbuch verfügt für das Deutsche über 480.000 Relationenpaare, davon wurden 180.000 Relationen in MILOS II auf Grundlage der Düsseldorfer OPAC-Daten (teil)automatisch generiert. Weiterhin wurde der gesamte Bestand der Sachschlagwörter der SWD (120.000 Datensätze) als Grundlage für die aus MILOS I bekannte mehrstufige Indexierung verwendet. Dabei wurden insgesamt etwa 100.000 Wörterbucheinträge neu erstellt oder überarbeitet. Ziel dieser Arbeiten unter Verwendung der SWD war es, möglichst viel semantische Information (Synonyme, Hypo- und Hyperonyme) zu integrieren, die in der SWD verfügbar ist. Darauf aufbauend wurde versucht, eine automatische Klassifikation von Titeldatensätzen relativ zur lokalen Bibliothekssystematik vorzunehmen (Lepsky, 1996a). Jedoch kommen die Projektteilnehmer zu dem Ergebnis, daß eine allein auf Titeldaten beruhende automatische Notationsvergabe auch auf einem um Informationen aus der SWD angereicherten Index keine Aussicht auf Erfolg hat.

Insgesamt erscheint das MILOS-Projekt in seiner Ausrichtung dem OSIRIS-Projekt sehr ähnlich, jedoch werden z.T. sehr unterschiedliche Methoden verwandt, um die gemeinsam erkannten Schwächen heutiger OPAC-Systeme zu überwinden. Am auffälligsten erscheint die starke Betonung wörterbuchbasierter Verfahren – der Ausdruck

„Wörterbuch“ wird dabei explizit und systematisch im MILOS-Kontext verwendet. Während in OSIRIS generell von „Lexika“ die Rede ist, um auch auf die interne Struktur sowie die integrierten Suchprozesse hinzuweisen, deutet „Wörterbuch“ einhergehend mit der explizit vorgesehenen manuellen Kontrolle und Pflege den statischen Charakter einer Wortliste an:

„[...] alle Manipulationen am Textmaterial beruhen auf einem Abgleich mit verschiedenen elektronischen Wörterbüchern. Gegenüber den algorithmisch basierten Verfahren [...] hat dieser Ansatz den deutlichen Vorteil einer höheren Genauigkeit, da es immer möglich ist, das Ergebnis der Indexierung über die Wörterbücher zu steuern.“ (Lepsky, 1994, :1237)

Ein weiterer wichtiger Unterschied zwischen OSIRIS und MILOS besteht in der Behandlung des Basic Index. Während OSIRIS den bestehenden Basic Index in die eigene Wissensbasis integriert und durch die Vorverarbeitung der Benutzereingabe für eine bessere Abbildung der Anfrage auf den Index sorgt, wird in MILOS der Index selbst erweitert, um der unveränderten Benutzeranfrage bessere Trefferchancen einzuräumen. Dabei werden in MILOS die konventionellen Suchmethoden eines Bibliothekssystems verwendet (pattern matching etc.), allerdings auf einer verbesserten Datengrundlage. In OSIRIS wird auf der systemeigenen Wissensbasis nicht mehr mit konventionellen Methoden gesucht, sondern Treffer werden in Relation zur Benutzeranfrage und zur Gesamttrefferlage bewertet. Und schließlich unterscheidet die Verwendung der lokalen Klassifikation die beiden Systeme: Während in OSIRIS der Index nur einen Teil der Wissensbasis ausmacht, die ansonsten noch über die Ergebnisse der statistischen Auswertung der lokalen Systematik verfügt, ist in MILOS der angereicherte Index die einzige Möglichkeit, die Suche zu verbessern. Einer der Vorteile, die eine Integration klassifikatorischer Elemente mit sich bringt, ist z.B. die Disambiguierung von Eingaben in thematisch unterschiedliche Klassen, wie sie in MILOS gewünscht wird (Gödert und Liebig, 1997). Auch die bereits angesprochene automatische Klassifikation von Titeldatensätzen in OSIRIS wird durch die Auswertung der lokalen Klassifikation erst möglich.

Die Arbeiten in OSIRIS an der Dialogschnittstelle, insbesondere die Erweiterung der Eingabemöglichkeiten hin zu einer beschränkt natürlichsprachlichen Eingabe haben in MILOS kein Gegenstück. Auch ist über die entwicklerzentrierte Evaluation hinaus keine Benutzeruntersuchung durchgeführt worden (vgl. für OSIRIS 6.1), so daß das OSIRIS-Projekt nicht nur in der Durchführung anders, sondern auch im Ansatz breiter eingeschätzt werden muß.

3.4.2 CANAL/LS

Das von der Europäischen Union 1995/96 im Rahmen des „Libraries-Program“ geförderte System CANAL/LS²⁰ ist, wie bereits erwähnt, eine „Ausgründung“ des MILOS-

²⁰Catalogue with Multilingual Natural Language Access / Linguistic Server

Projektes, an dessen Beginn die im Rahmen von MILOS nicht handhabbaren Schwierigkeiten einer Integration von Übersetzungen in die Retrievalfunktionalität standen. Darüber hinaus hat CANAL/LS aber den Entwurf einer universellen Schnittstelle für die Anbindung verschiedener linguistischer Analysekomponenten unternommen (Stegentritt, 1996). Ziel war es, ein konventionelles Bibliothekssystem über eine SGML-basierte Schnittstelle mit einem sogenannten „Linguistic Server“ zu verbinden. Diese Verwaltungskomponente soll für unterschiedliche Analyseprobleme mit einer Reihe verschiedener „Linguistic Engines“ in genanntem SGML-Protokoll kommunizieren können. Die bei Bedarf auch parallel von verschiedenen „Linguistic Engines“ erarbeiteten Ergebnisse werden beim „Linguistic Server“ gesammelt, ausgewertet und dem Bibliothekssystem übergeben. Schwerpunkt ist dabei die wortweise Übersetzung mit Hilfe verschiedener, auf bestimmte Übersetzungsrichtungen spezialisierter „Linguistic Engines“, die in der während der Projektlaufzeit entwickelten Version stark an die aus MILOS bekannten Wörterbücher erinnern, z.T. aber auch eine Fortentwicklung der Ergebnisse des Projektes EMIR²¹ (1990–94) (Stegentritt, 1992; Fluhr et al., 1996, 1997) zu sein scheinen. Offen bleibt, wie es mit Hilfe des SGML-Protokolls möglich ist, über die im Rahmen des Projektes angebotenen linguistischen Komponenten eines Projektpartners hinaus weitere fremde Komponenten anzubinden. Ob das entwickelte Protokoll hierzu universell und flexibel genug ist, scheint nicht einmal das Hauptproblem, vielmehr stellt sich die Frage, wie die „Linguistic Server“ genannte Verwaltungskomponente in die Lage versetzt werden soll, verschiedene Analyseprobleme unterschiedlichen Komponenten zu übergeben und deren z.T. vielleicht sogar konkurrierende Ergebnisse zu bewerten und zu vereinheitlichen. Der eigentliche Ausgangspunkt von CANAL/LS, die Unterstützung der Bibliotheksrecherche durch Übersetzungswörterbücher, konnte nicht überzeugen. Stegentritt kommt in einer Evaluation zu folgendem, angesichts der seit Jahrzehnten bekannten Komplexität der Problematik wenig überraschenden Ergebnis:

„Die Ergebnisse der mehrsprachigen Suche [...] zeigen vor allem, welche Probleme noch zu lösen sind. [...] Zusammengefaßt kann man sagen, daß sich hier alle wohlbekanntesten Probleme der automatischen Übersetzung zeigen.“ (Stegentritt, 1998, :45)

Insgesamt scheint CANAL/LS hinsichtlich der linguistischen Analyse breiter angelegt zu sein als OSIRIS : In CANAL/LS wird explizit versucht, eine Architektur mit offenen Protokollen für den Einsatz ganz unterschiedlicher Analysekomponenten zu definieren. In der Praxis ist von dieser möglichen Vielfalt der Analysekomponenten aber nichts zu sehen: CANAL/LS ist (soweit der Autor dies aufgrund der zwei zugänglichen Veröffentlichungen beurteilen kann) offenbar das geblieben, was es ursprünglich war, nämlich ein auf Übersetzungen spezialisierter Spin-Off des MILOS-Projektes.

Im Gegensatz zu OSIRIS wird in CANAL/LS selbst nicht an eine Veränderung des Retrievalvorganges gedacht. Vielmehr soll sich das System völlig transparent gegen-

²¹European Multilingual Information Retrieval

über einem konventionellen Bibliothekssystem verhalten und diesem optional behilflich sein. Auch wenn es sich dabei um ein bereits verbessertes System wie z.B. MILOS handelt, ist der Suchvorgang auf der Datenbank nach wie vor konventionell an pattern matching und Boolesche Operatoren gebunden.

3.4.3 VILIB

Mit dem Projekt VILIB²² fördert die Europäische Union 1998/99 wiederum im „Libraries-Program“ den weiteren Ausbau der CANAL/LS-Projektergebnisse. Ging es in CANAL/LS nur um eine einzelne Bibliothek, so soll mit der in VILIB entwickelten Lösung eine Suche in verschiedenen Bibliotheken über Sprachgrenzen hinweg möglich werden. Die Konzeption des in VILIB entwickelten Systems sieht eine wortweise Anfrageübersetzung vor, für die die Wörterbücher aus CANAL/LS um zusätzliche Sprachpaare erweitert und für zwei Beispielanwendergruppen („Management“ und „Wasser und Umwelt“) um Fachvokabular angereichert werden sollen. Die wortweise Übersetzung wurde offenbar aus CANAL/LS übernommen – deren Ergebnisse können nun unter einer HTML-Oberfläche vom Benutzer kontrolliert und manipuliert werden. So liegen die Suchanfrageerweiterung und die Disambiguierung in der Verantwortung des Suchenden. Gesucht wird in Bibliothekskatalogen, die mit dem Standard-Protokoll zum Austausch bibliothekarischer Daten Z39.50 angesprochen werden können. Welche Vorteile die Verwendung des sehr komplexen Z39.50-Protokolls gegenüber einer rein HTTP-basierten Kommunikation in diesem Zusammenhang bringen soll, wird nicht deutlich. Daß für die Suche in den Titeldaten verschiedener, verteilter Bibliothekskataloge auf die Verwendung von Z39.50 verzichtet werden kann, beweisen verschiedene Projekte, nicht zuletzt die CAI-Komponente in OSIRIS (vgl. 5.2.3). Neben der Suche auf Titeldaten ist auch die Suche in Volltexten möglich. Voraussetzung ist dabei allerdings, daß diese Texte mit einem bestimmten, kommerziellen Indexierprogramm bearbeitet wurden. Für die Zukunft ist die Unterstützung weiterer Indexierprogramme angekündigt (Stegentritt, 1999).

Die Ergebnisse des zum 31. Dezember 1999 ausgelaufenen Projektes werden unter <http://www.vilib.uni-koeln.de/> im WWW vorgestellt. VILIB ähnelt sehr stark dem CANAL/LS-Projekt – die Befragung verschiedener, im Internet verteilter Bibliothekskataloge macht den Eindruck einer Ausweitung des aus CANAL/LS bekannten Anwendungsszenarios. Neu erscheint lediglich die Volltextretrievalfunktionalität. Entsprechend gelten für den Vergleich zwischen VILIB und OSIRIS die bereits für CANAL/LS gemachten Bemerkungen.

3.4.4 KASCADE

Im MILOS-Folgeprojekt KASCADE²³ wurde 1997/98 der Versuch unternommen, im Spannungsfeld zwischen der Volltextsuche moderner Informationssysteme einerseits

²²Virtual Digital LIBrary on an Europe-wide level

²³Katalogerweiterung durch Scanning und Automatische Dokumenterschließung

und der Suche auf formal präzise erschlossenen aber nur spärlich vorhandenen Titeldaten in Bibliothekssystemen andererseits einen Schritt hin zu einer breiteren Datenbasis zu machen (Lepsky und Zimmermann, 1998). Anhand von Datensätzen des Fachgebietes Jura (Zeitschriften und Monographien) wurden dazu neue Möglichkeiten für die automatische Anreicherung und eine automatische Erschließung erprobt. Dazu gehörte die Erfassung und Verwertung von Inhaltsverzeichnissen, Abstracts und Registerbegriffen. Für etwa 3.000 Titel wurde versucht, die enthaltenen Informationen mit Hilfe von Scanning und OCR-Verfahren unter Zuhilfenahme der Wörterbücher aus dem MILOS-Projekt zu erschließen. Dies führte in der Praxis aber zu so großen Problemen mit der Erkennungsleistung, daß ein Einsatz im Routinebetrieb nicht in Frage kommt (Junger, 1999). Trotz der Schwierigkeiten mit der Scanning-Technik sehen die Projektinitiatoren keine handhabbare Alternative, um an die benötigten Daten zu gelangen. Zusätzlich sollten bereits elektronisch verfügbare Dokumente des in Frage stehenden Fachgebietes konvertiert und in die Datenbasis aufgenommen werden. Wesentliches Ziel war es, eine stärkere Anreicherung und darauf aufbauend eine bessere und automatische Erschließung der Daten zu erreichen.

Dazu wurde im Teilprojekt SELIX-JB damit begonnen ein Verfahren zu entwickeln, das eine Gewichtung der neu gewonnenen Daten ermöglicht und so ein unkontrolliertes Wachstum des Wortmaterials verhindern kann. Eine solche Gewichtung sollte in Abhängigkeit von der Relevanz des Ausdrucks erfolgen, die dieser Ausdruck für die Beschreibung eines Dokumentes hat. Zur Bestimmung der Relevanz und damit des Gewichtungsfaktors sollten verschiedene Parameter erforscht werden, von denen in der Literatur (Lepsky und Zimmermann, 1998; Junger, 1999) nur die Relation zwischen Auftretenshäufigkeit eines Ausdrucks in der gesamten Datenmenge und der innerhalb eines einzelnen Dokumentes genannt wird. Das Ziel dieser Relevanzberechnung wurde offenbar nicht erreicht: „Das SELIX-Verfahren zur Gewichtung von Indexaten erbrachte weder für die Quantität noch für die Qualität der Treffermengen die erwarteten Effekte. [...] Als Ergebnis bleibt festzuhalten, daß das SELIX-Verfahren in der bisherigen Form nicht für eine breitere Anwendung geeignet ist.“ (Junger, 1999, :89). Dies könne u.U. daran liegen, daß im vorangegangenen Indexierungsschritt, der aus dem MILOS-Projekt übernommen wurde, durch die Anreicherung des Vokabulars die spezifische Verteilung und damit die Grundlage der Relevanzbewertung eines Ausdrucks verschoben wird. Da die Berechnung der Gewichte ein sehr zeitaufwendiger Prozeß ist (mehrere Stunden für die erwähnten 3.000 Testdatensätze) und zudem nicht inkrementell vorgenommen werden kann, verbietet sich ein Einsatz im laufenden Betrieb eines größeren Bibliothekskataloges.

Im Teilprojekt THEAS-JB wurde mit Hilfe einer Themen-Aspekt-Analyse versucht, eine automatische Klassifikation der Dokumente vorzunehmen. Dieser Klassifikationsschritt, der bislang nur als Entwurf vorliegt, soll anhand dreier Analyseergebnisse eine Einordnung des in Frage stehenden Dokumentes in eine selbst konstruierte Systematik vornehmen. Dazu wird einerseits die Analyse des Volltextes herangezogen, der aufgrund vorhandener Phrasen wie *ein Kommentar zu* oder *ein praktisches Hand-*

buch über die „Zielgruppe“ Auskunft geben soll. Andererseits wird der Dokumenttyp zur Klassifikation herangezogen. Der Dokumenttyp ist entweder aus der formalen Erschließung (also über Schlagwörter) ersichtlich, oder muß aus Titel oder Abstract des Dokumentes erschlossen werden. Dazu wird nach Vorkommen von Wörtern oder Phrasen wie *Gesetzessammlung* oder *Festschrift* gesucht. Und schließlich soll eine Analyse des Dokumenttitels helfen, das Dokument zu klassifizieren. Diese Analyse stellt den Schwerpunkt des Klassifikationsvorganges dar: Hier wird versucht, auf Grundlage eines eigens erstellten Inventars von Phrasen wie *neue Erkenntnis, ein Vergleich zwischen ... und ...* oder *in Bayern* das Thema des Dokumentes zu bestimmen und andere, das Thema einschränkende Bestandteile des Titels näher in ihrer Funktion zu identifizieren.

Auf Grundlage dieser Ergebnisse soll später versucht werden, das Dokument in ein selbst entwickeltes „semantisches Bezugssystem“ einzuordnen, „das anhand der maschinenlesbaren Datensammlung durch statistische Analyse und intellektuelle Markierungen spezifischer lexikalischer Einträge aufgebaut wird“ (Lepsky und Zimmermann, 1998, :58). Ob es sich dabei um ein semantisches Netz oder ein Klassifikationssystem handelt, bleibt offen. Sicher ist lediglich, daß keine Informationen aus der klassifikatorischen Erschließung der Düsseldorfer Universitätsbibliothek verwendet werden: „Dahinter steckt nicht, wie man auf den ersten Blick vermuten könnte, die Verknüpfung von verbaler und klassifikatorisch-systematischer Information, wie sie in Form von Notationen in bibliothekarischen Titelaufnahmen enthalten ist“ (Junger, 1999, :89).

Insgesamt stellt sich das Projekt KASCADE als eine konsequente Fortsetzung der in MILOS begonnenen Arbeiten dar. Die Erweiterung des Problemfeldes auf Ansätze des automatischen Klassifizierens sowie der Analyse von Titeldaten zeigen in der Zielrichtung wiederum die Ähnlichkeit zwischen MILOS/KASCADE einerseits und OSIRIS andererseits. In der tatsächlichen Umsetzung (soweit sie in KASCADE bislang erfolgt ist) zeigen sich jedoch die bereits aus MILOS bekannten Unterschiede. Insbesondere ist es der Verzicht, im Rahmen von MILOS/KASCADE klassifikatorische Erschließungselemente sowie computerlinguistische Standardverfahren zu nutzen, der den größten Unterschied zu OSIRIS ausmacht.

Da die Vorschläge zur Themen-Aspekt-Analyse bislang nicht umgesetzt sind, erscheint ein echter Vergleich zwischen der in 3.2.4 für OSIRIS gezeigten Analyse des Parsers und dem für KASCADE vorgeschlagenen Weg nicht möglich. Auch sind keine Details des interessant klingenden Ansatzes bekannt. Lepsky und Zimmermann (1998) führen lediglich aus, daß aus einem manuell auf Basis der Testdaten erstellten lexikalischen Grundinventar ein offenes Regelsystem abgeleitet werden sollte, mit Hilfe dessen dann die Themen-Aspekt-Zuordnung ohne Unterstützung eines Lexikons durchführbar sei. Wie dies vor sich gehen soll und welche der klassischen Methoden aus Syntax- und Semantikanalyse dazu eingesetzt werden könnten, bleibt völlig offen. Junger merkt in ihrem Bericht über den KASCADE-Workshop an: „Der im Teilprojekt THEAS verfolgte Ansatz scheint überaus kompliziert zu sein und darüber hinaus für eine echte thematische Zuordnung von Dokumenten wenig auszutragen“ (Junger, 1999, :89). Ins-

besondere kritisiert Junger, daß in KASCADE die klassifikatorischen Erschließungsdaten nicht verwendet werden.

3.4.5 Die OKAPI-Projekte

Seit Anfang der achtziger Jahre fördert die British Library im Rahmen der OKAPI-Projekte eine kontinuierliche Entwicklung besserer Retrievalsysteme. Diese Projekte wurden zunächst an der politechnischen Universität London und später an der City University London durchgeführt. OKAPI zeichnet sich international durch den langen Zeitraum der Entwicklung aus, der (immer wieder durch Benutzerevaluationen korrigiert) von der zeichenorientierten VT100-Schnittstelle bis zur graphischen Tcl/Tk-Oberfläche ein breites Spektrum von Verbesserungen hervorgebracht hat (Beaulieu, 1997; Robertson et al., 1997; Robertson und Beaulieu, 1997).

Im Zentrum stehen dabei Verbesserungen des Retrievalvorganges durch den Einsatz verschiedener Methoden zur query expansion sowie unterschiedlicher Termgewichtungsverfahren. Auf der Oberfläche hat der Benutzer die Möglichkeit, als Suchanfrage Freitext einzugeben, der vom System wortweise behandelt wird (Robertson, 1997). Diese Behandlung umfaßt automatische morphologische Reduktion mit Hilfe eines Stemming-Programmes und die Gewichtung der einzelnen Wörter relativ zu ihrer Häufigkeit im Datenbestand. In einigen Varianten von OKAPI wurde die Möglichkeit vorgesehen, in der Benutzereingabe Phrasen zu identifizieren. Diese werden aber nicht mit computerlinguistischen Methoden behandelt, sondern mit Hilfe verschiedener Operatoren (Adjazenz u.a.) durch eine aufwendige Hintergrundsuche bearbeitet. Mit Hilfe des Thesaurus können zu einzelnen Bestandteilen der Benutzeranfrage Synonyme geliefert werden. Das Ergebnis einer solchen bearbeiteten und erweiterten Benutzeranfrage ist eine aufgrund der Termgewichtung sortierte Liste möglicherweise relevanter Dokumente, für die der Suchende jeweils eine Relevanzbewertung vornehmen kann („Is this the kind of thing you want?“). Die Markierung einzelner Dokumente als relevant kann dann den Ausgangspunkt einer nachfolgenden Suche (relevance feedback search) bilden, zu der das System das verfügbare Vokabular der vom Nutzer als relevant markierten Dokumente heranzieht (Robertson, 1997). Viel Entwicklungsarbeit wurde in die Verteilung der Kontrolle der Anfrageanreicherung investiert. In verschiedenen Ansätzen wurde den Suchenden unterschiedlich viel Kontrolle über den Expansionsprozeß gegeben. Dabei zeigte sich in der Evaluation, daß die Benutzer mit zuviel Kontrollmöglichkeiten überfordert waren und lieber ihre Ausgangsfrage modifiziert als neue Suchanfrage dem System stellten, als interaktiv ihre Anfrage zu erweitern. Die Benutzer, die sich auf die Systemvorschläge einließen und ihre Anfrage mit Hilfe des angebotenen Materials erweiterten, waren allerdings nicht bereit, mehr als einige wenige Suchbegriffe hinzuzufügen: 70 % der angebotenen Suchbegriffe wurden explizit von den Benutzern abgelehnt, die durchschnittliche Länge der Benutzeranfrage steigerte sich gegenüber 2.2 Suchtermen in der Ausgangsfrage nur auf 4.4 Suchterme in der expandierten Variante (Beaulieu, 1997). Auch für die überarbeitete Variante, in der die Belastung der Suchenden bei der Auswahl der zur Anfrageexpansion verfüg-

baren Terme deutlich reduziert wurde, stellt Beaulieu (1997) immer noch fest, daß enttäuschend wenige Benutzer die verfügbare neue Suchfunktionalität nutzen würde. Als vordringliches Ziel weiterer Entwicklungen werden deshalb zwei Faktoren bezeichnet, die nach Beaulieu (1997) wesentlich für eine Entlastung der Benutzer sind: Einerseits der Grad, zu dem die Komplexität des Systems sichtbar gemacht werden kann und andererseits die Aufteilung der Kontrolle zwischen Suchendem und System. Diese beiden Punkte seien nur durch eine detaillierte Untersuchung der Gestaltungsspielräume zwischen den Polen automatische und interaktive Anfrageexpansion, explizite und implizite Thesaurusverwendung sowie die Metakommunikation zwischen Mensch und Maschine über einerseits die Bestandteile der Benutzeranfrage selbst und andererseits die Ergebnisse der Suche nach Dokumenten lösbar.

Die Familie der OKAPI-Projekte zeichnet sich gegenüber anderen Projekten wie z.B. OSIRIS durch ihren Variantenreichtum und die zahlreichen parallel untersuchten Verbesserungen aus. Diese Vielfalt wird allerdings erst möglich durch die langfristige und kontinuierliche Förderung durch die British Library. Dabei werden neben der ergonomischen Oberflächengestaltung und der regelmäßigen Überprüfung der Projektergebnisse durch Auswertung von Log-Dateien (Jones et al., 1997a) und Evaluationen (Robertson et al., 1997; Robertson und Beaulieu, 1997) Fortschritte wesentlich durch Retrievalverbesserungen erzielt. Eine Eingabeverarbeitung findet nur in eingeschränktem Maße statt, der Freitext, den der Benutzer eingeben kann, wird lediglich als Liste von Suchtermen interpretiert, deren einzige Struktur im Fall erkannter Phrasen der Abstand des Auftretens (nahe beieinander ohne Einhaltung der linearen Folge oder linear direkt aufeinander folgend) ist. Obwohl bei Robertson (1997) der Ausdruck „parsing“ an einer Stelle fällt, kann doch nicht von Parsing im Sinne der Computerlinguistik gesprochen werden, weil weder eine Grammatik noch ein Lexikon bei der Zerlegung der Benutzereingabe genutzt werden. Große Probleme bereitet den OKAPI-Projekten die Entwicklung des Internet: Ursprünglich konzipiert für kontinuierliche Benutzerinteraktionen auf der Grundlage des aus der Bibliothekswelt bekannten Z39.50-Protokolls, haben die OKAPI-Projekte Schwierigkeiten, als verteilte Systeme mit Hilfe des HTTP-Protokolls stateless realisiert zu werden. Als Ausweg schlagen die Entwickler die Verwendung von Java vor (Jones et al., 1997b), was aber zu problematischen Anforderungen an die Client-Software führt. Wie bereits für die anderen Projekte gezeigt, unterscheidet sich der Ansatz von OSIRIS erheblich, insbesondere durch die natürlichsprachliche Eingabemöglichkeit und die computerlinguistische Analyse sowie hinsichtlich der Berücksichtigung klassifikatorischer Information für die Suche und Präsentation der Ergebnisse.

3.5 Fazit

Abschliessend soll es darum gehen, das OSIRIS-System mit den Eigenschaften zu vergleichen, die in 2.3 als Fazit der Untersuchung von Online-Schnittstellen genannt wurden. Dabei war der erhebliche intellektuelle Aufwand für die Konstruktion geeigneter

Suchanfragen als ein wesentliches Hindernis für den erfolgreichen Einsatz von Dialogschnittstellen genannt worden. In dieser Hinsicht stellt OSIRIS aus Entwicklersicht einen echten Fortschritt gegenüber den bislang existierenden Systemen dar. Die Konstruktion natürlichsprachlicher Eingaben in OSIRIS ist für den Benutzer ohne großen Lernaufwand möglich. Insbesondere war in 2.3 der Einsatz Boolescher Operatoren und regulärer Ausdrücke für Dialogschnittstellen als wenig hilfreich kritisiert worden, da diese weder für die Suche auf *sprachlichen* Daten geeignet sind noch von den Suchenden beherrscht werden. Gegenüber Systemen mit diesen Eingabeformen kann OSIRIS als überlegen betrachtet werden, da OSIRIS auf Boolesche Operatoren und reguläre Ausdrücke an der Eingabeschnittstelle verzichtet, ohne dabei aus Entwicklersicht die Anfragemöglichkeiten des Benutzers zu beeinträchtigen. Ein weiterer Punkt, der für bestehende Schnittstellen kritisiert wurde, war die Notwendigkeit, über die Struktur des Datenbestandes informiert zu sein, um erfolgreich Anfragen gestalten zu können. Hiermit verknüpft war das Problem, die Qualität von Anfrageergebnissen einzuschätzen: Ohne detailliertes Wissen über den Datenbestand und das verwendete System, so wurde in 2.3 festgestellt, ist es heutzutage nicht möglich abzuschätzen, von welcher Qualität die Suchergebnisse sind und wieviele vergleichbare Treffer es eventuell noch im Datenbestand gibt. Auch hinsichtlich dieses Punktes stellt das OSIRIS-System einen deutlichen Fortschritt dar. Durch die Berücksichtigung klassifikatorischer Information beinhalten die Systemantworten eine Vollständigkeitszusicherung an den Suchenden: Es werden nicht nur die aufgrund des Titelmaterials passenden Objekte gefunden, sondern durch die Abbildung der Suchanfrage auf die natürlichsprachliche Beschreibung der Klassen, alle aufgrund der Systematik als inhaltlich verwandt zu betrachtenden Titel. Dies stellt eine enorme Verbesserung der Situation des Suchenden dar.

Mit der Verbesserung der Eingabemöglichkeiten und der Qualität der Suchergebnisse einher gehen in OSIRIS Verbesserungen in Bezug auf die ergonomische Gestaltung der Schnittstelle. So stellt die hierarchische Präsentation in einem Ausschnitt der Klassifikation einen vielversprechenden Ansatz dar, Suchergebnisse übersichtlich und verständlich zu präsentieren. Wie für die Einfachheit der natürlichsprachlichen Schnittstelle gilt aber auch hier, daß die tatsächlichen Effekte auf die Benutzerzufriedenheit sich aus Entwicklersicht nur vermuten lassen. Ob und wenn ja wie sehr OSIRIS (verglichen mit herkömmlichen Dialogschnittstellen) von Benutzern einfacher zu bedienen ist, kann nur durch Nutzerbefragungen und Evaluationen geklärt werden. Auf diesen Punkt wird in Kapitel 6 eingegangen. An dieser Stelle ist ersteinmal nur von Belang, daß die bislang genannten Kritikpunkte in OSIRIS angegangen und aus Entwicklersicht deutlich entschärft, wenn nicht gar gelöst worden sind.

Insbesondere für Schnittstellen zu Bibliothekskatalogen wurde in 2.3 festgestellt, daß sie durch eine an den Zwängen vergangener Hard- und Softwaregenerationen orientierte Gestaltung der Schnittstellen geprägt seien. Weiterentwicklungen zielten lediglich auf kosmetische Verbesserungen ab, bei unverändert mangelhafter Funktionalität. OSIRIS verfolgt hier eine gemessen an Bibliothekssystemen radikale Abkehr von über-

kommenden Konzepten, indem es stateless operiert und auf Seiten des Endanwenders nur einen herkömmlichen WWW-Browser erwartet. Obwohl OSIRIS auf der Seite der Eingabeanalyse auf zahlreiche Ergebnisse früherer Arbeiten zurückgreift, handelt es sich bei der Datenaufbereitung, beim Retrieval und bei der Präsentation der Retrievalergebnisse um vollständige Neuentwicklungen. Diese Neuentwicklung war nötig und umfassend, weil die erwünschten Verbesserungen beim Retrieval zu einem großen Teil nur unter Zuhilfenahme neuer und eigens modifizierter Techniken möglich waren. Zahlreiche Publikationen und Systempräsentationen belegen, daß das entstandene System dem Anspruch, state of the art Technologie einzusetzen, gerecht wird. Der Vorwurf an konventionelle Bibliotheksrecherchesysteme, keine Fortschritte hinsichtlich der Funktionalität zu machen, läßt sich aufgrund der kurzen Einsatzdauer von bislang nur etwas über 2 Jahren nicht prinzipiell für OSIRIS ausschließen. Dennoch zeigen die in 6.1 dargestellten und aufgrund einer permanenten Evaluation möglichen Fortschritte von der Version 1.0 bis zur aktuellen Version 3.0, daß den Entwicklern die Notwendigkeit einer permanenten, nutzergesteuerten Funktionsverbesserung bewußt ist.

Weitere Kritikpunkte an Online-Schnittstellen waren in 2.3 die mangelnde Robustheit des Systems gegenüber Fehlern oder Ambiguitäten in der Eingabe. In beiden Punkten kann OSIRIS mit deutlichen Verbesserungen aufwarten. Durch die morphologische Reduktion im Verlauf der Eingabeverarbeitung werden verschiedene Fehlerquellen herkömmlicher Systeme wie Numerus- und Kasusmarkierungen sprachlicher Eingaben beseitigt. Durch Kompositazerlegung sowie die Integration weiterer Daten (z.B. Synonyme) erlangt OSIRIS zusätzliche Robustheit gegenüber Variationen in der Eingabe sprachlicher Daten. Die explizite Fehlererkennung und -korrektur (vgl. 3.2.7) innerhalb der Eingabeanalyse erkennt darüber hinaus eine ganze Reihe typischer Eingabefehler wie Vertauschungen und Auslassungen, auf die in Abhängigkeit vom identifizierten Fehler reagiert werden kann. Die Suche nach Eigennamen auf einem phonetischen Index (vgl. 3.2.9) reduziert Fehler bei der häufig nicht eindeutigen Schreibung von Eigennamen. Dabei ist die gesonderte Behandlung von Eigennamen einerseits als Vorteil zu werten, weil es sich um ein speziell auf diese Datenart abgestimmtes Korrektiv handelt. Andererseits ergibt sich so auf der Eingabeseite eine dem Suchenden aufgezwungene Trennung in Namen und Themen, weil OSIRIS nicht in der Lage ist, eigenständig zu erkennen, ob es sich bei einer Nutzereingabe um einen Personennamen oder ein Thema handelt. Insbesondere kommt es so zu einer unterschiedlichen Behandlung eines Eigennamen, wenn dieser als Autorennamen eingegeben wird und wenn er Thema einer Literatursuche ist.

Eine weitere Forderung aus 2.3 besagte, Online-Schnittstellen sollten so intuitiv, übersichtlich und verständlich gestaltet werden, daß sie ohne Training zu benutzen sind. Sowohl die Annäherung an dieses Ziel als auch die Überprüfung desselben konnte in OSIRIS nur durch eine wiederholte Systemevaluation sichergestellt werden, da auch die besten Vorsätze und Intuitionen der Systementwickler zur Oberflächengestaltung und Symbolverwendung sich als nur begrenzt praxistauglich erwiesen haben. Daß ent-

wicklerzentrierte Oberflächengestaltung meist nicht zu angemessenen Resultaten für unbedarfte Endnutzer führt, ist dabei keine Überraschung – die systematische Einbeziehung von Arbeits- und Organisationspsychologen für die Systemevaluation ist der eigentlich interessante Punkt, der für die Weiterentwicklung von OSIRIS über das aktuelle Projekt hinaus von Bedeutung ist.

Kritisiert wurden in 2.3 auch die teilweise starken Vorannahmen zum Wissensstand des Nutzers durch die Wortwahl auf der Systemoberfläche. In OSIRIS ist die Gestaltung der Oberfläche darauf ausgelegt, lange Erklärungen zu vermeiden: Wenn möglich wird ausschließlich die natürlichsprachliche Eingabemöglichkeit, die in der Vervollständigung des Satzanfangs liegt (vgl. 3.1), angeboten. Dabei werden andere, als unzureichend kritisierte Versuche der Verbesserung der Interaktion wie z.B. vorformulierte Schemata, in denen Nutzer sich selbst oder ihr Anliegen einordnen müssen sowie die detaillierte Antizipation der Benutzerbedürfnisse in der Struktur der Eingabefelder vermieden. Daraus folgt aber noch nicht, daß die eingeschränkt natürlichsprachlichen Eingabemöglichkeiten in OSIRIS optimal für menschliche Benutzer z.B. hinsichtlich der Lernbarkeit oder besonders geeignet für den Gegenstandsbereich der Suche wären (vgl. 4.1).

Die ebenfalls in 2.3 angesprochene stärkere Nutzung graphischer Repräsentationsmittel, um die Bedienbarkeit der Oberfläche zu verbessern, wurde in OSIRIS nicht über einfache Gestaltungsrichtlinien für Systemoberflächen (in Kombination mit den Ergebnissen der Evaluation) hinaus umgesetzt. Erste Ideen, wie der Benutzer durch rein graphische Mittel unterstützt werden könnte, werden in 5.2 angesprochen.

Die Unverständlichkeit allgemeiner Systemmeldungen, insbesondere aber der Fehlermeldungen war ein weiterer Kritikpunkt in 2.3. Wie bereits weiter oben für die allgemeine Verbesserung der Systemfunktionalität angeführt, wird auch für diesen Punkt in OSIRIS versucht, über Systemevaluationen und Nutzerbefragungen eine Verbesserung zu erreichen. Hintergrund dieser Vorgehensweise ist die Annahme, daß eine Verbesserung von Systemmeldungen wie z.B. Fehlermeldungen kein rein technisches Problem ist, sondern auch abhängig vom Bewußtsein der Entwickler ist. Als Beispiel möge Krause (1980) dienen, der über eine natürlichsprachliche Datenbankschnittstelle, das USL-System (vgl. S. 100) und dessen Evaluation berichtet. USL reagiert auf einen Tippfehler genauso wie auf die Verwendung einer im Sinne des Systems falschen Präposition mit der Meldung „Nicht verstanden. Bitte Prüfbefehle oder Leerzeile eingeben“. Krause schreibt in Hinblick auf diese Systemmeldung:

„Der Benutzer hat dann die Möglichkeit, eine Prüfroutine zu starten, die der Fehlersuche dient. Interessiert er sich nicht dafür, was auf 99,9 % der Fälle zutrifft, muß er [...] eine 'Leere Eingabe' abschicken, wenn er weiterfragen möchte. [Die Evaluation] zeigt, daß in etwa ein Viertel aller Fälle die Leerzeile nicht eingegeben wurde, was einen neuen Fehler verursachte, obwohl der Text der Fehlermeldung an Klarheit nichts zu wünschen übrig ließ.“(Krause, 1980, :215)

Krause kommt aber zu der Einsicht, daß es nicht reicht, eine Fehlermeldung (in seinem Sinne) „klar und deutlich“ abzufassen, sondern daß es dem Benutzer möglich sein muß, direkt und ohne Eingabe einer Leerzeile nach einem Fehler eine neue Anfrage zu stellen. Aber auch andere Fehlermeldungen wie „ONCODE: 8094“ (von Krause nicht problematisiert) können zumindest aus heutiger Perspektive nur als abschreckendes Beispiel dienen.

Um die Situation des Suchenden an einer Online-Schnittstelle zu verbessern, wurden in 2.3 neben den bislang angesprochenen Punkten die folgenden Vorschläge gemacht:

- Die Benutzung von Online-Schnittstellen sollte einheitlich sein. Dazu ist nicht unbedingt ein Standard vonnöten. Wenn intelligente Schnittstellen es dem Benutzer ermöglichen, sein Informationsbedürfnis natürlichsprachlich mitzuteilen, wird der Benutzer entlastet und die verschiedenen Systeme aus Benutzersicht einheitlicher, ohne daß es notwendigerweise auf Systemseite zu einer Standardisierung kommt.
- Eine intelligente Verarbeitung der Benutzereingaben, um Robustheit gegenüber Eingabefehlern sowie Entlastung des Suchenden bei der Anfrageformulierung zu erreichen.
- Systeminhärente Intelligenz, um Nutzern ohne einschlägige Kenntnis der Domäne die erfolgreiche Suche zu ermöglichen. Dazu müssen die Interessen und Ziele des Benutzers erkannt und selbständig durch das System auf die Besonderheiten des Datenbestandes abgebildet werden. Dabei sollte die Schnittstelle sich dynamisch den variablen Bedürfnissen des Nutzers anpassen, eigenständig die für das aktuelle Informationsbedürfnis adäquate Informationsquelle auswählen und diese unter Berücksichtigung ihrer spezifischen Eigenschaften befragen.
- Einbettung des Suchvorganges in den gesamten Arbeitsablauf des Nutzers durch Übernahme der Rechercheergebnisse in vereinbartem Format.

OSIRIS erreicht diese sehr weit gesteckten Ziele nicht, auch ist dem Autor kein anderes System bekannt, daß diesen umfassenden Ansprüchen gerecht werden würde.²⁴ Jedoch ist für OSIRIS festzuhalten, daß es sich an diesen Zielen orientiert und in einigen Punkten diesen Zielen deutlich nähergekommen ist als andere Systeme.

Zur geforderten Entlastung des Suchenden durch *Einheitlichkeit* des Systems trägt OSIRIS in mehrfacher Hinsicht bei. So werden zum einen in der speziellen Anwendungsdomäne, der Suche auf Bibliotheksdaten, dem Nutzerinteresse folgend die bislang stets getrenntgehaltenen Funktionen „Sachrecherche“ und „Themenrecherche“ erstmals unter einer Abfrage zusammengefaßt (vgl. 3.1). Für zukünftige Erweiterungen des Systems ist dabei auch die Integration weiterer Funktionen wie die Suche nach

²⁴Da der angeführte Kriterienkatalog mit dem Ziel entworfen wurde, den gegenwärtigen Zustand auf dem Gebiet der Dialogschnittstellen zu verbessern, darf ein gewisser utopischer Zug nicht verwundern.

Personen in einer Expertendatenbank oder nach Datenbanken in einem thematischen Verzeichnis von Datenbanken angedacht. Vereinheitlichung heißt in OSIRIS also zunächst einmal, verschieden strukturierte Datenbestände mit einer Anfrage parallel zu durchsuchen. Die Unterschiede zwischen den Datenbeständen können dabei natürlich nicht arbiträr sein (wenn man qualitativ hochwertige Rechercheergebnisse gewährleisten will), jedoch kann eine Verbesserung von OSIRIS auf dem erreichten Level quantitativ, d.h. durch stetige Integration zunehmend heterogenerer Datenbestände erfolgen. Einheitlichkeit heißt für OSIRIS in einem ganz anschaulichen Sinne darüber hinaus aber auch, daß die Systemoberfläche dem Benutzer bekannte Elemente enthält. Diesen Umstand erreicht OSIRIS durch die (im Anwendungskontext von OSIRIS bis dato unbekannt) alleinige Konzentration auf WWW-Browser und die von ihnen dargestellten HTML-Seiten. Potentiell ist damit jeder, der über Vorerfahrung als Nutzer des WWW verfügt, in der Lage, OSIRIS zu bedienen. Dabei spielt nicht nur die einfache und übersichtliche Gestaltung der Oberfläche eine Rolle, sondern vor allen Dingen der Wiedererkennungseffekt beim Nutzer, der Hemmschwellen abbauen hilft und Erfahrungen aus anderen Bereichen des Umgangs mit dem Computer auf OSIRIS übertragbar macht.

Dies ist zwar ein Schritt in die richtige Richtung, kann angesichts der Gestaltungsmöglichkeiten von HTML-Dokumenten²⁵ allein aber nicht ausreichen und muß deshalb durch Evaluationen unterstützt werden (vgl. 6). Einheitlichkeit im Sinne des angesprochenen Verbesserungsbedarfs von Online-Schnittstellen meint schließlich für OSIRIS die Möglichkeit des Suchenden, natürlichsprachliche Eingaben zu tätigen. So kommt es (wie angeregt) auf der Eingabeseite zu einer Vereinheitlichung, ohne daß davon die tatsächliche interne Verarbeitung der Eingabe betroffen sein muß (in der Tat ist OSIRIS ja modular aufgebaut und könnte die natürlichsprachliche Eingabeanalyse auch anderen nachfolgenden Komponenten zur Verfügung stellen, vgl. 3.1).

Damit ist aber noch nicht entschieden, ob die Möglichkeit, natürlichsprachliche Eingaben zu tätigen, wirklich die für den Suchenden beste Interaktionsform ist (auf diesen Punkt wird in Kapitel 6 weiter eingegangen). Insbesondere muß dabei berücksichtigt werden, daß wirklich freie, d.h. unbeschränkte Eingabe natürlicher Sprache derzeit noch Gegenstand der Forschung ist. Aus diesem Grunde muß für alltagstaugliche Systeme irgendeine Einschränkung der Eingabemöglichkeiten, z.B. eine Untermenge der natürlichen Sprache konstruiert werden. Diese Punkte werden in 4.1 weiter diskutiert, wenn dort ein Spektrum unterschiedlicher Arten von Schnittstellen vorgestellt wird.

Die geforderte intelligente Verarbeitung, um Robustheit gegenüber Eingabefehlern zu erlangen, wurde weiter oben in diesem Abschnitt (S. 90) bereits erläutert. In diesem Punkt ist OSIRIS allen anderen dem Autor bekannten Dialogsystemen überlegen. Da-

²⁵Hier sind natürlich nicht nur die Möglichkeiten von HTML, sondern auch die spezifischen Einschränkungen von Bedeutung, die u.a. zu Weiterentwicklungen wie den gerade im Standardisierungsprozeß befindlichen XML-Sprachen geführt haben.

bei handelt es sich aber nur um eine erste Verbesserungsstufe²⁶ – in Kapitel 4.2 werden die wünschenswerten Möglichkeiten einer Verbesserung und Ausweitung dieser Eingabeverarbeitung dargestellt.

Auch bei der angemahnten systeminhärenten Intelligenz, die es Suchenden erlauben soll, ohne Kenntnis der Struktur des Datenbestandes verlässliche Suchergebnisse zu bekommen, ist OSIRIS anderen Recherchesystemen entscheidend voraus. Wie in 3.1 dargestellt, ermöglicht es die Berücksichtigung klassifikatorischer Information, dem Suchenden eine gemessen am Datenbestand vollständige und zuverlässige Antwort zu geben. Diese neue Zuverlässigkeit der Suchergebnisse ist aber auf zwei Dinge angewiesen: das Vorhandensein einer zugrundeliegenden Klassifikation als Voraussetzung einer intelligenten Aufbereitung der Daten und das Vertrauen der Suchenden in das Ergebnis. Beides ist nicht selbstverständlich und, wenn es fehlt, nicht einfach zu bekommen. Auf die Schwierigkeiten, einen größeren Datenbestand anhand eines Kategorienschemas zu klassifizieren oder sogar erst ein geeignetes Kategorienschema dafür zu entwerfen, wird in 5.2 kurz eingegangen.

Die ebenfalls vorgeschlagene Einbettung der Systeme in größere Arbeitsabläufe durch z.B. Übernahme von Rechercheergebnissen in anderen Dateiformaten ist bislang für OSIRIS nicht realisiert. Derzeit kann ein Nutzer lediglich ASCII-Text in der vom System gelieferten Struktur während einer Sitzung speichern und entweder ausdrucken oder per cut-and-paste in eine andere Anwendung übernehmen. Der direkte Export in z.B. BIBTEX ist angedacht, aber noch nicht realisiert (vgl. 5.2).

Auch wenn OSIRIS in vielen Punkten den genannten Forderungen oder Verbesserungen noch nicht genügt, so schneidet das System doch im Vergleich zu anderen Schnittstellen sehr gut ab. Besonders positiv sind hier die Aufbereitung der Daten und die natürlichsprachliche Schnittstelle. Die Aufbereitung der Daten führt zu verlässlichen und vollständigen Suchergebnissen. Die natürlichsprachliche Eingabemöglichkeit reduziert die Belastung des Suchenden, indem sie ihm die Formulierung der Anfrage in einer künstlichen Eingabesyntax abnimmt und durch die Analyse und Fehlerkorrektur für eine bessere Abbildung auf den Datenbestand sorgt. Die natürlichsprachliche Schnittstelle ist zudem domänenunabhängig und trägt im obengenannten Sinne zu einer Vereinheitlichung der Systeme bei, die man als Ausrichtung an „kognitiven Standards“ bezeichnen könnte.

In Kapitel 5 soll es daher um die Grenzen und Verbesserungsmöglichkeiten der linguistischen Verarbeitung in OSIRIS gehen, sowie um die Anforderungen an die theoretische Behandlung der an natürlichsprachlichen Schnittstellen zu erwartenden Phänomene. Dazu werden in 4.1 zunächst die Eigenschaften unterschiedlicher Eingabeschnittstellen diskutiert, um die Möglichkeiten unterschiedlich umfangreicher und leistungsfähiger NL-Schnittstellen im Vergleich zu anderen Lösungen beurteilen zu können. Diese Diskussion ist damit genauso wichtig für den weiteren Ausbau der OSIRIS-Schnittstelle, wie die in 4.2 geführte Diskussion um den Phänomenbereich.

²⁶Zu den Grenzen der Eingabeverarbeitung, siehe 4.1.

Die in Kapitel 6 geschilderte Evaluation des Systems schließlich belegt, daß OSIRIS nicht auf dem Erreichten verharrt, sondern sich kontinuierlich weiter in Richtung auf die genannten Ziele einer intelligenten, bedienerfreundlichen Schnittstelle hin bewegt. Dabei wird es insbesondere um die Erwartungshaltung der Nutzer gehen, die aus Eigenschaften des Systems wie der komplexen internen Anfrageverarbeitung und der NL-Schnittstelle entstehen kann.

4 Eingabemöglichkeiten an Dialogschnittstellen

4.1 Dialogschnittstellen im Vergleich

Um die Eigenschaften der Eingabeverarbeitung von OSIRIS beurteilen und Aussagen über deren Angemessenheit treffen zu können, ist es sinnvoll, die Möglichkeiten des Systems mit denen anderer Arten von Dialogschnittstellen zu vergleichen. Dazu erscheint eine Betrachtung der Freiheitsgrade von Mensch-Maschine-Schnittstellen und deren Auswirkungen auf die Systemarchitektur wünschenswert.

Am einen Ende einer solchen Skala von Freiheitsgraden stehen die *rein passiv* zu nutzenden Schnittstellen. Nachrichtenticker, Wetterkarten, Börsenkurse und ähnliche Informationsangebote lassen sich (wie das derzeitige Fernsehprogramm) vom Betrachter nur dadurch beeinflussen, daß er sich bewußt für oder gegen ihre Nutzung entscheidet. Konkrete Beispiele für solche Systeme sind heute selten, weil (wie in 2.1 beschrieben) der Konkurrenzdruck im WWW stetig wächst und Angebote dieser Art angesichts der Möglichkeiten des Hypertextes zumeist als nicht mehr zeitgemäß angesehen werden. Allerdings gibt es zunehmend Übergangsbereiche, in denen sich das Internet mit anderen Medien wie dem Fernsehen vermischt, und die durch die Eigenschaften beider Medien gekennzeichnet sind. Web-Kameras sind ein Beispiel für die Vermischung von Fernsehen und Internet und dadurch zugleich ein Beispiel für ein rein passiv zu nutzendes Informationsangebot.

Einen kleinen Schritt weiter gehen Angebote mit *Auswahlmöglichkeiten*. Sie erlauben es dem Betrachter z.B. mittels eines Systemmenüs oder durch die Möglichkeiten des Hypertextes, aus dem vorgefertigten und vorstrukturierten Informationsangebot durch Klicken oder Anwahl von Systemoptionen aktiv eine Teilmenge des Angebotes zu wählen. Je nach Ausprägung der internen Struktur und Umfang der selektierbaren Objekte kann diese Wahlmöglichkeit von einer binären Entscheidung zwischen zwei Teilmengen (Endverbraucherpreise versus Großhandelspreise oder Damenbekleidung versus Herrenbekleidung) bis hin zur Konstruktion einer virtuellen Sicht auf die Daten (z.B. bei der Konstruktion eines umfangreichen Börsenportfolios) reichen. Wiederum einen Schritt weiter gehen Anbieter, die ihr *Informationsangebot personalisieren* und je nach Interessenlage des Nutzers Teile des Angebotes ausblenden oder aber den Grad der Detailliertheit der Ausgabe variieren. Eine wichtige Frage in diesem Zusammenhang ist natürlich, wie das für die erfolgreiche Personalisierung notwendige Wissen in das System kommt. Hierzu gibt es eine Vielzahl verschiedener Ansätze, angefangen von der Aufforderung an den Benutzer, sich beim ersten Systemkontakt zu bestimmten Fragen zu äußern oder sich vorgegebenen Benutzergruppen zuzuordnen, bis hin zu Methoden des maschinellen Lernens, die aus der Beobachtung des Benutzerverhaltens auf seine Interessenlage schließen sollen. Die in diesem Zusammenhang relevanten Fragen, ob und wie beispielsweise die Zufriedenheit eines Nutzers mit einem Informationsangebot sicher und effizient zu erkennen ist, können in dieser Arbeit nicht weiter behandelt werden. Beispiele für solche Systeme sind das Yahoo Portfolio und der Online-Buchhandel Amazon. Ersteres erlaubt es dem Benutzer, für jedes Wertpapier

explizite Angaben über den Detailgrad der gewünschten Informationen zu machen. Letzterer empfiehlt zu einem vom Benutzer selektierten Buch auf der Grundlage der Auswertung des Verhaltens anderer Käufer dieses Werkes weitere Bücher.

All diesen Schnittstellen ist gemeinsam, daß der Benutzer (wenn überhaupt) nur im Vorwege das Informationsangebot des Systems aktiv bestimmen kann. Die eigentliche Interaktion kann er nur innerhalb der dann vom System angebotenen Wahlmöglichkeiten und ausschließlich durch binäre Entscheidungen beeinflussen. Dadurch ist die explizite Auswahl von Objekten aus Datenbeständen möglich – nicht möglich sind jedoch komplexere Anfragen zum Datenbestand oder die Suche in den Daten selbst.

Eine Möglichkeit, solche Anfragen innerhalb der bisher beschriebenen Systeme umzusetzen, besteht in der Erweiterung um *explizit anzuwählende, vorformulierte Anfragen*. Diese Wahlmöglichkeiten werden vom Benutzer genauso bedient wie die genannten Möglichkeiten zur Informationsauswahl im Datenbestand. Ein Beispiel für solche Interaktionsformen ist im Online-Buchhandel Amazon ein Link auf alle Neuerscheinungen oder die zehn meistverkauften Taschenbücher. Solche Anfragemöglichkeiten sind genauso einfach zu bedienen wie die Standardauswahlmöglichkeiten und basieren auf Antizipation der Benutzerbedürfnisse. Der Nutzer kann wie bisher auch nur aus der Menge der ihm angebotenen Systemfunktionen wählen, d.h. es handelt sich um eine Wiederaufnahme des bereits besprochenen Prinzips der Interaktionsgestaltung auf einer Metaebene. Auch wenn diese Metaanfragen mit einer Personalisierung des Systems einhergehen, sind dem Suchenden dennoch enge Grenzen im Umgang mit dem System gesetzt. Insbesondere bei komplexeren Datenbeständen und stark variierenden Nutzerschichten erscheint eine erfolgreiche Antizipation der Systemanfragen nicht flächendeckend möglich zu sein. Dabei erscheint es nicht nur unrealistisch, für große Datenbestände alle möglichen Benutzerfragen zu antizipieren, sondern auch die Benutzbarkeit wird angesichts der Vielzahl möglicher Anfragen nicht zu gewährleisten sein.

Den Standard bei der Suche in reichhaltigen textuellen, vornehmlich in Datenbanken gespeicherten Beständen stellt derzeit die *Volltext-Stichwortsuche* dar. Sie wird vor allem von hochwertigen Datenbanken angeboten und ermöglicht es dem Benutzer, direkt in den Daten zu recherchieren. Dadurch ist einerseits deutlich weniger Antizipationsaufwand nötig als für die bislang gezeigten Lösungen, weshalb eine Volltextrecherche eigentlich „magerer“ erscheint als ein System mit vorformulierten Anfragen. Andererseits ist Volltextsuche aber aus Sicht der Benutzer flexibler als die anderen bislang genannten Lösungen, weil die Suchfunktionalität zum einen leicht verständlich ist und zum anderen die Suchmöglichkeiten nicht von vornherein beschränkt erscheinen. Daß Benutzer einer Datenbank normalerweise keine tiefergehende Kenntnis des Datenbestandes haben, führt allerdings häufig dazu, daß Volltextsuchanfragen scheitern. Ein typisches Beispiel für diese Art von Systemen ist der OPAC, dessen Eigenschaften – insbesondere die auf pattern matching beruhende volltextbasierte Suche in diesen Systemen – in 2.2 ausführlich beschrieben wurde.

Für typische Online-Datenbanken ist der Antizipationsaufwand im Sinne einer Vorga-

be möglicher Systemanfragen aufgrund der Größe des Datenbestandes und oft auch aufgrund seiner Dynamik nicht mehr zu bewältigen. Allerdings kann für einen kleinen Datenbestand und eine überschaubare Domäne sowie qualitativ hochwertige Antizipation die Qualität eines Systems mit vorformulierten Suchanfragen durchaus die eines Systems mit Freitextsuche übertreffen: Durch die Antizipation ist das Informationsbedürfnis des Benutzers vollständig festgeschrieben, und eine darauf abgestimmte Systemantwort ist garantiert. Im Gegensatz hierzu können isolierte Suchbegriffe nur stringbasiert auf einem Index gesucht werden – über die Besonderheiten dieses Suchverfahrens wurde ausführlich in 2.2 berichtet. Dabei kann über die Intentionen des Benutzers lediglich angenommen werden, daß das bloße Vorhandensein des Suchbegriffes in einem Datensatz den Datensatz als ganzes für den Benutzer relevant macht. NL-Schnittstellen erscheinen in diesem Zusammenhang vorteilhaft, weil sie die Flexibilität einer nicht vollständig antizipierenden Benutzerschnittstelle mit der Ausdrucksmächtigkeit einer dem Benutzer vertrauten Anfragesprache kombinieren. Einerseits ist der Suchende in seiner Wortwahl nicht eingeschränkt und kann ähnlich wie bei der Stichwortsuche das für ihn interessante Wortmaterial verwenden, andererseits ist die Anfragesprache ausdrucksstark genug, um die Benutzerintention sehr viel spezifischer als bei der reinen Stichwortsuche auszudrücken. Als Vorteile von NL-Interfaces im allgemeinen werden die kognitive Entlastung des Suchenden während der Arbeit, die einfache Bedienbarkeit des Systems ohne Training durch alle Anwendergruppen und der höhere Akzeptanzgrad der zugrundeliegenden Datenbank genannt (Krause, 1980; Carbonell et al., 1983; Perrault und Grosz, 1988). Natürliche Sprache als Anfragesprache an einen Datenbestand entlastet den Suchenden zudem von notwendigem Wissen über eventuell vorhandene Datenstrukturen oder verwendetes Vokabular, weil natürliche Sprache an sich domänenunabhängig ist und abstrakt verwendet werden kann. Hinzu kommt, daß Fehler in der Bedienung der Schnittstelle vermindert werden, weil eine dem Nutzer vertraute Interaktionsform praktiziert wird. Dies wiederum ist der Akzeptanz der Schnittstelle förderlich (vgl. auch Bates (1984)). Im Ergebnis sollen Datenbanken durch NL-Schnittstellen häufiger und für die Nutzer komfortabler eingesetzt werden, so daß die Kosten zur Wiedergewinnung von Informationen sinken.

Allerdings ist eine völlig unbeschränkte natürlichsprachliche Anfrage mit den verfügbaren (computer)linguistischen Analysemöglichkeiten derzeit (und auch in absehbarer Zukunft) nicht vollständig im allgemeinen Fall verarbeitbar (s. S. 103). Eine ganze Reihe an Phänomenen (Modalität, Implikaturen, Diskursstrukturen, Einstellungsberichte, Ereignisse u.a.) sind momentan bestenfalls im Rahmen der linguistischen und sprachphilosophischen Forschung verstanden, von einer robusten und integrierten maschinellen Verarbeitung aber noch weit entfernt. Aus diesem Grunde wird neben der Forschung an umfassenderen und effizienteren Analysemethoden auch nach natürlichen Einschränkungen gesucht, die die Verwendung einer Teilmenge der natürlichen Sprache (auch Subsprache genannt) an Schnittstellen ermöglicht. Eine solche Subsprache ist gekennzeichnet durch die Verwendung von definierten Teilen lexikalischer, morphologischer, syntaktischer und semantischer Information – sie soll so in die La-

ge versetzt werden, effizienter als im allgemeinen Fall Eingaben (aus der Domäne der Subsprache) zu analysieren. Außerdem soll sie unter Verwendung eines domänenspezifischen Lexikons für grammatikalisch analysierbare Sätze zwischen solchen, die für die Domäne relevant erscheinen, und anderen, für die Domäne irrelevanten unterscheiden können (Kittredge, 1982).

Ausgehend von der Idee, daß sowohl die Anfrage-Situation des Benutzers als auch die Domäne der in der Datenbank gespeicherten Informationen eine natürliche Beschränkung der verwendeten Sprache darstellen, wird angenommen, daß dem Benutzer diese, die Eingabeverarbeitung erleichternden, Einschränkungen schnell und einfach vermittelbar sind.

In wieweit diese Schnittstellen dann immer noch „natürlichsprachlich“ oder doch besser „eingeschränkt natürlichsprachlich“ heißen müssen, sei dahingestellt, aus Gründen der Einfachheit werden wir weiter von „natürlichsprachlichen Schnittstellen“ oder „NL-Schnittstellen“ sprechen und den Sprachumfang dieser Schnittstellen explizit als begrenzt oder eingeschränkt bezeichnen.

Das in 3.2 ausführlich dargestellte System OSIRIS akzeptiert eine *definierte Teilmenge der natürlichen Sprache* stellt damit die Synthese aus den auf Antizipation bauenden Systemen und denen mit stichwortbasiertem Volltextretrieval dar: Benutzerbedürfnisse werden antizipiert, indem ein unvollständiger Satz vorgegeben wird, der durch den Benutzer zu vervollständigen ist. Dabei werden aber nur abstrakte Ziele antizipiert, denn die genaue Beschreibung des gesuchten Gegenstandes bleibt dem Benutzer durch die Vervollständigung des Satzes überlassen. Zugleich werden aber die Nachteile der Eingabe bloßer Stichwörter vermieden, weil die Eingabe explizit nicht als eine Reihung einzelner Wörter, sondern als entsprechend zu verarbeitende strukturierte natürlichsprachliche Ausdrücke verstanden wird. Die Ausdruckskraft und die Präzision der Beschreibung des gesuchten Gegenstandes ist somit wie in 3.4 ausgeführt sehr viel höher als bei Stichwörtern. Da OSIRIS in der Domäne „Bibliothek“ angesiedelt ist, kann in lexikalischer Hinsicht der mögliche Vorteil einer Subsprache, einen eingeschränkten, domänenspezifischen Wortbestand aufzubauen (z.B. für Wetterberichte oder Börsen Nachrichten) nicht umgesetzt werden.

Neben den Vorteilen auf der Eingabeseite bietet die Möglichkeit der Verarbeitung natürlicher Sprache aber auch auf der Datenseite Vorteile: Wie in 2.2 ausführlich am Beispiel der Online-Bibliothekskataloge gezeigt, stellt das stringbasierte Volltextretrieval für sprachliche Daten kein adäquates Suchverfahren dar. Lexikalische Variation und morphologische Vielfalt waren die wesentlichen Beispiele, die zu diesem Problemkomplex genannt wurden. Durch die computerlinguistische Analyse der natürlichsprachlichen Eingabe kann nachfolgend entsprechend präziser in den Daten gesucht werden, denn es muß nicht mehr nur das bloße Vorhandensein der Suchbegriffe in einem Datensatz geprüft werden, sondern es ist eine höherdimensionale Suche möglich, die syntaktische Dekomposition und semantische Relationen sowie klassifikatorische Merkmale miteinbeziehen kann. OSIRIS ist deshalb eine Alternative für die stichwortbasierten OPAC Systeme, die sich durch den Einsatz computerlinguistischer Methoden

zur Eingabeanalyse und damit auch zur besseren Abbildung der Benutzereingabe auf den Datenbestand auszeichnet.

Eine Variation stellen Systeme dar, die die Eingabe *kontrollierter Sprache* erlauben. Gegenüber Systemen wie OSIRIS ist hier eine Erweiterung des Eingabebereiches bei gleichzeitiger Reduktion der Komplexität umgesetzt worden. Es gibt keine Vorgaben mehr, aus denen zu wählen oder die zu ergänzen wären, sondern es können vollständige natürlichsprachliche Anfragen gestellt werden. Diese müssen allerdings, um eine korrekte Analyse zu gewährleisten, gewissen syntaktischen Regeln gehorchen und sich auf ein domänenspezifisches Vokabular beschränken.

Krause (1980) beschreibt mit USL²⁷ ein System dieses Typus. USL ist eine Datenbank mit Angaben über Noten von Schülern sowie deren Abgang, Versetzung und Schichtzugehörigkeit. USL wurde Ende der siebziger Jahre von H. Lehmann, N. Ott und M. Zoeppritz am Wissenschaftlichen Zentrum der IBM Deutschland in Heidelberg entwickelt und baut zu großen Teilen auf Software auf, die bereits sehr viel früher entstand. USL erlaubt es Benutzern wie z.B. Lehrkräften, eingeschränkt natürlichsprachliche Anfragen zu tätigen, auf die eine Antwort in Form von tabellarisch aufbereiteten Daten erfolgt. Dabei zeigen sich in der von Krause (1980) beschriebenen Evaluation die zu erwartenden Probleme, wenn nämlich Benutzer sich in der Wahl der Präpositionen einschränken müssen: „Wieviele nichtversetzte Schüler gehen *aus/in Sexta ab?“²⁸ Dem Benutzer wird nicht klar, nach welchem Prinzip das System Präpositionen akzeptiert oder nicht. Andere Probleme ergeben sich bei der Einschränkung der Wortstellung – auch hier waren die untersuchten Benutzer nicht in der Lage, Regeln im Systemverhalten zu erkennen. Vom System zurückgewiesen wurden auch Versuche, mehr als ein Objekt erfragen zu wollen („Wieviele nichtversetzte Schüler gehen in welcher Klassenstufe ab?“).

Vermutlich, um die Ausdrucksmächtigkeit der Anfragesprache zu erhalten, ist es dem Benutzer möglich, Eigenschaften und Relationen im Systemdialog zu definieren und in späteren (syntaktisch einfachen) Anfragen zu benutzen. So gibt im von Krause angeführten Beispieldialog der Benutzer zunächst „nv=nichtversetzte Schüler“ ein, um anschließend dann „Wieviele nv gehen in Sexta ab?“ zu fragen. Dieser Mechanismus sprengt spätestens dann die gängige Vorstellung von natürlicher Sprache, wenn der Benutzer mit „sex2h=Differenz(sex2,sex1)“ sich eine Funktion definiert, die er in seiner folgenden Eingabe wieder aufruft: „Liste Durchschnitt der Notenschnitte von sex2h in Sexta.“

Der Vorteil kontrollierter Spracheingabe liegt in der freien Ausdrucksmöglichkeit des Benutzers, der sein Informationsbedürfnis ohne Systemvorgaben auf die ihm genehme Art und Weise formulieren kann. Als Nachteile kontrollierter Sprachen werden ein oft hoher Schulungsaufwand, aufwendige Etablierung und schwierige oder gar unmög-

²⁷User Speciality Language

²⁸Die Verwendung der zusammengescriebenen Form „nichtversetzte“ könnte auf generelle Einschränkungen des Systems in Hinblick auf Negation in Anfragen hinweisen. Krause geht auf diesen Punkt nicht ein.

liche Übertragbarkeit der Lösungen sowie mögliche Abwehrreaktionen der Benutzer genannt.

Der Schulungsaufwand entsteht, weil auch gut konstruierte kontrollierte Sprachen für die Produktion meist nicht über ausreichend selbsterklärende Regeln verfügen (Copestake und Sparck-Jones, 1990). Dies deckt sich mit dem Ergebnis der Evaluation von USL, in der Krause (1980) feststellt: „Der Benutzer hatte große Schwierigkeiten, wenn er durch USL gezwungen war, seine Wortstellungsmöglichkeiten einzuschränken.“ Selbiges gilt laut Krause auch für die eingeschränkte Wahl der Präpositionen. Krause (1980) berichtet weiter über den Gebrauch von Paraphrasierungen, „keine Oberflächenstruktur wurde ausschließlich verwendet, wenn dazu noch eine Variante zur Verfügung stand [...], obwohl der Benutzer wußte, daß sich in bezug auf die zu erwartende Antwort nicht der kleinste Unterschied ergeben konnte“ (Krause, 1980, :217). Krauses Beobachtung, daß Benutzer ihre Anfragen variieren, läßt sich auch ganz anders erklären: Den Fragenden war entweder nicht klar, daß sich „in bezug auf die zu erwartende Antwort nicht der kleinste Unterschied ergeben konnte“ oder aber die Fragenden mochten dies nicht glauben und wollten sich davon selbst überzeugen. Diese Erklärung paßt stärker zu den Beobachtungen zum Verhalten von OPAC-Nutzern, über die in 2.2 berichtet wurde.

Die Schwierigkeiten bei der Erstellung resultieren laut Kittredge (1982) aus dem hohen Aufwand, mit Hilfe von Korpora die natürlichen Grenzen einer Subsprache zu bestimmen. Dabei müssen linguistische Phänomene aus der in Frage stehenden Subsprache oft anders als im allgemeinen Fall behandelt werden, was die Übertragbarkeit der für eine Subsprache gefundenen Lösungen auf weitere Subsprachen in Frage stellt.

Die Abwehrreaktionen der Benutzer entstehen bei der Rezeption, wenn der Text als „Kindersprache“ empfunden werde.²⁹ Eine kontrollierte Sprache müsse so gut sein, daß Rezipienten nicht bemerken, daß es sich bei einem ihnen vorliegenden Text um kontrolliertes Deutsch handle. Für die Einsatzzwecke der technischen Dokumentation kann dies gewährleistet werden, wenn z.B. komplexe Nebensatzstrukturen im Deutschen vermieden werden, um später kostensparende automatische Übersetzungssysteme einsetzen zu können. So ist kontrollierte Sprache im anglo-amerikanischen Raum seit langer Zeit zur Dokumentenerstellung z.B. in der Luftfahrtbranche in verschiedenen Ausprägungen in Gebrauch. Die komplizierten Erstellungsregeln begrenzen den Einsatz dabei aber praktisch ausschließlich auf die Produktion von Texten. Kontrolliertes Deutsch scheidet als Anfragesprache für Information Retrieval-Systeme unter diesen Umständen aber natürlich aus. Zudem ist unklar, ob ein Ausschnitt der natürlichen Sprache so generell definiert werden kann, daß seine Grammatik auf andere Anwendungen übertragbar ist (Krause, 1980).

Eine interessante Variante stellen Systeme dar, die natürlichsprachliche Sätze als Systemanfragen antizipieren und dem Benutzer in einem Auswahlmenü zur Verfügung stellen. Was in der simpelsten Ausführung noch eine bloße Variation des oben

²⁹U. Block in einem eingeladenen Vortrag auf der Jahrestagung der Gesellschaft für Informatik (GI'98) in Magdeburg, 1998.

gezeigten Ansatzes vorformulierter Anfragen ist („Zeige mir die zehn meistverkauften Sachbücher der letzten Woche“ als mögliche Systemanfrage versus „Klicken Sie hier, um die zehn meistverkauften Sachbücher der letzten Woche zu sehen“ als Beschriftung eines Knopfes), wird in manchen Systemen durch eine Kombinationsmöglichkeit auf der Ebene von Teilphrasen deutlich erweitert. Tennant et al. (1983b,a) beschreiben ein System, in dem der Suchende in einem mehrfach unterteilten Fenster zunächst aus den drei verfügbaren Interaktionsarten „find“, „delete“ und „insert“ wählen kann, die dann als erstes Wort des entstehenden Anfragesatzes verwendet werden. Anschließend können aus weiteren, je nach Kontext aktivierten oder deaktivierten Fenstern weitere Teile selektiert werden: Aus dem Fenster *attributes* z.B. „color“, gefolgt von „and“ aus dem Fenster *connectors*, gefolgt von „name“ wiederum aus *attributes*, „of“ aus *connectors* und „parts“ aus *nouns*. Der Satz kann mit einer Phrase aus dem *modifiers* Fenster fortgesetzt werden („whose color is“). Der Suchende erreicht eine Metaebene, wenn die nun einzufügende konkrete Farbe bestimmt werden soll: Da aus verständlichen Gründen kein Fenster *color* existiert und auch unter *attributes* keine vollständige Liste aller Farbausdrücke gegeben werden kann, gibt es ein Fenster (*experts*) mit Ausdrücken wie „<specific colors>“, die nach Anwahl in einem separaten Fenster die der Datenbank bekannten Werte des Attributs „color“ bereitstellen. Auf diese Weise erzeugt der Suchende Anfragen wie „Find color and name of parts whose color is green or blue and which are supplied by suppliers whose supplier status is greater than or equal to 25“. Dieses System verwendet eine intuitiv verständliche Anfragesprache und erlaubt mit einer einfachen Konstruktionsmöglichkeit und einem kleinen Inventar an Ausdrücken die Bildung abstrakter Anfragen. Tippfehler gibt es genauso wenig wie Anfragen, die die Domäne des Systems verlassen. Aufwendige Grammatiken und große Lexika sind genauso überflüssig wie große Rechenleistung. Nachteilig ist aber natürlich die umständliche Art der Bedienung (die auf ihre Benutzbarkeit hin untersucht werden müßte). Obwohl Benutzeranfragen nicht als vollständiger Ausdruck antizipiert werden und ggfs. auf einer Metaebene der in der Datenbank aktuell verfügbare Wertebereich eines Attributes berücksichtigt wird, erscheint es für große und insbesondere für dynamische oder heterogene Datenbestände unrealistisch, alle Benutzeranfragen (strukturell) zu antizipieren.

Cohen et al. (1989) berichten von einem Versuch, durch Kombination mit graphischer Interaktion die Komplexität von NL-Eingaben so weit zu reduzieren, daß sie als Anfragemöglichkeit handhabbar werden. Zu den von Cohen et al. (1989) vorgeschlagenen Maßnahmen gehören u.a. Begrenzungen in Form separater Fenster und farbiges Hinterlegen verschiedener Ausdrücke, um Referenzambiguitäten im Diskurs zu vermeiden. Zudem wird eine baumartige Repräsentation des Diskurses erzeugt, in der der Benutzer navigieren kann, um an frühere Diskursthemen anzuknüpfen. Weiterhin werden Formblätter vorgeschlagen, in denen für bestimmte Handlungen (z.B. „move“) die nötigen (thematischen) Rollen zum einfachen natürlichsprachlichen Befüllen vorgesehen sind. Alles in allem erscheint die Kombination aus NL-Eingabe und direkter Manipulation sinnvoll. Allerdings wird die Schnittstelle durch die Vermischung der

Modi deutlich komplexer, und es stellt sich wie schon für Tennant et al. (1983b,a) die Frage, ob die Mittel, die zur Reduktion des Analyseaufwandes der natürlichsprachlichen Eingabe nötig sind, wirklich so flexibel und domänenunabhängig sein können, daß sie einfach zu erweitern und zu übertragen sind.

Das älteste, aber auch heute noch unerreichte Konzept einer natürlichsprachlichen Schnittstelle sieht die Möglichkeit vor, völlig unbeschränkt Eingaben in natürlicher Sprache tätigen zu können. Die Vorteile liegen angesichts der zuvor diskutierten, oftmals schwierig zu vermittelnden Einschränkungen anderer NL-Schnittstellen auf der Hand: Die volle Ausdruckskraft der natürlichen Sprache ermöglicht es dem Benutzer, tatsächlich so „naiv“ und ohne Training mit der Maschine zu kommunizieren wie mit einem Menschen.

Jedoch ist es genau diese Metapher „wie mit einem Menschen“ zu kommunizieren, die auch die Probleme dieses Ansatzes verdeutlicht. Zwischenmenschliche Kommunikation läßt sich nicht isoliert sprachlich betrachten, sondern ist vielfältig mit persönlicher Erfahrung, Vorstellungen über die Struktur der menschlichen Umwelt und sozialen Normen verbunden: „In particular, to use natural language fluently, a system must understand how the communication process itself is affected by language user's goals, plans, and beliefs“ (Hendrix, 1982, :61). Menschen haben aufgrund ihrer persönlichen Erfahrung Vorannahmen, die (auch) ihr sprachliches Verhalten bestimmen, indem sie z.B. Äußerungen tätigen, die nicht unmittelbar aus der sprachlichen Interaktion abzuleiten sind (Charniak, 1986). Menschen interagieren mit ihrer Umwelt nicht nur sprachlich, sondern aktiv auch durch Mimik und Gestik und passiv (unterschiedlich stark) mit den anderen verfügbaren Sinnen.

Diese vielfältigen Interaktionsformen sind entweder unterschwellig bei der Mensch-Maschine-Kommunikation beteiligt, wenn die Metapher „wie mit einem Menschen“ zu kommunizieren, ernstgenommen wird (vgl. 7.2), oder aber sie führen zu einem Bruch, wenn die Mensch-Maschine-Kommunikation nicht so natürlich und umfassend erlebt wird wie die zwischenmenschliche Kommunikation.

Doch auch ohne den „naiven“ Benutzer in Betracht ziehen zu müssen, bleiben Schnittstellen, die uneingeschränkte natürlichsprachliche Eingaben verarbeiten können, eine flexible Möglichkeit, um die Recherche in Datenbeständen einer beliebigen Domäne zu ermöglichen. Der Nutzer soll sich weder mit einer Anfragesprache vertraut machen müssen, noch soll er sich an unterschiedlichen Schnittstellen jeweils neu auf das System und den Datenbestand einstellen müssen. Das betrifft nicht nur den Wechsel zwischen einzelnen Schnittstellen, sondern auch die Unterbrechung der Nutzung durch größere Zeitabstände. Gerade für Gelegenheitsnutzer erscheinen natürlichsprachliche Schnittstellen ideal. In einigen Situationen kann natürliche Sprache als Anfragemöglichkeit aber auch für erfahrene Nutzer von Vorteil sein: Einige Phänomene wie Negation oder Quantifikation lassen sich natürlichsprachlich sehr viel kompakter und einfacher formulieren als in einer Anfragesprache wie z.B. SQL. Zudem können Folgefragen gestellt werden, in denen anaphorisch auf zurückliegende Äußerungen verwiesen werden kann.

Diesen Vorteilen stehen jedoch massive Probleme gegenüber, einerseits solche, die das theoretische Verständnis natürlichsprachlicher Phänomene betreffen, andererseits grundsätzliche Verarbeitungsprobleme (vgl. 4.2) und schließlich massive Performanzprobleme bei der Bearbeitung eines (mit heutigen Methoden) enorm großen Suchraumes möglicher Analysen. Der erste Problembereich führt dazu, daß eine ganze Reihe an Phänomenen (z.B. im Diskurs, im Zusammenhang mit sprachbegleitender Gestik oder in Hinblick auf Pragmatik) nicht zufriedenstellend von der Maschine behandelt werden und so die Idee der völlig uneingeschränkten natürlichsprachlichen Schnittstelle nicht gehalten werden kann.

Allerdings besteht hier die Hoffnung, daß auf die in Frage stehenden Phänomene für Standardanwendungen verzichtet werden kann und die Einschränkungen der Schnittstelle hinsichtlich z.B. der Verarbeitung von Gestik sich durch andere Mittel wie z.B. die Gestaltung der Schnittstelle vermitteln lassen. Dabei könnte es sich als nützlich erweisen, systematisch der Verwechslung der Mensch-Maschine-Kommunikation mit der zwischenmenschlichen Kommunikation entgegenzutreten (vgl. 7.2). Außerdem folgt aus den genannten prinzipiellen Vorteilen natürlichsprachlicher Schnittstellen nicht, daß diese auch für alle Einsatzgebiete und jede Situation optimal geeignet sind. Es ist vielmehr vorstellbar und wünschenswert, mit herkömmlichen Schnittstellen an solchen Stellen zu arbeiten, an denen entweder natürliche Sprache als Anfragesprache spezifische Nachteile hat, oder aber andere Interaktionsmöglichkeiten mit deutlichen Vorteilen bekannt sind. Aufgaben, zu deren Bearbeitung überwiegend Zahlen oder Formeln gebraucht werden, sind vermutlich besser mit einer strukturierten, graphischen Oberfläche zu bearbeiten, als mit natürlicher Sprache. Auch eine sehr eingeschränkte Systemdomäne wie die einer Preisagentur braucht vermutlich nicht die voll Flexibilität natürlicher Sprache, um korrekte Suchanfragen ausdrücken zu können.

Schwieriger zu handhaben erscheinen diejenigen der oben genannten Probleme, die auf grundsätzlichen Verarbeitungsschwierigkeiten beruhen. So ist es durchaus möglich, die Speicherung von Wissen über die Welt in Form einer (strukturierten) Knowledge-Base zu modellieren, zusammen mit Prozessen, die auf diesen Datenstrukturen Inferenzen ausführen (vgl. z.B. Klose et al. (1991); Herzog und Rollinger (1992)). Allerdings ist derzeit vollkommen unklar, wie dies über Beispieldomänen (die sogenannten „Toy Worlds“) hinaus einheitlich für wirklich relevante Ausschnitte der Welt geschehen soll. Weder kann diese Arbeit flächendeckend per Hand durchgeführt werden, noch ist angesichts der sehr unterschiedlichen Inferenzen z.B. zur Lage von Objekten in Zeit und Raum oder zu den Folgen bestimmter Ereignisse³⁰ klar, wie eine einheitliche Datenstruktur, die dies unterstützt, aussehen soll. Auch wird die Erweiterbarkeit der detailliert modellierten Beispielwelten grundsätzlich in Frage gestellt – eine Kritik, die in der Vergangenheit an der gesamten Forschung der Künstlichen Intelligenz geübt wurde (vgl. Dreyfus (1979)). Und schließlich stellt sich auch für eine (hypothetisch) vollständige Wissensbasis die Frage, wie die Dynamik einer sich stän-

³⁰Vgl. Tversky und Kahneman (1982) über menschliche Heuristiken zur Einschätzung von Wahrscheinlichkeiten und zur Korelation von Ereignissen.

dig ändernden Umwelt abgebildet werden kann: Eine Quelle mit Informationen über den Zustand der Welt und Methoden, mit Hilfe dieser Informationen im Diskurs mit dem Benutzer zu bestehen, nützen nur dann etwas, wenn das System, das im Dialog mit dem Benutzer die Welt verändert, sowohl auf diese wie auch auf externe Veränderungen entsprechend reagieren kann (vgl. hierzu die alte Diskussion in der Künstlichen Intelligenz zum Frameproblem, z.B. Hayes (1981); McCarthy und Hayes (1969) und die hierfür vorgeschlagenen Lösungsmöglichkeiten McCarthy (1979a); Hanks und McDermott (1986); Shoham und McDermott (1988)).

Zusätzliche Schwierigkeiten treten auf, wenn wie in VERBMOBIL (Wahlster, 1993), statt der Tastatureingabe die Verarbeitung gesprochener Sprache berücksichtigt werden soll. Tastatureingaben sind als Ausgangspunkt für die Analyse eindeutiger als die Ergebnisse eines Spracherkenners – zudem läßt sich ein vorhandenes System nicht einfach durch Vorschalten eines Spracherkenners zu einem sprachverstehenden System erweitern: Die einzelnen Analysekomponenten müssen für die neue Aufgabenstellung auch neu aufeinander abgestimmt werden. Je „natürlicher“ die Benutzer gesprochene Eingaben tätigen können, desto größer wird auch der Anteil der (vielfach agrammatischen) Spontansprache, für die andere Verarbeitungsmethoden erforderlich sind als für Tastatureingaben.

Vor diesem Hintergrund ist es nicht verwunderlich, daß es kaum Systeme gibt, die wirklich völlig unbeschränkte natürlichsprachliche Eingabe zulassen.

Forschungsprototypen wie HAM-RPM (Jameson et al., 1980; Hahn et al., 1980), LILOG (Geurts, 1990a; Herzog und Rollinger, 1992) oder PARNAX (Comino et al., 1983) können für den gewählten Weltausschnitt (Hotelzimmerbuchung, Sehenswürdigkeiten in der Düsseldorfer Altstadt, Mitarbeiter einer Firma) natürlichsprachliche Anfragen verarbeiten, beantworten und dabei eine ganze Reihe komplexer linguistischer Phänomene wie z.B. anaphorische Referenz und Raumsemantik behandeln (vgl. Geurts (1990a); Habel (1988)). Trotz der beeindruckenden Leistungen kommen diese Systeme für den Praxiseinsatz nicht in Betracht,³¹ da die sprachlichen und konzeptuellen Grenzen dieser und vergleichbarer Systeme dem Benutzer nur sehr schwer zu vermitteln wären. Hierzu wäre nicht nur eine substantielle Ausweitung der ohnehin bereits aufwendigen Domänenmodellierung zwingend erforderlich, sondern zudem auch die Etablierung einer für naive Benutzer deutlichen Grenze in behandelbare und nicht behandelbare Anfragen.³² Allerdings ist es alles anderes als klar, wie eine solche Grenze einmal im Phänomenbereich, dann hinsichtlich der linguistischen Konstruktionen und schließlich auch in Richtung eines Metadialoges gezogen werden könnte. Auch ist die eingeschränkte Übertragbarkeit der für die Modellierung des Weltausschnittes verwendeten Methoden auf größere Domänen, wie bereits erwähnt, ein wesentlicher Kritikpunkt.

Menke et al. (1998) führen in Hinblick auf die Möglichkeit, nicht nur Satzteile, son-

³¹ Als Forschungsprototypen sind sie auch nicht dafür gedacht.

³² Wenn man denn den Traum einer universellen natürlichsprachlichen Schnittstelle zumindest vorübergehend zurückstellt.

dern auch vollständige natürlichsprachliche Sätze verarbeiten zu können, zwar aus, daß „this approach gives the user the full expressiveness of his native language“, aber wie dies wirklich flächendeckend durchgeführt werden könnte, bleibt offen. Dieses Zitat ist wohl eher als eine generelle Aussage gedacht, daß der Suchende sein Informationsbedürfnis nicht in Bruchstücken, sondern in ganzen Sätzen auszudrücken vermag. Das hat mit der vollständigen Ausdrucksmächtigkeit der natürlichen Sprache, wie sie in diesem Abschnitt zur Diskussion steht, aber wenig zu tun. Dabei wäre das Anwendungsbeispiel, eine natürlichsprachliche Schnittstelle für die im Rahmen des MeDoc Projektes entstehende digitale Bibliothek (Helbig et al., 1997; Endres und Fuhr, 1998), geeignet, anhand des Retrievals in elektronischen Volltexten aus sehr unterschiedlichen Domänen die Vorteile unbeschränkten NL-Zugriffs zu demonstrieren.

Die einzige Schnittstelle, die ernsthaft den Anspruch auf unbeschränkten natürlichsprachlichen Input erheben kann, ist VERBMOBIL. In der ersten Projektphase von 1993 bis 1996 wurde für die Domäne „Terminvereinbarung zwischen Geschäftsleuten“ ein Prototyp entwickelt, der der Spontansprache zweier Gesprächsteilnehmer in annähernd Echtzeit folgen können soll und auf Anforderung der Gesprächsteilnehmer Teile übersetzt. Dabei sieht das Szenario vor, daß die beiden Beteiligten (ein Deutscher und ein Japaner) gemeinsam auf Englisch kommunizieren und im Falle eines Kommunikationsproblems gezielt auf Knopfdruck eine Übersetzung aus ihrer Muttersprache in die Interlingua Englisch anfordern können. Diese Übersetzung wurde für rund 2.500 deutsche und nicht ganz 400 japanische Wörter umgesetzt, wobei die eigentlich anspruchsvolle Aufgabe des Systems darin besteht, für eine adäquate Übersetzung dem Gespräch die ganze Zeit über zu folgen und so nicht einfach eine wörterbuchbasierte Übersetzung vorzunehmen, sondern eine dem Kontext angepaßte. In einer weiteren, bis zum Jahr 2000 angelegten Projektphase sollte die Übertragung des Systems auf eine neue Domäne („Verhandlungsdiallog zwischen Reisebüromitarbeiter und Kunde“) erfolgen. Am Ende dieser Projektphase soll das System einen Wortschatz von 10.000 deutschen und 2.500 japanischen Wörtern haben.

VERBMOBIL kann, obwohl nur für eine sehr restringierte Domäne konzipiert, lediglich Teilen der (thematisch stark eingeschränkten) Unterhaltung folgen. Der große Fortschritt besteht aber in der Robustheit des Systems: Anstatt zu versuchen, das gesamte Gespräch zu verstehen, ignoriert VERBMOBIL nicht verarbeitbare Teile und versucht, auf Grundlage der übrigen Teile zu einer konsistenten Interpretation des Dialoges zu gelangen.

Der zeitliche Rahmen, die Förderungshöhe (mehrere hundert Millionen Mark) und die Ähnlichkeit der Szenarien in Phase 1 und 2 (klare Rollenverteilung, Vereinbarung eines Datums und eines Ortes für ein Unterfangen, dessen Zweck und Ziel bereits definiert ist) machen allerdings die Schwierigkeiten deutlich, die bei der Entwicklung einer unbeschränkten und robusten Schnittstelle zu bewältigen sind.

4.1.1 Fazit

Von den oben beschriebenen Schnittstellen werden die rein passiven (S. 96), die mit Auswahlmöglichkeit (S. 96), die personalisierbaren (S. 96) und die auf Stichwortsuche im Volltext basierenden Schnittstellen (S. 97) heute in der Praxis eingesetzt. Dabei handelt es sich um keine echte Rangfolge, denn die Verwendung jeder einzelnen Dialogschnittstelle kann je nach Einsatzgebiet durchaus sinnvoll sein. Schnittstellen mit vorformulierten Anfragen (S. 97) erscheinen in der Praxisanwendung schwierig wegen des hohen Antizipationsaufwandes, der (wenn überhaupt sinnvoll zu bewältigen) sehr teuer ist. Ebenfalls ungeeignet für den Einsatz in Online-Schnittstellen zu Datenbanken erscheinen die kontrollierten Sprachen (S. 100). Der Bedarf an einfach zu bedienenden Schnittstellen für Online-Systeme nimmt stetig zu (vgl. 3.5), und vor dem Hintergrund der Betrachtungen in diesem Abschnitt sind solche Schnittstellen, die eine definierte Teilmenge der natürlichen Sprache (S. 98) oder freie Eingaben (S. 103) akzeptieren, prinzipiell geeignet, um in diesem Bereich eingesetzt zu werden. Diese Systeme (gemeinsam mit denen, die kontrollierte Sprache einsetzen) verfügen gegenüber den anderen betrachteten Möglichkeiten auch zumeist über eine interne Repräsentation der Benutzereingabe. Dies macht es möglich, die Eingabemöglichkeiten systematisch zu erweitern und die Schnittstelle auf neue Anwendungsgebiete zu übertragen – ohne interne Repräsentation erscheint dies für die anderen genannten Schnittstellen nur schwer oder ad hoc möglich.

Bemühungen, vollkommen neue Anfragesprachen zu entwerfen, erscheinen angesichts der schlechten Erfahrungen mit künstlichen Anfragesprachen (siehe 2.1.2 und 2.2) wenig erfolgversprechend. Auch ambitionierte Projekte, die auf Mengentheorie setzen (Navarro und Baeza-Yates, 1995, 1997) oder mit Hilfe von Diagrammen versuchen, Anfragen zu stellen (Catarci et al., 1993), erscheinen höchstens für spezielle Einsatzgebiete und Nutzer mit besonderen Bedürfnissen oder Vorkenntnissen geeignet.

Es erscheint daher auch vor diesem Hintergrund der richtige Weg, die beiden eben beschriebenen, für eine definierte Teilmenge der natürlichen Sprache bzw. für freie Eingaben geeigneten Schnittstellentypen näher zu untersuchen. Natürlich existieren diese beiden Schnittstellentypen nicht in reiner Form, sondern es gibt für beide Arten eine enorme Bandbreite möglicher Ausprägungen, die, unabhängig vom Einsatzgebiet, nur in Bezug auf Endbenutzervorteile zu sehen sind. Diese Vorteile gilt es im Folgenden auszuloten, wobei insbesondere die bislang nicht betrachtete Benutzerinteraktion von besonderem Interesse sein wird.

Ausgehend von der Funktionalität des OSIRIS-Systems als Vertreter der Dialogschnittstellen mit eingeschränkt natürlichsprachlicher Eingabemöglichkeit wird der weitere Ausbau des Systems in Richtung auf eine unbeschränkte Eingabe sowie in Richtung auf eine ausgeprägtere Benutzerinteraktivität betrachtet. Dabei werden die Ergebnisse aus 2.3 und 3.5 berücksichtigt. Ein Ausbau der Benutzerinteraktivität ermöglicht es dem System einerseits, mit dem Benutzer in einen echten Diskurs zu treten, d.h. auf der Eingabeseite sind zusätzliche linguistische Phänomene wie Anaphern oder El-

lipsen zu erwarten. Andererseits kann durch Benutzerinteraktivität sehr viel gezielter Feedback gegeben werden, ein bislang in OSIRIS noch nicht ausreichend genutztes Mittel, um Nutzerverhalten zu steuern und den Suchprozeß zu unterstützen. Feedback, dessen Bedeutung auch in 6.2 thematisiert wird, kann offenbar nicht nur über die Verbesserung des Suchprozesses, sondern auch einfach als Systemeigenschaft an sich die Benutzerzufriedenheit verbessern:

„The basic issues are whether the program can interpret the query the way the user intended it and – even if it does – whether the user knows the program interprets it that way.“ Shneiderman et al. (1998, :97)

Weiterhin werden als unterstützende Maßnahmen in 5.2 Verbesserungen bei der Interfacegestaltung, der Benutzerführung und bei der Struktur der verfügbaren Datenbestände angesprochen.

4.2 Linguistische Phänomene

Im Folgenden wird es um eine Reihe natürlichsprachlicher Phänomene gehen, die für OSIRIS (als Vertreter der in 4.1 beschriebenen Klasse von Dialogschnittstellen mit eingeschränkt natürlichsprachlicher Eingabemöglichkeit) derzeit nur zum Teil oder gar nicht behandelbar sind. Dabei wird keine Vollständigkeit angestrebt, sondern es werden die Phänomene genannt, die bei einer möglichen Erweiterung eines Systems wie OSIRIS an der Schnittstelle zusätzlich zu erwarten sind. Für einige Bereiche werden Bemerkungen zum Stand der Forschung gemacht.

Ausgangspunkt der Betrachtungen sind Phänomene des Diskurses: In dem Augenblick, in dem Systeme beginnen, explizite sprachliche Rückmeldungen an den Benutzer zu geben (also nicht nur mit einer graphischen Anzeige zu reagieren), kommen rückverweisende Ausdrücke und damit äußerungsübergreifende Bezüge ins Spiel, die z.T. mit pragmatischen Phänomenen interagieren.

4.2.1 Diskursstrukturen

Die Diskursforschung beschäftigt sich mit zwei Diskursformen, Text und (gesprochenem wie geschriebenem) Dialog. Beide Diskursarten weisen große Überschneidungen auf: Sie bestehen aus Segmenten, enthalten querverweisende Ausdrücke, sind kohärent und haben übergeordnete Strukturen. Text wird im Zusammenhang mit Dokumentenanalyse und Dokumentenretrieval als Diskursphänomen betrachtet, während für die Untersuchung von Dialogschnittstellen der Dialog (gesprochen und geschrieben) im Zentrum des Interesses steht. Diese Unterscheidung resultiert aus den unterschiedlichen Anforderungen des Einsatzkontextes „Mensch-Maschine-Schnittstelle“ und „Textanalysekomponente“ (Grosz, 1996). Der Dialog als Diskursform wird im Folgenden daher im Mittelpunkt der Betrachtungen stehen und wenn nicht ausdrücklich anders beschrieben, werden „Diskurs“ und „Dialog“ synonym verwendet werden. Ein Dialog als Beispiel eines Diskurses besteht typischerweise aus mehreren Diskurssegmenten, die wiederum aus mehreren Äußerungen verschiedener Dialogpartner bestehen, wobei sowohl der Dialog insgesamt, als auch die einzelnen Segmente eine Struktur haben, die mehr ist als nur die Kombination ihrer Bestandteile. So verfügt der Dialog über eine *Gesamtstruktur*, die neben der eigentlichen Behandlung des Diskursthemas auch einen Eröffnungsteil und einen Schlußteil beinhaltet. Dieser offensichtlichen Diskurssegmentierung steht die Tatsache gegenüber, daß ein Diskurs einen inneren Zusammenhalt, die Diskurskohärenz, aufweist, von der zu klären ist, wie sie entsteht.

Ein weiteres Dialogphänomen ist die *verteilte Äußerungsinitiative*, d.h. an bestimmten Punkten des Dialoges kann die Kontrolle von einem Partner auf den anderen übergehen (siehe 4.2.10). Weiterhin gehören *Überbeantwortung*, d.h. das Angebot von nicht speziell erfragter, aber aus dem Zusammenhang als für den Dialogpartner informativ erkannter Information (Wahlster et al., 1983; Owen, 1986), *kontextuelle Interpretation*, das ist der explizite Bezug auf kontextuelle Information bei der Auswertung von z.B.

Anaphern (siehe 4.2.2), Ellipsen (siehe 4.2.3), Deixis (siehe 4.2.8) oder ambigen Äußerungen (siehe 4.2.3, 4.2.15, 4.2.13) und *interaktive Fehlererkennung und -behebung* zu den Merkmalen des Diskurses.

Theorien über Diskursstrukturen haben durchaus unterschiedliche Ansprüche wenn sie versuchen, einen Diskurs zwischen einem menschlichen Benutzer und einem Computer oder aber einen zwischenmenschlichen Diskurs zu modellieren. In der Kommunikation zwischen Mensch und Maschine müssen Diskursphänomene nicht zwangsläufig im gleichen Umfang wie zwischen Menschen auftreten, da Menschen gegenüber einer Maschine offenbar in vielen Fällen ihre sprachlichen Eingaben simplifizieren (siehe 6.2). Dennoch werden viele dieser Phänomene (wenn auch vielleicht seltener) in der Mensch-Maschine-Kommunikation verwendet, so daß die Übertragung eines (ev. reduzierten) Ansatzes zur Modellierung des Mensch-Mensch Dialoges auf den Mensch-Maschine Dialog gerechtfertigt erscheint.

Ausgehend von der Idee, daß Dialoge eine wiederkehrende Struktur haben (Frage/Antwort-Paare oder Vorschlag gefolgt von Zustimmung oder Ablehnung), wird in vielen Systemen versucht, einen Dialog mit dem Benutzer mit Hilfe einer Dialoggrammatik zu modellieren. Grundbestandteil dieser Modellierung ist dabei der Sprechakt, den der Sprecher mit seiner Äußerung umsetzen möchte. Problematisch ist dabei die automatische Erkennung des vom menschlichen Sprecher intendierten Sprechaktes durch die Maschine.

Einfache Ansätze im Bereich der Dialoggrammatiken sehen ein meist auf endlichen Automaten beruhendes Modell eines Dialoges vor, in dem jeder Knoten eine endliche Liste möglicher Nutzeraktionen darstellt. Für diese Ansätze sind solche Aufgaben geeignet, deren Struktur so klar gegliedert ist, daß für sie eindeutige Zuordnungen zu Knoten eines endlichen Automaten möglich sind. Weiterhin ist der Verzicht auf Metadialoge, sowie die Gesprächsinitiative auf Seiten des Systems notwendig, d.h. dem Nutzer werden explizit die möglichen Aktionen zur Auswahl gestellt. Auf diese Weise werden schwierige Probleme wie turn taking (siehe 4.2.10), Einschübe (siehe 4.2.11) u.ä. einfach umgangen.

Nachteile eines solchen Vorgehens sind die mangelnde Flexibilität und die fehlende Übertragbarkeit auf andere Domänen. Um die genannten Nachteile zu umgehen, können generelle Dialogakte verwendet werden, in denen Leerstellen mit dem konkreten Thema des Dialogabschnittes gefüllt werden (Hulstijn et al., 1996).

Andere Ansätze im Bereich der Dialoggrammatiken sehen „Dialogbibliotheken“ oder Sammlungen häufig verwendeter Dialogakte vor (Novick und Sutton, 1996), wobei die Übergänge zwischen den einzelnen Knoten für jede Anwendung neu bestimmt werden müssen.

Elaboriertere Diskurstheorien, die sich eher mit dem (zwischen)menschlichen Dialog beschäftigen, versuchen die Ziele und Pläne hinter den konkreten Äußerungen der Diskursteilnehmer mit einzubeziehen. So hat ein Sprecher in der Regel Absichten, die über die aktuelle Äußerung hinaus wirken und z.B. auf eine Meinungsänderung beim Hörer hinwirken sollen. Ein solches Ziel kann sogar über verschiedene Dialoge hinweg

konstant bleiben. Die derzeit bekanntesten Theorien dieser Art sind Hobbs (1985), Grosz und Sidner (1986), Mann und Thompson (1988) und Litman und Allen (1987). Zu einem Vergleich verschiedener Theorien siehe Grieszl (1995). Bemerkenswert ist auch der Ansatz von Wahlster et al. (1993) die versuchen, den Phänomenbereich auf andere Objekte im multimodalen Dialog zu erweitern (s. a. 4.2.2).

Grosz und Sidner (1986) stellen eine nach wie vor vielbeachtete Theorie der Diskursstruktur vor, die drei Aspekte als zentral für die Beschreibung eines Diskurses annimmt: die Struktur der Äußerungssequenz (*linguistic structure*), die Absichten des Sprechers (*intentional structure*) und den Fokus der Aufmerksamkeit des Sprechers (*attentional state*).

- Die *linguistic structure* beschreibt auf der Ebene der Äußerungen die einzelnen Diskurssegmente. Jede Äußerung erfüllt einen Zweck innerhalb des Segmentes, in dem sie auftritt und trägt zum Zweck des Segmentes bei, den dieses gegenüber höheren Strukturen, in die es eingebettet ist, innehat. Damit findet die Segmentierung des Diskurses in Abhängigkeit von den identifizierten Zielen ab. Die *linguistic structure* zeigt die Segmente des Diskurses sowie die Relationen, die zwischen eingebetteten und einbettenden Elementen bestehen.

- Die *intentional structure* behandelt die Ziele des Sprechers und zwar sowohl die, die innerhalb eines in der *linguistic structure* gelisteten Segmentes bestehen, als auch die, die zwischen den Segmenten, also in den Relationen ausgedrückt werden.

Die *intentional structure* liefert mit der Repräsentation der Sprecherintentionen die Erklärung für die aktuelle Struktur des Diskurses und die Kohärenz innerhalb des Diskurses. Da für Grosz und Sidner (1986) jeder Diskurs zielgerichtet ist, verfügt jeder Diskurs über ein in der *intentional structure* repräsentiertes Gesamtziel (*discourse purpose*, DP). Mit Hilfe des DP lassen sich die Grenzen des Diskurses innerhalb eines Komplexes mehrerer Diskurse bestimmen. Dabei widerspricht die Annahme nur eines DP nicht der Überzeugung, daß Diskursteilnehmer durchaus mehrere, verschiedene Ziele verfolgen können, jedoch gehen Grosz und Sidner (1986) davon aus, daß eines dieser Ziele grundlegender und wichtiger ist als die anderen. Der DP soll Teil des mutuellen Wissens sein und daher allen Diskursteilnehmern bekannt werden. Jedes Segment des Diskurses verfügt darüber hinaus über ein eigenes Diskurssegmentziel (*discourse segment purpose*, DSP), das den spezifischen Beitrag dieses Segmentes zum DP liefert.

- Der *attentional state* repräsentiert die gerade im Verlauf des Diskurses salienten Objekte, Relationen und Eigenschaften und macht so Aussagen über den Aufmerksamkeitsfokus der Diskursteilnehmer. Je nach Situation der Diskursteilnehmer kann dieselbe *intentional structure* einen unterschiedlichen *attentional state* in verschiedenen Diskursen bewirken.

Die Theorie von Grosz und Sidner (1986) unterscheidet durch die klare Trennung in linguistische Struktur, intentionale Struktur und Fokusstruktur zwischen den Objekten,

wie sie z.B. nach der Analyse an der Eingabeschnittstelle eines Dialogsystems erkannt wurden, den Intentionen, die hinter den Äußerungen stehen und die von den jeweils anderen Diskurspartnern erkannt werden sollen, und den Aufmerksamkeitszuständen, die Referenz und andere Phänomene beeinflussen können. Während die Analyse der linguistischen Struktur für eine aktuelle Dialogschnittstelle umsetzbar ist, gibt es für die Erkennung von Sprecherintentionen erst wenige Ansätze (vgl. z.B. Carberry und Lambert (1999)). Eine Modellierung der Fokusstruktur im rein linguistischen Sinne ist im Rahmen von Theorien wie der DRT möglich (vgl. Geurts und Van der Sandt (1997)), eine Erkennung von Aufmerksamkeitszuständen bei Menschen durch eine Maschine ist, von wenigen Ansätzen zur Blickverfolgung abgesehen, noch nicht realisiert.

Carberry und Lambert (1999) stellen einen Ansatz zur Modellierung von kooperativen Verhandlungsdialogen vor. Aufbauend auf Allen und Perrault (1980), Grosz und Sidner (1986), Cohen und Levesque (1990) und Perrault (1990) zeichnen sich ihre Arbeiten im Gegensatz zu den genannten Ansätzen dadurch aus, daß unter Zuhilfenahme von linguistischem, kontextuellen und domänenabhängigen Wissen komplexe Diskursakte erkannt und Verhandlungs(sub)dialoge zwischen Diskurspartnern planbasiert modelliert werden. Dabei ist sowohl implizite Akzeptanz einer Proposition P_1 als auch der Ausdruck von Zweifel an P_1 durch Äußern einer Proposition P_2 möglich (dabei ist P_2 nicht bereits oberflächlich als im Widerspruch zu P_1 stehend erkennbar). Implizite Akzeptanz wird dann angenommen, wenn der Hörer von P_1 die Gelegenheit, P_1 zu bezweifeln, nicht wahrnimmt. Eine solche Gelegenheit besteht, wenn der Hörer von P_1 die Gesprächsinitiative hat und keine eingebetteten Verhandlungsdialoge mehr bestehen, die vor einer Verhandlung von P_1 zuende gebracht werden müßten. Damit sind die Autorinnen nicht nur in der Lage, (wie verwandte Ansätze) die Konflikte zwischen den Zielen mehrerer Diskurspartner zu modellieren, sondern können darüber hinaus auch die Verhandlungsdialoge der Diskurspartner darstellen. Dazu wurde auch die Möglichkeit unterschiedlich starker Überzeugungen in Bezug auf eine Proposition vorgesehen. Mit Hilfe der unterschiedlichen Wissensquellen und dem zugrundeliegenden planbasierten Modell können so komplexe Diskursakte wie „Zweifel“ erkannt werden.

Damit sehen Carberry und Lambert einen großen Schritt in Richtung auf ein intelligentes, sprachverstehendes System getan. Denn damit ein solches System (unbeschränkte) Benutzeräußerungen verstehen und adäquat darauf reagieren kann, muß es in der Lage sein, nicht nur die Struktur des Dialoges zu erkennen, sondern auch die kommunikativen Akte, die mit einzelnen, untereinander zusammenhängenden Äußerungen des Benutzers verbunden sind. Insbesondere muß ein solches System in die Lage versetzt werden, das natürliche Verhalten von Benutzern zu analysieren, die (zunächst unabhängig von Zielen) über unterschiedliche Überzeugungen mit dem System oder untereinander verhandeln. Dabei geht der Ansatz von Carberry und Lambert über z.B. Allen und Perrault (1980) hinaus. Bei diesem sorgt die Kooperativität der Diskurspartner dafür, daß die mit einem *inform act* übermittelte Proposition vom Hörer stets

den eigenen Überzeugungen hinzugefügt wird, wohingegen bei Carberry und Lambert (1999) ein Diskursteilnehmer eine Proposition ablehnen kann. Auch gegenüber einem Ansatz wie Perrault (1990) sind Carberry und Lambert im Vorteil, da Perraults Ansatz zwar die Zurückweisung einer Proposition durch einen Diskursteilnehmer vorsieht, wenn diese im Widerspruch zu eigenem Wissen steht, nicht aber eine Revision dieser Ablehnung vorsieht. Hingegen kann bei Carberry und Lambert eine Proposition zunächst bezweifelt und nach einem Verhandlungsdiallog dann akzeptiert werden.

In VERBMOBIL (vgl. 4.1) wird versucht, verschiedene Ansätze zur Dialogmodellierung und -verarbeitung zu verbinden. Die Dialogakte werden mit einer Kombination aus flacher und tiefer Erkennung bestimmt, wobei die tiefe Analyse Kontextinformation mit einbezieht, während die flache Analyse im wesentlichen stichwortbasiert arbeitet. Zur Vorhersage der nachfolgenden Dialogakte werden statistische Verfahren eingesetzt, und planbasierte Operatoren helfen bei der Einordnung der aktuellen Situation in den angenommenen Verlauf des Gesamtdialoges sowie auf der Ebene der Sprechakte (Maier, 1997; Alexandersson et al., 1997a,b).

Die Dialog-Komponente besteht aus drei Teilen:³³ einer statistischen Komponente, in deren Umfeld Vorhersagen über mögliche Fortsetzungen des Dialoges getroffen werden (Reithinger et al., 1996; Wermter und Löchel, 1996), einem Dialoggedächtnis (Maier, 1996), in dem die Dialog-Sequenz, die thematische Struktur des Dialoges und die intentionale Struktur gespeichert werden, sowie einem Plan-Erkenner, der für die Konstruktion bspw. der intentionalen Struktur verantwortlich ist. Die intentionale Struktur wiederum besteht aus vier Ebenen: Auf der untersten Ebene werden die Dialogakte repräsentiert, die für sich wiederum innerhalb einer Dialogakthierarchie repräsentiert werden (Jekat et al., 1995). Dialogakte auf der Dialogaktebene der intentionalen Struktur sind z.B. *reject_date*, *reject_duration* und *reject_location*. Diese Dialogakte (42 Stück) wurden aus einem handannotierten, domänenspezifischen Dialog-Korpus manuell ausgewählt (Jekat et al., 1995; Maier, 1997; Jekat et al., 1997). Auf der nächsthöheren Ebene der intentionalen Struktur werden turns im Dialog repräsentiert, wobei z.B. der turn *reject* aus mehreren Dialogakten bestehen kann. Auf der nächsthöheren Ebene der intentionalen Struktur wird für einen turn festgelegt, in welcher Phase des Dialoges er sich befindet, wobei in VERBMOBIL (aufgrund der Anwendungsdomäne) nur die drei Phasen „Begrüßung“, „Verhandlung“ und „Beendigung“ unterschieden werden. Die höchste Ebene der intentionalen Struktur ist die Dialogebene – hier wird repräsentiert, ob ein Dialog z.B. mehrmals in die Phase der „Verhandlung“ eintritt.

Diskursrepräsentation Herkömmliche syntaktische und semantische Theorien sind für die Beschreibung eines Diskurses (ohne wesentliche Erweiterungen) unbrauchbar, da sie sich am Satz als der zugrundeliegenden Einheit der Analyse orientieren. Die Struktur eines Diskurses zeichnet sich aber gerade durch Querverweise zwischen den aus mehreren Sätzen bestehenden Äußerungen eines Sprechers aus. Weiterhin ist

³³Alexandersson et al. (1994); Alexandersson (1996).

für die Diskursanalyse neben der Repräsentation der semantischen und syntaktischen Strukturen auch eine Repräsentation der Pläne und Ziele der Diskursteilnehmer nötig: Sowohl für die Auflösung stark elliptischer Äußerungen wie für kooperatives Systemverhalten sind Informationen über die den einzelnen Diskurssegmenten zugrunde liegenden Intentionen nötig. Eine semantische Theorie, auf die sich die Modellierung eines Diskurses stützen kann, muß daher die formalen Mittel bereitstellen, um die im Diskurs auftretenden syntaktischen, semantischen und pragmatischen Phänomene aufeinander abgestimmt zu repräsentieren.

Eine Theorie zur Repräsentation eines Diskurses ist zunächst ein Ansatz zur Beschreibung der Semantik natürlicher Sprache, dessen auffälligstes Merkmal in der Wahl des Diskurses (im Gegensatz zum Satz) als Grundeinheit der semantischen Analyse liegt. Die wohl bekannteste Theorie zur Repräsentation eines Diskurses, die auch für den maschinellen Einsatz geeignet ist,³⁴ ist die „Diskursrepräsentationstheorie“ (DRT), die Anfang der 80er Jahre von Hans Kamp (Kamp (1981); Kamp und Reyle (1990)) entwickelt wurde. Als Input dient der DRT ein natürlich-sprachlicher Satz, der durch eine beliebige Syntaxkomponente aufbereitet wurde. Auf der Grundlage eines solchen Syntaxbaumes arbeiten nun Konstruktionsregeln an der Überführung in eine Diskursrepräsentationsstruktur (DRS). Eine DRS K ist dabei ein Paar der Form $\langle U(K), \text{Con}(K) \rangle$, wobei $U(K)$ die Menge der Diskursreferenten (das sogenannte Universum) in der DRS K und $\text{Con}(K)$ die Menge der Bedingungen über den Diskursreferenten in K ist. Angenommen, es gibt eine endliche Menge an Diskursreferenten x_1, \dots, x_n und ein n -stelliges Prädikat R , dann ist $R(x_1, \dots, x_n)$ eine einfache Bedingung.

Einfache Bedingungen repräsentieren lexikalische Elemente des natürlichsprachlichen Satzes wie Substantive oder Verben, komplexe Bedingungen, sogenannte Sub-DRSen, werden zur Repräsentation syntaktischer Konstruktionen wie z.B. Konditionalsätze, quantifizierende Ausdrücke, Negationen und Disjunktionen verwendet.

Insbesondere zeichnet sich die DRT dadurch aus, daß sie als eine dynamische Semantik konzipiert ist. D.h. daß es einen Mechanismus gibt, um jederzeit bestehende Strukturen um neu hinzukommende Informationen inkrementell zu erweitern. Am Ende eines jeden Satzes kann im Verlauf des Diskurses die Repräsentation ausgewertet werden. Ein Diskurs kann daher so repräsentiert werden, wie er in der Realität auch erscheint: ohne festes Ende, jederzeit zu unterbrechen und stets wieder aufzunehmen. Die Repräsentation des Beispieldiskurses

(17) „Ein Mann besitzt einen Hund. Er liebt ihn.“

mit Hilfe der DRT geht wie folgt vor sich: Der indefinite Artikel im ersten Satz führt einen neuen Diskursreferenten x im Universum U der aktuellen DRS K ein. Da der zu repräsentierende Satz nicht Teil eines bereits begonnenen Textes oder Diskurses ist, ist die aktuelle DRS leer und der Diskursreferent wird im Universum der Haupt-DRS

³⁴Die DRT wurde in modifizierter Form z.B. in LILOG (Herzog und Rollinger, 1992) und in VERB-MOBIL (Wahlster, 1993) eingesetzt.

eingeführt. Das Nomen wird als Bedingung des Diskursreferenten im Bedingungsteil Con der DRS eingeführt. Genauso wird mit der NP *einen Hund* verfahren, wobei natürlich ein neuer Diskursreferent, z.B. *y* verwendet wird. Das Verb *besitzt* schließlich wird ebenfalls als Bedingung im Bedingungsteil der DRS eingetragen, und zwar als zweistellige Relation zwischen den Diskursreferenten *x* und *y*. Damit ergibt sich eine Repräsentation dieser Art:

<i>x,y</i>
<i>Mann(x), Hund(y), besitzt(x,y)</i>

Die Repräsentation des um den Satz *Er liebt ihn* erweiterten Ausdrucks erfolgt in der DRT durch die Einführung eines Diskursreferenten für jedes Personalpronomen sowie dessen Bindung an ein passendes Antezedens mittels Gleichsetzung. Da der zu repräsentierende Satz Teil eines bereits begonnenen Textes oder Diskurses ist, ist die aktuelle DRS diejenige, die eben konstruiert wurde. Zusätzlich wird *lieben* als zweistellige Relation repräsentiert, analog zu *besitzt*. Es ergibt sich somit im ersten Schritt folgende DRS:

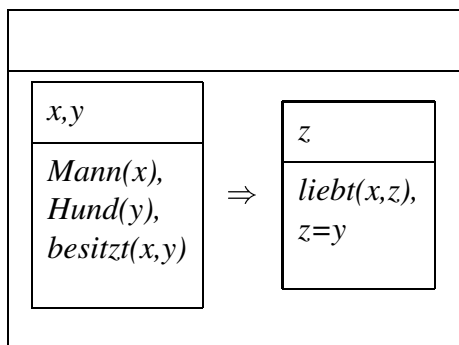
<i>x,y,q,r</i>
<i>Mann(x), Hund(y), besitzt(x,y), q=?, r=?, liebt(q,r)</i>

Das einfache Anfügen der neuen Diskursreferenten ist völlig kanonisch: Jede DRS ist darauf ausgelegt, daß zu jeder Zeit neue Information hinzugefügt werden kann. Diese Eigenschaft ist es, die die DRT zu einer *dynamischen* Semantiktheorie macht. In einem zweiten Schritt erfolgt nun die Auflösung der anaphorischen Referenz, als deren Ergebnis sich ergibt:³⁵

³⁵Natürlich hat dieser Satz auch eine zweite Lesart, die durch $q=y$ und $r=x$ ausgedrückt werden kann, s. a. (28), S. 119.

x, y, q, r
<i>Mann</i> (x), <i>Hund</i> (y), <i>besitzt</i> (x, y), $q = x$, $r = y$, <i>liebt</i> (q, r)

Neben einfachen gibt es in DRSEN auch komplexe Bedingungen: wenn K_1 und K_2 DRSEN sind, dann ist die Verknüpfung $K_1 \Rightarrow K_2$ sowie $K_1 \vee K_2$ eine komplexe Bedingung. Ebenso verhält es sich mit einer DRS K die unter einem Negationsoperator eingebettet ist: $\neg K$. Diese komplexen Bedingungen werden für die Konstruktion der DRSEN von z.B. universell quantifizierten Sätzen wie „Jeder Mann der einen Hund besitzt, liebt ihn.“ benötigt. Bei der Konstruktion der dem Satz entsprechenden DRS werden zwei SubDRSEN eingeführt, eine für den Antezedentsatz und eine für den Konsequenzsatz:



Dabei ist die Sichtbarkeit der einzelnen Diskursreferenten, je nach Ort ihres Auftretens, beschränkt. Generell kann aus eingebetteten DRSEN auf umliegende, höher gelegene DRSEN zugegriffen werden, aber nicht umgekehrt. Die Auflösung anaphorischer Referenzen kann so durch die beschränkte Sichtbarkeit der einzelnen Diskursreferenten gesteuert werden.

Die besonderen Merkmale der DRT (im Vergleich zur Prädikatenlogik erster Stufe) werden im Folgenden zusammengefaßt:

- Die DRT besitzt keinen Allquantor und der Mechanismus, der die entsprechenden natürlichsprachlichen Strukturen repräsentiert, funktioniert anders als die Allquantifikation in der PL1. In der PL1 ist es die totale Belegungsfunktion, die für alle Belegungen einer Variablen hinsichtlich des Universums des Modells verantwortlich ist. Dabei muß auch *jede* Belegung der Variablen betrachtet werden. In der DRT ist über die Semantik der allquantifizierende Mechanismus anders geregelt. In K_1 wird der Kontext zunächst auf die in Frage kommenden Entitäten eingeschränkt und mit K_2 wird dann gefordert, daß grundsätzlich eine Erweiterung der Einbettungsfunktion von K_1 existieren muß, die auch K_2 einbettet.
- Es gibt keinen Existenzquantor in der DRT. Was der Existenzquantor mit einer Belegungsfunktion in der PL1 leistet, wird in der DRT über die Einschränkung der Domäne erreicht. Die Belegungsfunktion der PL1 ist total, deshalb wird der Existenzquantor für eine bestimmte Variable benötigt, in der DRT ist das Pendant zur Belegungsfunktion, die Einbettungsfunktion dagegen partiell. Auf dieser Ebene der Betrachtung sind die Ergebnisse der jeweiligen Repräsentation äquivalent, aber die Mechanismen, die zu deren Erstellung führten, sind unterschiedlich.
- Die DRT ist eine dynamische Semantik in dem Sinne, daß neu hinzukommende Information inkrementell ausgewertet wird. Im Gegensatz dazu verfügt die PL1 nicht über eine Strategie zur inkrementellen Erweiterung. Neu hinzukommende Sätze können durch ein logisches *und* verknüpft dem bestehenden Ausdruck angefügt werden, wobei im Fall anaphorischer Bindung ggfs. der Skopus eines Quantors erweitert und der gesamte Ausdruck komplett neu ausgewertet werden muß.
- Auffälliger unterscheiden sich beide Theorien in der Behandlung von Eigennamen. Es gibt in der DRT keine Individuenkonstanten. *Theo* ist einfach eine Eigenschaft eines Diskursreferenten. Dennoch wird der Eigenname hier besonders behandelt,³⁶ indem er sofort in der Haupt-DRS eingeführt wird.
- Die DRT ist in der Lage, in Konditionalsätzen dynamisch zu binden. Das Problem der PL1 liegt nicht darin begründet, daß sie unfähig wäre, Konditionalsätze zu repräsentieren. Vielmehr besteht der Unterschied darin, daß die Variablenbindung verlorengelht, wenn man nach der Devise „indefiniten Artikel in der natürlichen Sprache meint Existenzquantor in der PL1“ vorgeht. Damit kann die DRT im Gegensatz zur PL1 mit impliziter Allquantifizierung wie in „Wenn ein Student ein Buch liest, dann bildet er sich“ adäquat umgehen.

³⁶Weil Eigennamen Präsuppositionsinduzierer sind, vrgl. 4.2.4.

- Ein letzter Unterschied besteht in den Möglichkeiten der beiden Theorien, Diskurse zu repräsentieren. Im Gegensatz zur PL1, die nur über eine Erweiterung des Quantorenkopos und damit einer Neuberechnung *aller* Belegungen einen Diskurs repräsentieren kann, ist die DRT als dynamische Semantik konzipiert und ist zur inkrementellen Auswertung am Ende eines jeden Satzes innerhalb eines Diskurses in der Lage.

4.2.2 Anaphorische Referenz

Anaphern sind Elemente die bereits im Diskurs eingeführte Objekte wieder aufgreifen und weitere Aussagen über sie treffen. Anaphern sind häufig Pronomen oder definite Nominalphrasen, die im klassischen Fall auf ein indefinit im Diskurs eingeführtes Objekt rückverweisen:

(18) „*Ein Mann* steht am Fenster. $\left. \begin{array}{c} \textit{Er} \\ \textit{Der Mann} \end{array} \right\}$ beobachtet die Wolken.“

In (18) wird für *ein Mann* und *er* angenommen, daß sie denselben Referenten haben, d.h. daß sie koreferieren. *Ein Mann* wird, weil es dem Pronomen *er* im Diskurs vorangeht, auch als dessen Antezedent bezeichnet.

Hirst (1981) beschreibt Anaphern als Referenzobjekte, deren Antezedenten Konzepte oder Entitäten seien, die explizit oder implizit in vorangegangenen Sätzen oder Situationen herausgestellt wurden. Anaphern werden, so Hirst, von Sprechern nur dann eingesetzt, wenn der Hörer (nach Meinung des Sprechers) grundsätzlich in der Lage ist, die Referenz aufzulösen. Es gibt für Hirst mehrere Wege, wie das Antezedens zu einer Anapher für den Hörer zugänglich sein kann, grundsätzlich unterschieden danach, ob das Antezedens sprachlich, d.h. im Text, oder außersprachlich zur Verfügung gestellt wird:

- das Antezedens wird explizit im Text genannt (und zwar „nicht allzuweit“ zurückliegend)
- das Antezedens ist implizit im Vortext gegeben (ebenfalls „nicht allzuweit“ zurückliegend)
- das Antezedens wird außersprachlich bezeichnet (durch Gestik o.ä.) oder ist (zumindest nach Meinung des Sprechers) prominent in einer Situation

Es ergibt sich für Hirst für die im Text aufzulösenden Referenzen³⁷ eine Anaphern-Taxonomie, angefangen bei Pronomen,

(19) „Maria kaufte einen Kaktus und schenkte *ihn* Peter zu *seinem* Geburtstag.“

³⁷Zu Referenzen auf außersprachliche Objekte siehe S. 121.

Proactions (im Englischen „do + so“ oder „do + it“, im Deutschen „machen/tun + das/dies/es“),

(20) „Maria fährt Rad. Peter *tut das* auch.“

über Proadjektive (im Englischen „such“, im Deutschen „so ein“, „ein solches“ etc.)

(21) „Peter lief grün an, röchelte etwa drei Minuten, bevor der Schaum auf seinen Lippen weniger wurde und er sich wieder beruhigte. *Solche* Ereignisse waren nichts Ungewöhnliches, seit er wieder englisches Rindfleisch aß.“

sowie Lokal- und Temporalreferenzen,

(22) „Peter suchte schließlich im Kohlenkeller. *Dort* fand er das Tagebuch.“

(23) „Im 19. Jahrhundert wurde die Dampfmaschine erfunden. *Damals* sorgte diese Erfindung für einen wirtschaftlichen Aufschwung.“

bis hin zur „ultimativen Anapher“, der Ellipse:

(24) „Maria liest 'Financial Times'. Peter \emptyset auch.“

Die Auflösung der Referenzbeziehung zwischen Anapher und Antezedens ist zum Zeitpunkt des Auftretens der Anapher nicht immer eindeutig durchführbar. So kann in (25) sowohl „Peter“ als auch „Hans“ Antezedens des Pronomens sein, wie die möglichen Fortsetzungen (26) und (27) zeigen:

(25) „Peter sah, wie Hans mit dem Ballon abhob.
Er war unauffällig gekleidet, ...“

(26) „... aber Peter verfolgte ihn über Minuten mit dem Fernglas.“

(27) „... aber Hans hatte ihn vor dem Start entdeckt.“

Während die Fortsetzungen (26) und (27) durch die explizite Namensnennung noch zu einer eindeutigen Festlegung der Referenz von „er“ führen, zeigt (28), daß eine solche Disambiguierung nicht immer zu erwarten ist,³⁸ sondern häufig nur durch Weltwissen aufgelöst oder zumindest weiter eingeschränkt werden kann. In (28) kommen als Antezedenten für „er“ sowohl der Mann als auch der Hund in Frage:

(28) „Ein Mann besitzt einen Hund. *Er* beißt ihn.“

Aufgrund von Weltwissen kann geschlossen werden, daß die Wahrscheinlichkeit, daß ein Hund ein anderes Lebewesen beißt sehr viel höher ist, als daß ein Mensch dieses tut. Auch bei definiten Nominalphrasen ist die Auflösung der Referenz vom Weltwissen abhängig:

³⁸Vgl. Schiehlen (1997) zu diesem Problem im System VERBMOBIL.

(29) „Boris Becker gewann das ATP-Turnier.“

(30) „Der Leimener spielte überragend.“

(31) „Der gestern gebrochene Oberschenkel hielt das Spiel über.“

(32) „Der Schiedsrichter versagte völlig.“

Während in (30) aufgrund des Weltwissens³⁹ „der Leimener“ als koreferent zu „Boris Becker“ erkannt werden kann, ist „der gestern gebrochene Oberschenkel“ in (31) vermutlich ein Teil von „Boris Becker“, aber nicht koreferent mit ihm und deshalb kein Beispiel für innertextuellen anaphorischen Bezug (wohl aber für anaphorischen Bezug auf einen durch die Nennung von „Boris Becker“ salient gewordenen Teil desselben). Und schließlich ist in (31) „der Schiedsrichter“ nicht koreferent mit „Boris Becker“, weil aufgrund des Weltwissens eindeutig ist, daß der „Schiedsrichter“ unparteiisch ist, also ein Spiel nicht gewinnen kann.

Andere Faktoren, die die Auflösung von Anaphern beeinflussen sind z.B. Betonung und Kontext (Bosch, 1983; Kasper, 1996):

(33) „Watson fand die Pfeife auf dem Tisch.“

(34) „Er hatte sie dort am Tag zuvor abgelegt.“

(35) „ER hatte sie dort am Tag zuvor abgelegt.“

Während „er“ in (34) auf „Watson“ im Vorsatz referiert, kann eine besondere Betonung von „er“ (in (35) durch die Großschreibung und Unterstreichung ausgedrückt) zu einer Interpretation führen, bei der nicht auf „Watson“, sondern auf eine andere, nicht direkt genannte Person verwiesen wird. Diese Person kann z.B. durch den Kontext gegeben sein („der große Unbekannte“, „der berühmte Meisterdetektiv“ usw.).

Andererseits gibt es syntaktische Konstruktionen, in denen die Referenz des Pronomens eindeutig bestimmt ist:

(36) „*Peter* sah Hans und kämmte *sich* die Haare.“

In (36) ist eine Referenz des reflexiven „sich“ nur auf „Peter“, nicht aber auf „Hans“ möglich.

Anaphorische Referenz wird an natürlichsprachlichen Schnittstellen nur dann virulent, wenn komplexe Sätze mit satzinternen Referenzen verarbeitet werden sollen oder aber eine kontinuierliche Interaktion zwischen Benutzer und Maschine möglich ist, d.h. mehrere aufeinanderfolgende und aufeinander bezugnehmende Eingaben erfolgen. Obwohl diese Möglichkeit per Definition eine Eigenschaft des menschlichen Dialoges ist, sind dennoch nicht viele Dialogsysteme dazu in der Lage.

³⁹z.B. aus den Sportteilen der Zeitung, in denen beschreibende Nominalausdrücke wie „der dreifache Deutsche Meister“, „der letztjährige Olympiasieger“, „der frisch nach Madrid gewechselte Profispieler“ usw. als Stilmittel beliebt sind.

Wie Ellipsis (siehe 4.2.3) führt auch Anaphorik zu mehr Ökonomie im zwischenmenschlichen Diskurs, weil nicht alles bereits Gesagte wiederholt werden muß, sondern auf im Diskurs bereits eingeführte Objekte referiert d.h. verwiesen werden kann. Wie Ellipsis ist auch Anaphorik im Mensch-Maschine-Dialog unabhängig von einem Selbstanspruch des Systems, anthropomorph (oder auch nicht) zu wirken – vielmehr ist Anaphorik für einen natürlichen d.h. unangestregten Dialog mit dem Benutzer unter den oben ausgeführten Rahmenbedingungen unabdingbar.

Verarbeitung von Anaphern Mit Hilfe von Diskursrepräsentationstheorien (siehe 4.2.1) ist die automatische Erkennung und Behandlung klassischer Anaphern (wie die in (18) gezeigte) möglich (siehe auch 4.2.4 für einen Ansatz zur Beschreibung von Präsuppositionen als anaphorischer Referenz). Die Erkennung möglicher anaphorischer Elemente wird über das Lexikon gesteuert, die Auswahl eines geeigneten Antezedenten erfolgt auf der Grundlage der Diskursobjekte, über die eine Diskursrepräsentationstheorie Buch führt. Damit ein Diskursobjekt als Antezedens einer Anapher in Frage kommt muß untersucht werden, ob es sich in einem „erreichbaren“ Abschnitt des Diskurses befindet, da ein Diskurs eine interne Struktur besitzt, die die behandelten Themen wiedergibt.

Informationen über die Diskursstruktur und deren Interpretation für die Auflösung z.B. anaphorischer Referenzen müssen aus Theorien über den Aufbau des menschlichen Diskurses (siehe 4.2.1) in die eigentlichen Diskursrepräsentationsformalisten übernommen werden.

Ist für eine erkannte Anapher eine Menge prinzipiell erreichbarer Diskursobjekte identifiziert, so kann auf Grundlage von Kasus-, Numerus- und Genuskongruenz eine weitere Auswahl getroffen werden. In der Regel werden dann immer noch eine ganze Reihe möglicher Antezedenten übrig bleiben, die nun über Parameter wie Nähe zur Anapher oder Auftreten in bestimmter syntaktischer Position geordnet werden können. So wurde in LILOG (Herzog und Rollinger, 1992) auf der Basis erweiterter Diskursrepräsentationsstrukturen (EDRSen) mit der Heuristik gearbeitet, nach Untersuchung der syntaktischen Kongruenz bevorzugt ein im Vorsatz auftretendes Diskursobjekt als Antezedens zu wählen, im Zweifelsfall das Subjekt des Vorsatzes.

Die nach der Prüfung der syntaktischen Kongruenz verbleibende Menge kann auch weiter über semantische Kriterien sowie über Weltwissen eingeschränkt werden. Dies stellt allerdings für den Einsatz in Dialogsystemen einen hohen Anspruch an die beteiligten Komponenten. Zu den speziellen Schwierigkeiten, Weltwissen zu sammeln, zu repräsentieren und zu pflegen siehe auch S. 105.

Die sprachliche Referenz, d.h. der Verweis innerhalb des Diskurses auf sprachliche Elemente, ist deutlich einfacher zu handhaben, als die außersprachliche Referenz. Bei letzterer ist zu unterscheiden zwischen der Referenz auf Objekte im Weltwissen des Benutzers und Referenz auf Objekte, wie sie z.B. auf graphischen Benutzeroberflächen vom System dem Benutzer dargeboten werden. Während erstere sofort in die aus der Künstlichen Intelligenz bekannt schwierige Modellierung und Begrenzung von Welt-

wissen des menschlichen Benutzers führt, erscheint letztere handhabbar und auch nur folgerichtig zu sein (vgl. Krüger und Ronthaler (1997) zur Referenz auf Objekte der graphischen Benutzeroberfläche in CALL-Systemen).

Menschen kommunizieren normalerweise über mehrere Kanäle gleichzeitig, neben Sprache z.B. auch mit Zeigegesten. Eine Dialogschnittstelle, die über eine (Tastatur-)Eingabemöglichkeit hinaus auch eine detaillierte graphische Schnittstelle zur Verfügung stellt (siehe 5.2.2), setzt sich somit fast automatisch dem Anspruch aus, ein integriertes Gesamtsystem zu sein. Hier können den Zeigegesten im zwischenmenschlichen Dialog vergleichbar bspw. Mauszeigerbewegungen interpretiert werden, um deiktische Referenzen (*dieser dort*) aufzulösen. Die Chancen des Systems sind dabei solange gut, wie sich die Referenzen auf zum System gehörige Teile z.B. auf der Oberfläche beziehen (siehe 5.2.2). Was derzeit eingesetzten Systemen dazu fehlt, ist die interne Repräsentation der Außenansicht des Systems selbst, d.h. die Repräsentation, wie es dem Benutzer erscheint, sowie innerhalb dieser Repräsentation die Auszeichnung der dem Benutzer präsentierten und damit potentiell als Referenzobjekte in Frage kommenden Systemteile.

Ob sich das Konzept der Anapher überhaupt auf graphische Elemente im multimodalen Dialog ausdehnen läßt, ist umstritten. So argumentieren Wahlster et al. (1991) und Wahlster et al. (1993), daß durch die Kombination von Text und Graphik anaphorische Bezüge zwischen sprachlichen Elementen und Teilen der Graphik zum Tragen kommen (vgl. auch André und Rist (1993) und 5.2.2). Dies wird von Lee und Stenning (1998) bestritten, die argumentieren, daß Koreferenz als Voraussetzung von Anaphorik nur in type-referentiellen Systemen wie z.B. Sprache auftreten kann. Da Abbildungen nach Auffassung von Lee und Stenning (1998) grundsätzlich token-referentiell seien, könnten sie nicht als Anaphern betrachtet werden.

4.2.3 Ellipsen

Definitionen für das, was eine Ellipse ist, gibt es zahlreich und in unterschiedlich restriktiver Form. Eine allgemeine Definition für „Ellipse“ bezeichnet diese als nicht in der Oberflächenform des Satzes realisiertes Element.

Beispiele⁴⁰ für verschiedene Ellipsenarten sind die folgenden, wobei \emptyset das getilgte Element repräsentieren soll:

(37) „Peter sah Maria, \emptyset warf die Arme in die Luft und rannte zu ihr.“
 \emptyset =Peter

(38) „Wer entdeckte Amerika?“
 „Die Wikinger \emptyset .“
 \emptyset =entdeckten Amerika

⁴⁰Vgl. Busch (1990) für eine weitergehende Übersicht und Klassifikation.

(39) „Peter fuhr nach Rom und Maria \emptyset nach Prag.“

\emptyset =fuhr

(40) „Peter liebt Jazz, aber Maria hat nie verstanden, warum \emptyset .“

\emptyset =Peter Jazz liebt

Nach Thomas (1979) verführt die eben gegebene sehr allgemeine Definition für „Ellipse“ dazu, all jene Elemente als Ellipse zu bezeichnen, die in einem Satz nicht auftauchen, dort aber eventuell hätten auftauchen können. Die Annahme hypothetischer, getilgter Elemente habe dazu geführt, daß sogar Elemente als Ellipsen eines Satzes betrachtet worden seien, die nie in diesem Satz hätten vorhanden gewesen sein können. Als Warnung vor einer zügellosen Auslegung des Begriffs Ellipse zitiert Thomas Jespersen, der schon 1937 gewarnt habe, vor

„ellipsomania [...] a panacea to explain a great many things which either need no explanation or else are not explained, or not sufficiently explained, by the assumption that something is understood.“

(zitiert nach Thomas (1979), S.43-44)

Aus den genannten Überlegungen zu den Folgen einer zu weit gefaßten Definition von Ellipse versucht Thomas eine schwächere Definition, nach der Ellipsen ein kommunikatives Phänomen seien, bei dem etwas ungesagt bliebe, aber aus dem Kontext heraus verstanden werde.

Dies führt ihn zu drei Phänomenklassen, die einander sehr ähnlich sind, von denen Thomas aber nur eine als Ellipse gelten lassen will. Indem er diese drei Klassen gegeneinander abgrenzt, bestimmt er die seiner Ansicht nach wesentlichen Merkmale von Ellipsen näher und gibt Methoden an, wie diese einander oft sehr ähnlichen Klassen unterschieden werden können.

Die erste dieser Phänomenklassen ist die Elision. Der folgende Satz ist für Thomas ein Beispiel für Elision:

(41) „Got the tickets ?“

Dieser Satz ist unvollständig, aber im Gegensatz zum folgenden kontextuell unabhängig:

(42) „Yes, I have.“

Hier muß das getilgte Element kontextuell verfügbar sein. Im vorangegangenen Satz hingegen sorgt das sprachliche Wissen des Hörers, nicht der Kontext, dafür, daß das getilgte „Have you“ ergänzt werden kann. „Yes, I have.“ ist für Thomas ein Beispiel für Ellipse. Elision, als ein von der Ellipse klar zu trennendes Phänomen, ist an Sprachkonventionen gebunden. Elisionen sind nicht selbst-definierend, wie Thomas am Beispiel von „plane“ zeigt, das sowohl als „aero-plane“ als auch als „hydro-plane“ interpretiert werden kann. Jedoch sorgt die Konvention für die Präferenz der ersten Interpretationsmöglichkeit. Nicht anders für den Satz „Got the tickets ?“, der durchaus unterschiedlich interpretierbar ist:

- (43) „ $\left. \begin{array}{l} \text{Have you} \\ \text{Have they} \\ \text{Has he} \end{array} \right\}$ got the tickets ?“

Sprachliche Konventionen sorgen bei subjektlosen Fragen für eine Interpretation mit zweiter Person Singular.

Der dritte Fall linguistischer Abwesenheit (neben Elision und Ellipse) ist für Thomas „Nonrealization“. Ein Beispiel für Nonrealization ist die Passivierung:

- (44) „Arthur’s been murdered.“

In diesem Satz fehlt das Agens, die Person, die Arthur umgebracht hat. Daß eine solche Person existiert, folgt logisch aus unserem Wissen über die Handlung „morden“. Nonrealization unterscheidet sich von den beiden anderen Phänomenen dadurch, daß weder das fehlende Element (wie im Fall der Ellipse) obligatorisch aus dem Kontext bestimmt werden kann, noch (wie im Fall der Elision) durch sprachliche Konventionen festgelegt ist, wie dieser Satz um das fehlende Element expandiert werden könnte.

Das für Ellipsen charakteristische Merkmal ist laut Thomas also deren obligatorische Rekonstruierbarkeit aus dem Kontext. Dabei ist der Unterschied zwischen Ellipse und Nonrealization für Thomas genau der, den Shopen (1973) als Unterschied zwischen definiten und indefiniten Ellipse angibt.

Thomas faßt den Unterschied zwischen Ellipse und Nonrealization noch einmal in einem Satz zusammen:

„The difference between ellipsis and nonrealization lies in the difference between obligatory and optional elements [...].“
(Thomas, 1979, :52)

Damit beschreibt Thomas zum einen den Charakter der obligatorisch identifizierbaren Ellipse und zum anderen die Nonrealization als eine *Option* zur Expansion eines Satzes in einen längeren, semantisch reicheren, anderen Satz. Dementsprechend sollen nach seiner Meinung auch durch Ellipsen hervorgerufene Lücken als solche in einer Satzstrukturbeschreibung repräsentiert werden. Hingegen sollte ein durch Nonrealization abwesendes Element nicht in der syntaktischen Beschreibung des Satzes vermerkt werden, da es ja nicht Teil der aktuellen, sondern einer potentiellen Erweiterung der aktuellen Satzstruktur sei.

Fehlt einem Verb ein Argument, so wird dieses Argument in einigen Fällen existentiell interpretiert, in anderen jedoch als vom Kontext abhängig gesehen. Letzteres gibt Anlaß zur Interpretation einiger Verben wie „zustimmen“ oder „ablehnen“ als mit einer leeren Argumentstelle versehen, die sich anaphorisch verhält.

Es geht dabei um die Frage, warum (46), isoliert gesehen, verständlich wirkt, während (45), wiederum isoliert betrachtet, mindestens die Frage aufwirft, was oder wem Peter denn nun zugestimmt hat:

- (45) „Peter hat zugestimmt.“

(46) „Peter hat einen Brief bekommen.“

Der erste Satz ist, verglichen mit dem zweiten, offensichtlich zusätzlich vom Kontext abhängig. Diese Referenz auf eine Information aus dem Kontext, ohne daß eine Anapher zu sehen wäre, wird deshalb auch als „Null“-Anapher bezeichnet.

(47) „Peter betrachtet sich im Spiegel. Maria \emptyset_1 auch \emptyset_2 .“

\emptyset_1 =betrachtet sich *oder*

\emptyset_1 =betrachtet Peter

\emptyset_2 =im Spiegel

Das Beispiel (47) ist besonders interessant durch die offensichtliche Ambiguität (siehe auch 4.2.13) zwischen den beiden Interpretationen, was getilgt wurde. Dabei muß für die Interpretation des Satzes zwischen der sogenannten *sloppy* Lesart

(48) „Maria betrachtet sich im Spiegel“

und der *strict* Lesart entschieden werden:⁴¹

(49) „Maria betrachtet Peter im Spiegel“

Diese Ambiguitäten entstehen, wie (47), (50) und (51) zeigen, bei der Auflösung von reflexiven Verben bzw. Pronomen:

(50) „Peter vertraut seinem Versicherungsvertreter.“⁴²

Hans auch.“

(51) „Peter glaubt, daß er zum 'Manager des Jahres' gewählt wird.

Hans auch.“

Ellipsen sind also weder durch bloßes Löschen des einen von zwei identischen Satzteilen zu erklären, noch durch bloßes Kopieren eines Satzteilens aufzulösen. (38), (40) und (47) sowie (50) und (51) zeigen deutlich, daß das getilgte Element zwar in einer Relation zu einem anderen steht, mit diesem jedoch nicht identisch sein muß. Insbesondere zeigt (38), daß das getilgte Element Kongruenz zu beachten hat, also mehr sein muß, als ein bloßer Verweis auf ein anderes Element.

⁴¹Zu einem Überblick über die verschiedenen Ansätze zur Unterscheidung, siehe Dalrymple et al. (1991).

⁴²In diesem Beispiel wird eine mögliche Lesart, bei der „seinem“ auf jemand anderen als „Peter“ referiert, nicht beachtet.

Ellipsen an NL-Schnittstellen: Elliptische Konstruktionen sind an NL-Schnittstellen von Dialogsystemen immer dann zu erwarten, wenn die Mensch-Maschine-Kommunikation über einen bloßen Austausch von isolierten Frage-Antwort-Paaren hinausgeht: An Datenbankschnittstellen kann es häufig zu einer Folge ähnlicher oder nahezu identischer Anfragen kommen, so daß Benutzer geneigt sein können, ab der zweiten Anfrage durch elliptische Konstruktionen sich effizienter auszudrücken:

- (52) „Wo finde ich Bücher über Goethes Jugend ?“
Im dritten Stock, in der Abteilung THQ.
- (53) „Und über Schiller ?“
Ebenfalls im dritten Stock, in der Abteilung THQ.
- (54) „Finde ich dort auch die Zeitschriften ?“
Nein, die Zeitschriften befinden sich im Keller.
- (55) „Auch die Tonträger ?“

Sobald an einer erweiterten Version der Eingabeschnittstelle des OSIRIS Systems ein echter Dialog möglich ist, ist daher mit elliptischen Konstruktionen zu rechnen.

Wie auf Seite 124 ausgeführt, ist das charakteristische Merkmal für eine Ellipse deren Rekonstruierbarkeit aus dem Kontext. Damit sind wir, nach den Betrachtungen zu verschiedenen Arten der Ambiguität, bei der Rolle des Kontextes (siehe 4.2.9) als wichtige Informationsquelle für ein sprachverstehendes System. So beschreiben Carbonell et al. (1983) für das Xcalibur System eine Reihe von „discourse expectation rules“ und „contextual substitution rules“, um das fehlende Element in elliptischen Anfragen zu erschließen.

„Discourse expectation rules“ beschreiben für Systemreaktionen die möglichen, erwarteten Reaktionen des Nutzers. In einer Situation, in der der User mit dem System über ein (als *type* aber nicht als *token*) spezifiziertes Objekt kommuniziert und das System mit der Angabe eines Attribut-Wert-Paares den User zu einer (eindeutigen) Identifikation veranlassen will, erwartet das System vom User in der Folge entweder eine explizite Bestätigung oder Zurückweisung des identifikationsunterstützenden Attributes, die Angabe eines anderen, semantisch zulässigen Attributes als Ersatz für das vom System vorgeschlagene, die Angabe eines (zum gegebenen Attribut passenden) Wertes als Ersatz für den vom System vorgeschlagenen Wert oder die explizite Frage des Benutzers nach möglichen Werten des vom System angeführten Attributes. Vor dem Hintergrund dieser Erwartungen können elliptische Reaktionen des Benutzers erwartungsgesteuert, d.h. in sehr eingeschränkter Form kontextabhängig aufgelöst werden. Eine beispielhafte „contextual substitution rule“ ist z.B.

Wenn: Der Name eines Attributes (oder eine Liste von Attributnamen) ist vorhanden, ohne daß ein Objekt vorhanden ist, auf das dieses Attribut verweisen würde.

Zudem ist das Attribut semantisch zulässig, d.h. wenn ein Objekt im Rahmen des bislang geführten Diskurses genannt wäre, würde das Attribut passen.

Dann: Ersetze das Attribut der letzten Äußerung durch das aktuelle Attribut und übernehme das entsprechende Objekt.

Mit Hilfe einer solchen Regel ist es möglich, die elliptische Äußerung des Benutzers in (56) zu behandeln:

- (56) Benutzer: „Wie hoch ist der Preis der drei größten SCSI-Festplatten?“
 System: 3.580 DM für ...
 Benutzer: „Und die Größe und Zugriffszeit?“

In vielen existierenden Dialogsystemen (so auch in OSIRIS) werden aber schon durch das Systemdesign Bezüge auf bereits zuvor getätigte Eingaben verhindert. Wird nach Reaktion auf eine Benutzeranfrage wie *Haben Sie Bücher von Goethe?* wieder der Eingangsbildschirm gezeigt und der Eindruck erweckt, das System begegne dem Benutzer zum ersten Mal, werden weitergehende Fragen wie *Und von Schiller?* unterbunden. Ein solches Systemverhalten ist aber weit entfernt von einem als natürlich empfundenen Dialog.

Ellipsen werden seit langer Zeit als ein den Anaphern verwandtes Problem betrachtet, da es ähnlich wie im anaphorischen Bezug eines Pronomens auch zu einem getilgten Element einen Antezedenten und zwischen beiden eine Relation geben müsse (siehe 4.2.2).

Hirst (1981) bezeichnet Ellipsen als „ultimativ“, da hier die Anapher nicht mehr an einen sprachlichen Ausdruck wie z.B. ein Pronomen gebunden, sondern völlig leer sei (siehe auch (24), S. 119). Wenn nun aber, wie Hirst schreibt, Ellipsen genau wie Pronomen und andere sprachliche Ausdrücke Anaphern sind, dann wird im Zuge eines umfassenden Ausbaus der Theorien zur Behandlung von Anaphern, wie es in den auf der DRT (Kamp und Reyle, 1990) (siehe auch 4.2.1) aufbauenden Theorien von Van der Sandt und Geurts der Fall ist (vgl. Van der Sandt (1988); Van der Sandt und Geurts (1991); Van der Sandt (1992); Geurts (1995)), es auch zu weiteren Fortschritten bei der Behandlung von Ellipsen kommen. Allerdings stellt die erwähnte Ambiguität bei der Ellipsenauflösung ein Problem dar, das (s. a. „Ambiguität“ 4.2.13), vermutlich nur interaktiv mit dem Benutzer gelöst werden kann.

4.2.4 Präsuppositionen

Präsuppositionen sind die Voraussetzungen eines Satzes, die unbedingt erfüllt sein müssen, damit sinnvoll über Wahrheit oder Falschheit des Satzes nachgedacht werden kann. In dem Satz

- (57) „Marc beginnt zu arbeiten.“

lautet die *Aussage*, daß die mit „Marc“ bezeichnete Person vom Zeitpunkt der Äußerung an arbeitet. Die *Präsupposition* lautet:

(58) „Marc hat vor dem Zeitpunkt der Äußerung nicht gearbeitet.“

Ohne diese Voraussetzung wäre eine sinnvolle Verwendung von „beginnen“ und damit eine Entscheidung über die Wahrheit oder Falschheit der Aussage von (57) nicht möglich.

Daß die Präsupposition kein Teil der Aussage des Satzes ist, läßt sich durch Manipulation der Aussage, typischerweise durch einen Negationsoperator, belegen:

(59) „Marc beginnt nicht zu arbeiten.“

Auch für (59) gilt die in (58) wiedergegebene Präsupposition als Voraussetzung für die Entscheidung, welchen Wahrheitswert Satz (59) in einer bestimmten Situation bekäme.

Pragmatisch gesehen erlangen Sprecher durch den Einsatz von Präsuppositionen einen sprachökonomischen Vorteil. Damit sind Präsuppositionen (wie Ellipsen, s. 4.2.3) nicht nur in der zwischenmenschlichen, sondern auch in der Mensch-Maschine-Kommunikation von Bedeutung.

Die Verwendung eines Satzes, der eine Präsupposition enthält, setzt normalerweise voraus, daß die Äußerung in einem Kontext stattfindet, in dem der Hörer ebenfalls über die vorausgesetzte Information verfügt. Hierbei kann auf alles referiert werden, was gemeinsam im Diskurs vereinbart wurde. Ist einmal in einem Gespräch die Existenz eines bestimmten, für beide Gesprächspartner eindeutig zu identifizierenden Buches geklärt worden, so muß im Verlauf des weiteren Gespräches nicht erneut die Existenz dieses Buches besprochen, sondern kann vorausgesetzt werden. Darüber hinaus kann ein Sprecher aber auch die Existenz beliebiger anderer Dinge, von denen er meint, daß sie gemeinsames Wissen zwischen ihm und dem Hörer seien, voraussetzen.

Aber auch in dem Fall, daß der Hörer nicht über die Information verfügt, die der Sprecher präsupponiert, muß das Verständnis des Hörers nicht notwendig leiden. Ein kooperativer Hörer, der nicht weiß, daß Maria einen Learjet besitzt, wird, wenn er (60) hört,

(60) „Peter weiß, daß Maria einen Learjet besitzt.“

sein Wissen um die präsupponierte Information anreichern, um auf der Basis dieses erweiterten Wissens den Satz verarbeiten zu können. Diese Erweiterung des Wissens ist natürlich nur dann möglich, wenn der resultierende Wissenszustand widerspruchsfrei ist (vgl. Lewis (1979)).

Den Verlauf eines Dialoges zwischen zwei Menschen oder zwischen Mensch und Maschine (die Möglichkeit hinreichend komplexer Eingaben vorausgesetzt), kann man sich also durchaus als eine gemeinsame Arbeit vorstellen. Jede Äußerung reichert das

gemeinsame Wissen weiter an und wird vor dem Hintergrund der bisher getätigten Äußerungen verarbeitet. Wird nun eine Äußerung wie (60) getätigt, so muß die enthaltene Präsupposition aus dem bisher Gesagten folgen, oder sich widerspruchsfrei in das Wissen des Hörers (Mensch oder Maschine) integrieren lassen.

Damit ein Sprecher Präsuppositionen erfolgreich im Dialog einsetzen kann, muß ein Hörer zunächst in der Lage sein, sie zu erkennen. Ein solches Verfahren, mit dessen Hilfe für eine gegebene Äußerung überprüft werden kann, ob und wenn ja wo eine Präsupposition enthalten ist, wird *Präsuppositionstest* genannt und ist eine unmittelbare Voraussetzung für die theoretische Erklärung dieses Phänomens. Zu diesem Punkt sowie weitergehenden Anforderung an eine Präsuppositionstheorie wie der Frage nach der Entstehung einer Präsupposition, d.h. *woher* die Präsupposition kommt (ob aus dem Lexikon, der Grammatik oder anderswoher) und *wie* diese Information aktiviert wird, vergleiche Geurts (1995); Van der Sandt (1988, 1989); Van der Sandt und Geurts (1991); Van der Sandt (1992); Ronthaler (1994).

Für die vorliegende Arbeit von speziellerem Interesse ist die Notwendigkeit eines Präsuppositionstests als Voraussetzung für den effektiven Einsatz von Präsuppositionen an Mensch-Maschine-Dialogschnittstellen.

Präsuppositionstests Präsuppositionen sind unabhängig von der Aussage des Satzes. Um zu zeigen, daß die Präsupposition wirklich eine Vorbedingung ist, die einer wie auch immer gearteten Aussage erst eine sinnvolle Verwendung ermöglicht, wurde in (59) die Aussage von (57) mit Hilfe des Negationsoperators manipuliert. Angesichts der durch die Negation unveränderten Präsupposition könnte man eine Präsupposition allgemein als die Implikation eines Satzes betrachten, die sich gegenüber der Negation invariant verhält.

Der Versuch, Präsuppositionen generell über einen Präsuppositionstest auf Basis der Invarianz gegenüber der Negation zu diagnostizieren, schlägt aber aus zwei Gründen fehl. Einerseits, weil es neben der präsuppositionserhaltenden Negation auch eine zweite, präsuppositionsnegierende Negation gibt, andererseits, weil z.B. im Kontext polarer Elemente entweder keine oder nur eine markierte Form der Negation möglich ist.

Der *erste Einwand*, daß die negierte Form eines Präsuppositionen enthaltenden Satzes im allgemeinen neben der präferierten, präsuppositionserhaltenden Lesart (61) eine zweite, präsuppositionsnegierende Lesart hat, wird durch (62) illustriert:

(61) „Es ist nicht der Fall, daß der Weihnachtsmann einen Bart trägt.“

(62) „Es ist nicht der Fall, daß der Weihnachtsmann einen Bart trägt, denn es gibt gar keinen Weihnachtsmann !“

Die präferierte Lesart von (61) verneint die Existenz eines Bartes. In der zweiten Lesart, durch Ergänzung explizit gemacht in (62), wird allerdings die Existenz desjenigen,

um dessen Bart es geht, verneint – mit der Verneinung der Existenz des Weihnachtsmannes einher geht damit selbstverständlich auch die Verneinung der Aussage, er würde einen Bart tragen.

Ob es sich bei dieser markierten Art der Negation um ein wahrheitsfunktionales Phänomen handelt, das dann als starke (präsuppositionsnegierende) Negation von der schwachen (präsuppositionserhaltenden) Negation zu unterscheiden ist, oder ob es sich bei der markierten Negation um ein pragmatisches Phänomen, also eine zweite Verwendungsweise der Negation handelt, ist noch Streitpunkt der Forschung. Allerdings gibt es gewichtige Gründe zu der Annahme, daß es sich nicht um ein wahrheitsfunktionales Phänomen handeln kann, u.a. deshalb, da sonst Präsuppositionen auf der Ebene der logischen Form über verschiedene Negationsoperatoren oder unterschiedlich großen Skopus eines Negationsoperators erklärt werden müßten (vgl. u.a. Horn (1989) und Seuren (1985)).

Die markierte Form der Negation, die auch Präsuppositionen negieren kann, wird in der natürlichen Sprache oft auf die gleiche Art und Weise erzeugt, so daß eine echte Ambiguität (durch die Möglichkeit der Erweiterung von (61) nach (62)) besteht.

Der *zweite Einwand* gegen einen Einsatz des Negationstests bezieht sich auf das Verhalten polarer Elemente gegenüber der Negation. Polare Elemente sind Wörter, zu denen es keine negierte oder auch nicht-negierte Form gibt. Dementsprechend wird zwischen positiv polaren Elementen wie „immer noch“

(63) „Bill glaubt immer noch an Theos Theorie.“

(64) „Bill glaubt nicht $\left\{ \begin{array}{l} \text{*immer noch} \\ \text{mehr} \end{array} \right\}$ an Theos Theorie.“

(65) Bill glaubte früher an Theos Theorie.

und negativ polaren Elementen wie „(be)kümmern“ unterschieden:

(66) „Chris kümmert es nicht, daß seine Theorie falsch ist.“

(67) „Chris $\left\{ \begin{array}{l} \text{*kümmert es} \\ \text{ärgert es} \end{array} \right\}$, daß seine Theorie falsch ist.“

(68) Chris' Theorie ist falsch.

Satz (63) präsupponiert (65) und enthält mit „immer noch“ ein positiv polares Element, zu dem es keine sinnvolle, grammatische negierte Form gibt, die dem Negationstest eine Möglichkeit gäbe, (65) als Präsupposition von (63) zu identifizieren. In (66) mit der Präsupposition (68) liegt mit „kümmern“ ein negativ polares Element vor. Die Negation von (66), wiedergegeben in (67), kann nun aber nicht mehr als Argument für den Negationstest verwendet werden, da sie nur mit Hilfe der positiven Umformulierung „ärgert es“ möglich ist.

Auch hier scheitert der Negationstest bei dem Versuch, die durch „immer noch“ und „kümmern“ induzierten Präsuppositionen des Satzes zu identifizieren, weil es keine sinnvoll negierte Form gibt.

Die Tatsache, daß in bestimmten Kontexten, z.B. als Erwiderung in einem Dialog, dennoch die markierte Variante aus (64) verwendet werden kann,

(69) **A:** „Bill glaubt immer noch an Theos Theorie.“

B: „Bill glaubt nicht immer noch an Theos Theorie. Er hat nie an sie geglaubt.“

steht nicht im Widerspruch zu dieser Schlußfolgerung, sondern ist Ausdruck anderer Phänomene. Von einigen Forschern wird in diesem Zusammenhang von einer markierte Form der Negation ausgegangen, die von Van der Sandt (1991) als „Echo-Effekt“ und von Horn (1989) als „metalinguistische Negation“ analysiert wird. Geurts (1998) hingegen bestreitet die Existenz eines einheitlichen Phänomens als Grundlage sowohl der metalinguistischen Negation von Horn, als auch des Echo-Effektes von Van der Sandt und führt Effekte wie in (69) auf einzelne, unabhängige Phänomene zurück.

Es gibt zwei weitere Möglichkeiten, die Präsuppositionen eines Satzes (sofern vorhanden) zu testen. Wie der Negationsoperator, sorgt auch ein (schwacher) Modaloperator dafür, daß innerhalb seines Skopus die Implikationen eines Satzes verschwinden, während die Präsuppositionen, wie in (70) und (71) ersichtlich, davon unbeeinflusst bleiben:

(70) „Peter bedauert, daß die Gans tot ist.“

(71) „Es ist möglich, daß Peter bedauert, daß die Gans tot ist.“

Im Gegensatz zu (70) impliziert (71) nicht mehr den, mit „Bedauern“ beschriebenen, Gemütszustand von Peter, jedoch präsupponieren beide Sätze den Tod der Gans. Alternativ zum Modaltest kann eine Einbettung in ein Konditional vorgenommen werden. Einen Satz wie (70) fasse man als Antezedent auf, bilde ein passendes Konsequens und erzielt so mit (72) denselben Effekt, wie in (71):

(72) „Wenn Peter bedauert, daß die Gans tot ist, dann wird er eine neue kaufen.“

Mit Hilfe dieser Tests, der Einbettung unter einen Modaloperator und der in ein Konditional, läßt sich für einen gegebenen, isolierten Satz zuverlässig die Präsupposition bestimmen. Die beim Negationstest beschriebenen Schwierigkeiten mit Ambiguitäten und polaren Elementen treten hier nicht oder zumindest nicht im gleichen Ausmaß wie beim Negationstest auf. Verbleibende Ambiguitäten können dabei durch geschickte Kombination verschiedener Tests ausgeräumt werden.

Als Definition für das Phänomen Präsupposition, etwa in der Form „Eine Präsupposition ist eine Implikation eines Satzes, die sogar unter einem Modaloperator bestehen bleibt.“ eignen sich allerdings sämtliche Tests *nicht*. Tests helfen nur für *isolierte* (im leeren Kontext befindliche) Sätze wie (73) zuverlässig das Vorhandensein von Präsuppositionen zu testen:

(73) „Peter bedauert, daß sein Auto defekt ist.“

Unter dem Einfluß von Kontextinformationen können Präsuppositionen dagegen ein gänzlich anderes Verhalten an den Tag legen:

(74) „Wenn Peter ein Auto besitzt, dann bedauert Peter, daß sein Auto defekt ist.“

Im Gegensatz zu (73) präsupponiert (74) (das ist der gleiche Satz in einem erweiterten Kontext) nicht mehr das Vorhandensein eines Autos, das Peter gehört.

Präsuppositionsinduzierer Mit Hilfe der beschriebenen Präsuppositionstests ist man nun in der Lage, nach systematischen Zusammenhängen zwischen lexikalischen Elementen oder auch syntaktischen Strukturen und dem Vorkommen von Präsuppositionen zu suchen. Diese werden Präsuppositionsinduzierer genannt, weil sie (im leeren Kontext) stets zusammen mit einer Präsupposition auftreten.⁴³

Zu den am häufigsten genannten Präsuppositionsinduzierern gehören Eigennamen, definite Beschreibungen in Subjektsposition, Quantoren, aspektuelle Verben (*beginnen, aufhören, fortfahren ...*), präsuppositionale Adverbien (*nur, auch, sogar ...*), faktive Verben (*erkennen, bedauern, wissen ...*), implikative Verben (*gelingen, es fertigbringen ...*), Spaltsätze und Spannsätze, bestimmte lexikalische Elemente wie *beschuldigen, kritisieren* oder *Junggeselle*, irreale Konditionalsätze, temporale Nebensätze und W-Fragen (Ergänzungsfragen).

Gerade für W-Fragen, die im Zusammenhang mit Dialogsystemen von besonderem Interesse sind, ist umstritten, ob es sich tatsächlich um Präsuppositionsinduzierer handelt (vgl. u.a. Geurts (1995)). Piwek (1999) argumentiert, ein Präsuppositionsinduzierer erzeuge eine Lücke, die durch gegebene oder für gegeben gehaltene Information gefüllt werden soll, so wie eine Frage eine Lücke erzeuge, die vom Befragten gefüllt werden solle, weshalb er Fragen als Präsuppositionsinduzierer einstuft.

Obwohl es einleuchtend erscheint, daß (75) die Information in (76) präsupponiert,

(75) „Wer hat den Kuchen gegessen?“

(76) Jemand hat den Kuchen gegessen.

bleibt der Phänomenbereich „Fragen“ dennoch problematisch, da ein zuverlässiger Test fehlt: Einbettung unter Operatoren kommt bei Fragen nicht vor

(77) *„Ist es möglich, wer hat den Kuchen gegessen?“

(78) *„Wenn wieso/weshalb/warum/wie/wo/wann hat Peter den Kuchen gegessen, (dann) müssen wir einkaufen fahren?“

⁴³Van der Sandt (1988) und Grewendorf et al. (1988) geben ausführliche Übersichten über Präsuppositionsinduzierer. Vergleiche aber auch die neuere Arbeit Geurts (1997a).

(79) *„Es ist nicht der Fall, daß wer hat den Kuchen gegessen?“

und auch die Annahme, (81) könnte die Negation einer Frage wie (80) sein (womit wir einen Negationstest für Fragen hätten), erscheint intuitiv unsinnig, da der Negationsoperator nicht *in* die Frage aufgenommen, sondern die Frage unter dem Negationsoperator eingebettet werden sollte:

(80) „Wer hat den Kuchen gegessen?“

(81) „Wer hat den Kuchen nicht gegessen?“

Ebenso gut wie (81) könnte (82) als Verneinung von (80) angenommen werden. Das charakteristische Merkmal einer Frage ist es, nach einer bestimmten Information zu verlangen. Da dieses Merkmal in (82) in sein Gegenteil verkehrt wird, indem die in (80) explizit verlangte Information explizit gegeben wird, könnte man statt (81) ebenso gut (82) als Negation der Frage (80) betrachten:

(82) „Peter hat den Kuchen gegessen.“

Das Projektionsproblem Eine weiteres Problem, das eine Theorie der Präsuppositionen beantworten können muß, stellt sich unmittelbar, wenn, statt der bisher betrachteten einfachen, komplexe Sätze und das Verhalten der Präsuppositionen in ihnen untersucht werden. Zunächst soll (83) zeigen, daß eine Einbettung von Präsuppositionen offenbar keine maximale Tiefe hat, sondern einzig durch den Verarbeitungsaufwand des Hörers/Lesers begrenzt ist:

(83) „Peter weiß, daß Hans es bedauert, daß Werner verärgert darüber ist, daß es seine Zigarre war, die der Bundeskanzler, bevor er Marias VW auslieh, zur Hälfte rauchte.“

Der (zugegeben) komplexe Satz (83) besteht aus Einbettungen einfacher Sätze, deren Präsuppositionen zusammen die in (84)–(90) wiedergegebene Menge der Präsuppositionen von (83) ausmacht:

(84) Es gibt einen Bundeskanzler.

(85) Maria besitzt einen VW.

(86) Werner besitzt eine Zigarre.

(87) Der Bundeskanzler lieh Marias VW.

(88) Der Bundeskanzler rauchte vor (87) die Zigarre (siehe (86)) zur Hälfte.

(89) Werner ist verärgert über (88).

(90) Hans bedauert (89).

Aus solchen Beispielen, in denen die Präsuppositionen der einfachen Teilsätze unverändert an der Oberfläche des Matrixsatzes auftauchen, entwickelte sich die Annahme, daß generell für komplexe Sätze gelte, daß deren Präsupposition die Summe der Präsuppositionen der konstituierenden (einfachen) Teilsätze sei (zuerst vorgeschlagen in Langendoen und Savin (1971)).

Diese „kumulative Hypothese“ genannte Annahme ließ sich sehr bald durch eine Vielzahl von Beispielen widerlegen. Wie schon für (74) gezeigt, gilt auch für (91), daß die Präsupposition (92) in den Fällen (93)–(95) nicht mehr existiert:

(91) „Marias Learjet macht Lärm.“

(92) Maria besitzt einen Learjet.

(93) „Es ist nicht der Fall, daß Maria einen Learjet besitzt, und daß Marias Learjet Lärm macht.“

(94) „Es ist möglich, daß Maria einen Learjet besitzt und daß Marias Learjet Lärm macht.“

(95) „Falls Maria einen Learjet besitzt, (dann) macht Marias Learjet Lärm.“

Das Problem, wann und warum Präsuppositionen aus eingebetteten Teilsätzen an den Matrixsatz vererbt werden, ist als Projektionsproblem⁴⁴ bekannt und stellt (neben der Frage nach dem Mechanismus der Präsuppositionsentstehung) die zweite zentrale Frage dar, auf die eine Theorie der Präsuppositionen eine Antwort zu geben in der Lage sein muß.

Um unabhängig von der konkreten technischen Ausarbeitung zu verstehen, was sich in (93)–(95) im Gegensatz zu (91) abspielt, muß man sich vor Augen halten, daß Präsuppositionen in der Kommunikation der Steigerung der Effizienz dienen. Durch ihre Verwendung vermeidet ein Sprecher unnötige Aufzählungen von Fakten und verweist direkt auf Informationen, die Sprecher und Hörer gleichermaßen bekannt sind. Wenn z.B. in einer Familie, in der es eine Tochter gibt, die Maria heißt, und die vor vielen Jahren einen Learjet von ihren Eltern geschenkt bekam, Marias Mutter gegenüber Marias Vater (91) äußert, dann setzt sie zu Recht voraus, daß Marias Vater die Existenz des Düsenflugzeugs bekannt ist. Die (91) eigene Präsupposition, daß Maria einen Learjet besitzt, könnte aber nun gegenüber jemandem, der Maria und ihren Jet nicht kennt, unangebracht sein und führt dazu, daß Marias Mutter statt (91) wohl lieber (96) äußert.

(96) „Maria besitzt einen Learjet, und Marias Learjet macht Lärm.“

⁴⁴da man auch von der *Projektion* einer Präsupposition an die Satzoberfläche spricht.

In (96) wird die in (91) präsupponierte Information im Satz *ausgesagt* und nicht mehr vorausgesetzt, da die Sprecherin davon ausgeht, die in (91) vorausgesetzte Information sei ihrem Gegenüber nicht bekannt.

Für (93) gilt nun eine ähnliche Feststellung: Die Präsupposition aus dem zweiten Teilsatz, daß Maria einen Learjet besitzt, widerspricht der Aussage des ersten Teilsatzes und „verschwindet“ deshalb.⁴⁵ Eine Präsupposition, kann nur dann an die Oberfläche eines Satzes projiziert werden, wenn dort Platz für sie ist – d.h. nicht eine Aussage bereits dort sitzt, die genau das aussagt, was die Präsupposition ausdrückt – und dort keine Widersprüche entstehen. In (93)–(95) aber wird ja genau der Information, die mit der Präsupposition aus dem eingebetteten Teilsatz an die Oberfläche gelangt, im ersten Teilsatz widersprochen. Um einen Widerspruch zu verhindern, kann die Projektion der Präsupposition also in diesen Fällen nicht stattfinden.

Verarbeitung von Präsuppositionen Präsuppositionen lassen sich im Rahmen der DRT (vgl. 4.2.1) auf der Grundlage des inzwischen mehrfach weiterentwickelten (z.B. Geurts (1993); Geurts und Van der Sandt (1997)) Ansatzes von Van der Sandt (1988, 1992) beschreiben. Fortschritte im Bereich dieser Diskursrepräsentationstheorie haben damit auch immer mittelfristig positive Folgen für die Verarbeitbarkeit linguistischer Phänomene an Dialogschnittstellen.⁴⁶

Im Gegensatz zur ursprünglichen DRT wird im Ansatz Van der Sandts für eine Äußerung φ zunächst eine Zwischenstruktur erzeugt, die sogenannte *Sentence DRS* (SDRS). Im Gegensatz zu einer normalen DRS sind in einer SDRS anaphorische Bindungen nicht aufgelöst worden und die Teile von φ , die später anaphorisch aufgelöst werden müssen, sind in einer Substruktur der SDRS, strikt getrennt vom nicht-anaphorischen Material, gespeichert.

Eine SDRS K ist somit ein Triple der Form $\langle U(K), \text{Con}(K), A(K) \rangle$, wobei $U(K)$ und $\text{Con}(K)$ die Menge der Diskursreferenten bzw. die Menge der Bedingungen sind. $A(K)$ ist eine Menge von SDRSen, auch die A -Struktur von K genannt, die jene Teile von K bezeichnet, die anaphorisch sind (vgl. S. 114). Ein Beispiel:

(97) „His dog sleeps.“

$$(98) \left\langle \underbrace{\emptyset}_{U(K)}, \underbrace{\{\text{sleep}(x)\}}_{\text{Con}(K)}, \underbrace{\left\langle \underbrace{\{x\}}_{U(K')}, \underbrace{\{\text{dog}(x), \text{poss}(y,x)\}}_{\text{Con}(K')}, \underbrace{\{\{y\}, \emptyset, \emptyset\}}_{A(K')} \right\rangle}_{A(K)} \right\rangle$$

(97) enthält zwei Anaphern: die Subjekt NP „dog“ und das „he“ in „his“. Diese sind Teil der A -Struktur von (97). $U(K)$ bleibt damit leer und $\text{Con}(K)$ enthält nur das

⁴⁵Ausdrücke wie „verschwinden“ sind in diesem Zusammenhang nur intuitiv zu verstehen, die Präsupposition in (93) löst sich nicht in Luft auf, sondern läßt sich andernorts nieder, siehe dazu S. 137.

⁴⁶Zum Einsatz einer Variante der DRT in der automatischen Sprachverarbeitung siehe Geurts (1990a); Herzog und Rollinger (1992).

„sleep(x)“.⁴⁷ $A(K)$ ist, per Definition, eine SDRS und kann also selbst wieder eine A-Struktur enthalten. Diese rekursive Struktur wird benötigt, um die Einbettung verschiedener Anaphern untereinander zu repräsentieren und damit die Reihenfolge der Verarbeitung (am tiefsten eingebettete Anaphern zuerst) zu garantieren. $A(K)$ ist also ein Triple $\langle U(K'), Con(K'), A(K') \rangle$, wobei $U(K')$ für (97) nur aus x und $Con(K')$ aus „dog(x), poss(y,x)“ besteht, da das „he“ in „his dog“ eingebettet ist. Diese Einbettung wird dadurch repräsentiert, daß erst in $A(K')$ ein y für „he“ eingeführt wird.

Weiterhin muß im Ansatz von Van der Sandt die Definition des *Projektionspfades* neu eingeführt werden.

Ein Projektionspfad ist die Route, die eine Anapher aus ihrer eingebetteten Position in Richtung auf die Haupt-DRS nehmen kann, um einen Referenten zu finden. Anders ausgedrückt: mit dem Bild des Projektionspfades wird die Menge der zugänglichen Referenten beschrieben, wobei das Bild des Pfades dadurch gerechtfertigt ist, daß es eine strikte Ordnung auf dieser Menge gibt, die durch die Entfernung⁴⁸ des Referenten zur Anapher bestimmt wird.

Nach der separaten Speicherung der unaufgelösten anaphorischen Elemente in der neu eingeführten A-Struktur muß im zweiten Schritt die SDRS mit der DRS (das ist die Repräsentation des Vortextes) zu einer neuen DRS vereinigt werden. Dabei werden die in der A-Struktur enthaltenen Anaphern gebunden und die A-Struktur wird aufgelöst. Grundsätzlich werden bei diesem Prozeß zuerst die *eingebetteten* Anaphern verarbeitet. Nur so ist gewährleistet, daß zu dem Zeitpunkt, da die *einbettende* Anapher verarbeitet wird, die *eingebettete* schon gebunden und die Gefahr, daß eine freie Variable entstehen könnte, somit gebannt ist.

Ein anaphorischer Ausdruck kann sowohl an einen passenden, zugänglichen Antezedenten gebunden werden oder (und das ist neu), wenn es sich um eine Präsupposition handelt, dann kann diese in einer zugänglichen DRS akkommodiert werden. Dabei gilt generell: Bindung hat Vorrang vor Akkommodation.

Als potentielle Antezedenten kommen dabei aus der Menge der zugänglichen Diskursreferenten alle die in Frage, die mit der Anapher verträglich sind (bei einer Identifikation also keine Inkonsistenzen erzeugen). Von diesen potentiellen Antezedenten wird stets der präferiert, der sich am dichtesten an der Anapher befindet: Der Diskursreferent aus der A-Struktur wird mit dem Antezedens-Diskursreferenten gleichgesetzt und dieser übernimmt die gesamte mit dem anaphorischen Diskursreferenten verbundene Information (z.B. die Conditions).

Wird für einen anaphorischen Diskursreferenten kein passendes Antezedens gefunden und handelt es sich um eine Präsupposition, dann wird versucht, zu akkommodieren. Dieser Versuch findet erst dann statt, wenn alle Möglichkeiten zur Bindung überprüft und verworfen wurden.

Akkommodiert werden kann grundsätzlich an allen zugänglichen und geeigneten Po-

⁴⁷Solange die anaphorischen Bindungen noch nicht aufgelöst wurden, darf x frei in $Con(K)$ vorkommen.

⁴⁸„Entfernung“ ist als Anzahl der Einbettungsschritte zu verstehen.

sitionen, d.h. in allen DRSen, die die ursprüngliche Position des anaphorischen Diskursreferenten subordinieren (s. a. 116). Die Zugänglichkeit einer Position ist aus der Theorie heraus definiert, die Eignung einer Position soll eventuell durch die Akkommodation entstehende Widersprüche ausschliessen. So ist die Haupt-DRS in (93) für „Maria’s Learjet“ zwar eine zugängliche, aber keine geeignete Position, da die resultierende DRS in sich widersprüchlich wäre.⁴⁹ In der so entstandenen Menge potentieller Akkommodationspositionen für eine Präsupposition wird die präferiert, die sich am weitesten von der ursprünglichen Einbettungsposition der Anapher entfernt, d.h. am dichtesten an der Haupt-DRS befindet. In Beispiel (93) wird, da die Haupt-DRS keine geeignete Position darstellt, die Präsupposition unterhalb der Haupt-DRS akkommodiert. Da aber nur Präsuppositionen auf Ebene der Haupt-DRS nach außen sichtbar sind, entsteht der Eindruck, die Präsupposition sei „verschwunden“.

Der Vorgang der Akkommodation selbst besteht darin, daß der anaphorische Diskursreferent mit all seinen Conditions von der Stelle des ursprünglichen Auftretens weg (also aus der A-Struktur heraus) an die präferierte Akkommodationsposition bewegt wird. Dadurch entsteht an dieser Akkommodationsposition ein neuer, als Antezedent verfügbarer Diskursreferent, während die A-Struktur um eine Ebene reduziert wird.

4.2.5 Implikaturen

Neben Phänomenen wie Ellipsen (4.2.3), Anaphern (4.2.2) und Präsuppositionen (4.2.4), die alle aufgrund ihrer Kontextabhängigkeit zur Pragmatik gezählt werden, müssen weitere pragmatische Phänomene wie z.B. Konversationsimplikaturen berücksichtigt werden. Ein System, das hilfreiche Antworten gibt, indem es über die möglichen Pläne und Ziele des Fragenden inferiert, wäre für den Benutzer ein enormer Gewinn, wie Cohen (1996) zeigt:

- (99) Customer: „Where are the steaks you advertised ?“
Butcher: „How many do you want ?“

Obwohl es einige Versuche gibt, z.B. die informative Überbeantwortung von Ja/Nein-Fragen zu modellieren (Wahlster et al., 1983) oder den Mechanismus hilfreicher Antworten auf modale und hypothetische Fragen zu klären (De Roeck et al., 1991), sind Systeme die ernsthaft konversationelle Implikaturen berücksichtigen selten.

Marks und Reiter (1990) stellen z.B. zwei Systeme vor, die in der automatischen Produktion natürlicher Sprache bzw. in der automatischen Produktion von Diagrammen unbeabsichtigte konversationelle Implikaturen im Sinne von Grice (1989) zu vermeiden suchen. Das hier interessierende Textgenerierungsprogramm „FN“ erzeugt mit nicht weiter genannten Mitteln verschiedene mögliche Objektbeschreibungen, die mittels einer Präferenzfunktion auf der Grundlage dreier Prinzipien bewertet werden.

⁴⁹Zu Beschränkungen der Menge der Akkommodationspositionen, siehe Van der Sandt (1992); Ronthaler (1994).

Prinzip 1 besagt, daß keine unnötigen Komponenten verwendet werden dürfen. Wenn ein Objekt als „Auto“ beschrieben werden kann, dann ist in der Beschreibung „Auto mit vier Rädern“ die Präpositionalphrase überflüssig, weil sie zu einer konversationellen Implikatur führen kann, daß das in Frage stehende Auto durch die Eigenschaft vier Räder zu haben in irgendeiner Form besonders ausgezeichnet sei.

Prinzip 2 orientiert sich an der Prototypensemantik von Rosch (1978), indem eine lexikalische Präferenzhierarchie angenommen wird, in der ein lexikalisches Element als umso besser eingestuft wird, je ähnlicher es der sogenannten basic-level class ist.

Prinzip 3 schließlich soll bei der Suche nach der präferierten Lösung lokale Minima vermeiden helfen, indem anhand einer domänen-spezifischen Taxonomie überprüft wird, ob es kürzere Ausdrücke gibt, die das gleiche Objekt oder den gleichen Sachverhalt beschreiben. Mit Hilfe des Prinzips 3 sollen lokale Minima vermieden werden. Lokale Minima können beispielsweise auftreten, wenn Nominalphrasen wie „großes braunes zutrauliches schnalzendes Känguru aus dem Vorabendprogramm der späten sechziger Jahre“ gemäß der anderen Regeln behandelt und dabei um ein oder zwei Adjektive reduziert werden. Prinzip 3 soll hier mit der Ersetzung durch „Skippy“ für eine stärkere Reduktion sorgen. Während die Notwendigkeit der Behandlung dieser Grice'schen Phänomene in textgenerierenden Systemen sofort einleuchtet, erscheint die konkrete Umsetzung des gezeigten Ansatzes gerade in Systemen mit breitem Anwendungsbereich aufgrund der starken Voraussetzungen an domänenspezifischem Wissen derzeit nicht möglich zu sein.

Ein anderes Phänomen, das insbesondere bei der Reaktion des Systems auf Anfragen eine Rolle spielen kann, sind skalare Implikaturen (Horn, 1972, 1989).

Skalare Implikaturen bezeichnen verschiedene Ausdrücke derselben grammatischen Kategorie die eine linguistische Skala beschreiben und die in einer Richtung in einer Inklusionsbeziehung zueinander stehen:

(100) alle < die meisten < viele < einige < wenige

Äußert ein Sprecher (101), dann folgt logisch (102) aus seiner Äußerung:⁵⁰

(101) „Alle Kinder sind nach Hause gegangen.“

(102) „Einige Kinder sind nach Hause gegangen.“

Bekommt ein Hörer hingegen (102) mitgeteilt, so kann er daraus normalerweise schließen, daß nicht alle Kinder nach Hause gegangen sind, d.h. daß (101) negiert wird. Allerdings kann die Beziehung zwischen (102) und der Negation von (101) keine Implikation sein, da der Schluß von (102) auf die Negation von (101) explizit zurückgenommen werden kann. D.h. (102) kann auch als (103) geäußert werden, ohne daß diese Äußerung in sich widersprüchlich wäre (siehe (103)). Die Implikation von (101) auf (102) verhält sich hier ganz anders, wie das als ungrammatisch markierte Beispiel (104) zeigt:

⁵⁰(101) impliziert (102), so wie logisch zwingend aus der Tatsache, zehn Kekse zu besitzen, folgt, auch neun, acht, sieben usw. Kekse zu besitzen.

(103) „Einige, ja sogar alle Kinder sind nach Hause gegangen.“

(104) *,„Einige, ja sogar wenige Kinder sind nach Hause gegangen.“

D.h. der Schluß auf den, auf der Skala höher gelegenen, Bereich (von (102) auf (101)) ist eine Implikatur, da der Schluß rücknehmbar ist. Motiviert wird diese Implikatur im Grice'schen Sinne zumeist durch folgende Argumentation:

1. Der Sprecher äußert „die meisten“.
 - (a) Der Sprecher hätte „alle“ sagen können, wenn es für den Dialog angemessen gewesen wäre. „Alle“ ist informativer als „die meisten“ und aus „alle“ folgt auch „die meisten“.
 - (b) „Alle“ ist nicht aus irgendwelchen anderen Gründen wie z.B. der Länge des Ausdrucks etc. weniger geeignet als „die meisten“.
2. Dem Sprecher sind die genannten Eigenschaften von „die meisten“ bewußt und er hat trotzdem „die meisten“ geäußert, d.h. er möchte, daß der Hörer inferiert, daß der Sprecher nicht weiß ob „alle“ der Fall ist oder daß der Sprecher sogar weiß, daß „nicht alle“ der Fall ist.

D.h. aus der Abwesenheit einer stärkeren Aussage wird in Anwesenheit einer schwächeren Aussage geschlossen, daß die stärkere Aussage nicht gilt.

Ähnliche Erwägungen lassen sich für andere Konstruktionen anstellen. So kann für eine Äußerung wie „Ich glaube, daß ...“ gefolgert werden, daß der Sprecher nicht über gesicherte Informationen verfügt, weil er das stärkere „Ich weiß, daß ...“ nicht verwendet hat.

4.2.6 (Indirekte) Sprechakte

Sprechakte sind Veränderungen der Welt die direkt und unmittelbar durch sprachliche Äußerungen eintreten (Austin, 1962; Searle, 1975b,a). Ein bekanntes Beispiel ist die Veränderung, die ein Brautpaar erfährt, wenn es vom Pastor oder vom Standesbeamten mit den Worten „Hiermit erkläre ich euch zu Mann und Frau“ angesprochen wird. Weitere Beispiele sind die Veränderung der Welt durch das (gesprochene) Urteil eines Richters („Der vierte Senat stellt fest, daß das in Frage stehende Bundesgesetz verfassungswidrig ist und deshalb die auf seiner Grundlage getroffenen Entscheidungen nichtig sind.“) oder durch eine Eröffnungsrede („Hiermit erkläre ich die 50. Frankfurter Buchmesse für eröffnet.“). Voraussetzung für die Wirksamkeit eines Sprechaktes ist eine bestehende Konvention oder soziale Norm, sowie die ausreichende Legitimation desjenigen, der den Sprechakt vollzieht.

Sprechakte können an natürlichsprachlichen Schnittstellen immer dann auftreten, wenn neben einer reinen Retrievalfunktion auch andere Aktionen (z.B. in einem Verwaltungssystem einer Behörde) möglich sind. Probleme bestehen dann vor allem in

der ausreichenden Identifikation des (sprachlich) Handelnden, der expliziten Rückmeldung über die Folgen der getätigten Eingabe usw. Für ein Auskunftssystem wie OSIRIS wird die Unterscheidung verschiedener⁵¹ Sprechakte vermutlich nicht von Bedeutung sein.⁵²

Indirekte Sprechakte sind zumeist Aufforderungen an den Hörer und verlangen von diesem eine aktive Interpretation der Äußerung. Perrault und Grosz (1988) geben als Beispiel für indirekte Sprechakte an Dialogschnittstellen folgenden Beispieldiskurs zwischen Nutzer und System:

(105) User: „Ich würde gerne wissen, welche Abteilungen in 1999 einen Umsatz von mehr als 2 Millionen Mark machten.“

(106) System: „Die Softwareabteilung.“

(107) User: „Berücksichtige die Personalentwicklung in den letzten fünf Jahren. Kann ich eine tabellarische Darstellung je Monat bekommen?“

(105) ist oberflächlich als Aussage des Nutzers über seine Befindlichkeit zu interpretieren. Natürlich möchte der Nutzer (zumindest in einem zwischenmenschlichen Dialog) seine Äußerung (105) als Aufforderung verstanden wissen, die Bezeichnungen der spezifizierten Abteilungen zu nennen. Daß (105) etwas anderes ist als nur eine Aussage über einen persönlichen Zustand ist an der Unangemessenheit der beiden folgenden Antworten auf (105) zu erkennen:

(108) „Das tut mir leid für Sie.“

(109) „Ok, verstanden.“

Das Verständnis von (105) als Aufforderung an den Hörer kann aber nicht auf der linguistischen Form von (105) beruhen, da es sich weder um eine Frage, noch um einen Befehl handelt. Die Leistung des Hörers besteht in Reaktion auf (105) darin zu erkennen, daß der Sprecher aussagt, daß ihm eine Information fehlt, daß der Sprecher aussagt, daß er diese Information gerne hätte und als Hörer auf diese Aussagen so kooperativ zu reagieren, daß er versucht, die fehlende Information bereitzustellen. (106) zeigt mit der Systemreaktion eine angemessene Reaktion auf (105), die zudem auch ein Beispiel für elliptische Phänomene (siehe 4.2.3) an einer natürlichsprachlichen Schnittstelle ist.

Der zweite Satz aus (107) stellt wiederum einen indirekten Sprechakt dar, da es sich oberflächlich um eine Ja/Nein-Frage handelt, tatsächlich aber der Benutzer den

⁵¹Natürlich liegt jeder Eingabe des Benutzers an OSIRIS ein Sprechakt zugrunde – OSIRIS interpretiert aber aufgrund der Systemkonzeption jede Anfrage als Aufforderung, Literatur zu suchen.

⁵²Vgl. aber Rousseau et al. (1996), die aufbauend auf den Erkenntnissen aus dem Umfeld der Sprechakttheorie die Entwicklung eines Modells für das Kommunikationsverhalten von Softwareagenten in einem Multi-Agenten System betreiben.

Wunsch äußert, die im ersten Satz in (107) spezifizierten Informationen in einer bestimmten Aufbereitung zu sehen. Auch hier gilt wieder, daß der Hörer aus der Frage des Sprechers nach der *Möglichkeit* einer Handlung auf den Wunsch des Sprechers, daß die Handlung ausgeführt werden solle, schließen muß und sich zudem so kooperativ verhalten soll, daß er den erkannten Wunsch des Sprechers zu erfüllen versucht.⁵³

4.2.7 Indirekte Reaktionen

Ein besonders auffälliges Phänomen im menschlichen Dialog sind indirekte Reaktionen:

(110) **A:** „Es klingelt.“

(111) **B:** „Ich stehe gerade unter der Dusche.“

(112) **A:** „In Ordnung, ich mache auf.“

Dieser Beispieldialog ist allein aufgrund der direkten Bedeutung der Äußerungen nicht verständlich. Dahinter steht in (110) die Aufforderung des Sprechers an den Hörer, auf das Klingelzeichen zu reagieren und die Tür zu öffnen (siehe *Indirekte Sprechakte*, 4.2.6), in (111) gibt der Hörer Gründe an, weshalb er der in (110) geäußerten Aufforderung zu handeln nicht nachkommen kann und in (112) bestätigt der Sprecher, daß er diese Gründe akzeptiert und die in (110) angesprochene Handlung selbst ausführt.

Obwohl solche Sequenzen theoretisch sehr interessant erscheinen, spielen sie dennoch für Dialogschnittstellen derzeit keine Rolle: Solange Computersysteme nicht eigenständig, bewußt und vor allen Dingen überzeugend in sozialen Kontexten agieren, so lange wird die Maschine nicht durch den Benutzer mit beispielsweise indirekten Reaktionen konfrontiert werden. Natürlich sind verschiedene Möglichkeiten denkbar, dem System indirekte Reaktionen unter bestimmten, vom Entwickler antizipierten Bedingungen zu ermöglichen. Solche Versuche, das System durch kontextabhängige Lösungen mit ausgeprägtem ad hoc Charakter in die Lage zu versetzen, bestimmte indirekte Reaktionen zu erkennen und angemessen auf sie zu reagieren, sind kontraproduktiv, weil sie in der Außendarstellung der Fähigkeiten des Systems irreführend sind (vgl. 7.2, 6.2.6 sowie Copestake und Sparck-Jones (1990)).

Angesichts der stark beschränkten Möglichkeiten einer tiefergehenden Modellierung der kognitiven Leistungen des Menschen kann das Ziel in Hinblick auf solche Phänomene zur Zeit nur deren explizite Zurückweisung oder die Vermeidung ihres Auftretens sein.

⁵³In OSIRIS hat jede Eingabe die Form eines indirekten Sprechaktes, da es sich stets um die Vervollständigung des Satzes „Ich suche Literatur zum Thema ...“ handelt. Dennoch kann natürlich nicht davon die Rede sein, OSIRIS könne (indirekte) Sprechakte interpretieren. Vielmehr interpretiert OSIRIS jede Eingabe als Aufforderung, Literatur zu suchen.

4.2.8 Deiktische Ausdrücke

Deiktische oder indexikalische Ausdrücke wie „nächste Woche“, „Ich“ oder „dort drüben“ sind häufig Bestandteil des zwischenmenschlichen Dialoges. Sie zeichnen sich dadurch aus, daß sie nur aus dem Kontext heraus zu interpretieren sind: Ohne Kenntnis des zeitlichen oder räumlichen Bezugsrahmens sind sie bedeutungslos, da z.B. jede Person von sich als „ich“ sprechen kann und zu jeder Zeit der aktuelle Augenblick als „jetzt“ bezeichnet werden kann. Für eine Dialogschnittstelle ergeben sich diese zusätzlichen Anforderungen insbesondere dann, wenn in heterogenen Datenbeständen (z.B. neben Texten auch Bilder und Audiodateien) gesucht wird (siehe 5.1.4 und 5.2.1): In diesem Fall ist nicht nur textuelles Material verfügbar (innerhalb dessen mit „inner-textuellen deiktischen Ausdrücken“ wie Anaphern, siehe 4.2.2 agiert wird), sondern es wird durch das System explizit ein nicht-sprachlicher, aber systeminterner Kontext eingeführt,⁵⁴ in dem zur Orientierung im Raum (in Bildern mittels „hier“, „dort“, „oben“ etc.) und in der Zeit (in Audiomaterial mittels „eben“, „vor X Minuten“, „gleich“ etc.) deiktische Ausdrücke verwendet werden. Zur Diskussion, ob es sich dabei noch um deiktische Ausdrücke oder um den neuen Typus einer multimodalen Anapher handelt, siehe (Wahlster et al., 1991; Lee und Stenning, 1998) sowie 4.2.2.

Voraussetzung für die erfolgreiche Verarbeitung deiktischer Ausdrücke an einer NL-Schnittstelle ist die Existenz eines objektiven Bezugsrahmens, auf den das System Ausdrücke wie „hier“ oder „heute“ abbilden kann. Weiterhin ist für die Systemreaktion auf Anfragen, die deiktische Ausdrücke mit Bezug auf graphische oder andere Objekte enthalten, eine Repräsentation der Eigenschaften dieser Objekte nötig (vgl. Krüger und Ronthaler (1997) zur Fehlererkennung und -behandlung bei der Referenz auf Objekte in graphischen Benutzeroberflächen von Sprachlehrsystemen).

4.2.9 Kontext

Im Sinne von Diskurskontext kann der Kontext einer Äußerung einen wichtigen Beitrag für eine angemessene Reaktion des Systems auf eine Benutzerfrage sein. So führt Hendrix (1982) als Beispiel für die Funktion des Kontextes an, daß ein Benutzer, der dem System mitgeteilt hat, daß er nur deutschsprachige Bücher zu sehen wünscht, auf eine spätere Aufforderung „Finde Bücher zum Thema Morphologie“ als Ergebnis nur deutschsprachige Bücher präsentiert bekommen sollte. Allerdings müsse diese Einschränkung des Suchergebnisses für den Benutzer erkennbar und vor allem revidierbar sein.

(Diskurs)Kontext ist zudem auch eine Möglichkeit, um Ambiguitäten in der Eingabe zu reduzieren oder aufzulösen (siehe 4.2.13). Auch kann der Kontext einer Eingabe helfen, erkannte oder vermutete Eingabefehler zu korrigieren. So stellen Carbonell et al. (1983) ein System vor, daß die Benutzereingabe „prot“ im Kontext von Disketten auf „port“ (von „dual port disk“) zurückführt und entsprechend korrigiert. Diese Korrektur erfolgt ähnlich wie die in OSIRIS, daß nämlich aus einer ungrammatischen

⁵⁴Zusätzlich zum sowieso immer vorhandenen außersprachlichen Kontext des Benutzers.

(genauer: einer nicht im Lexikon verzeichneten) Eingabe durch Ähnlichkeitsuntersuchungen wie lokale Buchstabenvertauschungen auf ein kontextuell zulässiges Wort geschlossen wird.⁵⁵

Risikoreicher ist laut Carbonell et al. (1983) die „Korrektur“ eines korrekten, aber semantisch auffälligen Wortes wie „flies“ in „Copy the flies in my directory“ zu „files“. Um bei diesem Vorgehen einen zu rigiden Eingriff des Systems in das Eingabeverhalten des Nutzers zu verhindern, ist neben einer tiefen Analyse der Nutzereingabe unbedingt eine interaktive Behandlung der Korrektur vonnöten (siehe 5.1.4). Um eine Belästigung empfindlicher Nutzer durch das System auszuschließen, erscheint die Möglichkeit, sich als Nutzer solche weitgehenden Korrekturvorschläge durch das System generell zu verbitten, hilfreich.

4.2.10 Turn taking

Turn taking, das ist der Wechsel zwischen Sprecher- und Hörerrolle im Dialog, funktioniert normalerweise so gut, daß er im zwischenmenschlichen Dialog kaum bemerkt wird. Untersuchungen von Ervin-Tripp (1979) z.B. zeigen, daß nur 5% der Redebeiträge überlappen, während die Pausen im Turn-Wechsel z.T. nur wenige $\frac{1}{10}$ Sekunden lang sind.

Auch wenn man nicht von einer Maschine erwartet, daß sie im Mensch-Maschine-Dialog dieselbe Performanz wie ein Mensch in dieser Hinsicht an den Tag legt, so bleiben dennoch zwei grundlegende Probleme. Einerseits sind die grundsätzlichen Mechanismen des turn taking auch unabhängig von der Performanz wenig verstanden, andererseits muß ein System, das Teilnehmer eines Gespräches mit zwei oder mehr untereinander kommunizierenden Menschen ist, wenigstens passiv turn taking in dem Sinne beherrschen, daß es dem Gespräch folgen und, wenn schon nicht von sich aus, dann doch auf explizite Ansprache hin adäquat inhaltlich reagieren kann.

Dabei ist zu beachten, daß die turn taking Mechanismen unterschiedlich anspruchsvoll für die Maschine sein können, je nachdem, wie stark vorstrukturierend die Domäne auf den Dialog wirkt. Den (theoretischen) Extremfall stellt hier eine ritualisierte oder liturgische Kommunikationssituation dar, in der die Sprecherwechsel genau vorgeschrieben oder (z.B. am Ende von frei gestaltbaren Passagen) eindeutig markiert sind.

In Dialogsystemen wie auch OSIRIS stellt sich die Frage nach dem turn taking nicht, weil sie bereits in der Systemarchitektur antizipiert ist. Ein turn des Benutzers geht durch einen Mausklick zuende und der turn der Maschine endet mit der vollständigen Anzeige der gefundenen Information. Die Verwendung z.B. der „Zurück“-Taste im Browser, in dem OSIRIS für den Benutzer läuft, gibt dem Benutzer zwar die Möglichkeit, die Initiative zu beliebigen Zeitpunkten zu erlangen, jedoch um den Preis der

⁵⁵Interessant ist in diesem Zusammenhang, daß die Korrekturfunktion nicht nur auf den Äußerungskontext zurückgreift, sondern an sich bereits kontextabhängig ist: Die Eingabe *dual port disks* ist im Kontext des Jahres 1983 wahrscheinlich, nicht aber heute mehr.

Störung oder des Abbruchs der Kommunikation: Eine Aktion am Browser ist gerade eine Aktion *außerhalb* des System-Dialoges – die Verwendung der „Zurück“-Taste im Browser ist eine nutzerseitige Manipulation des Dialoges und kann von OSIRIS nicht als Handlung innerhalb eines Diskurses erkannt werden. Daher wird erst mit der Möglichkeit von einerseits Rückfragen und andererseits einer inkrementellen Analyse auch komplexerer Eingaben turn taking relevant. Dann sind auch Situationen denkbar, in denen das System intervenieren könnte (*Darf ich an dieser Stelle einhaken ?*).

Voraussetzung für die Realisierung eines Phänomens wie turn taking an einer Schnittstelle ist ein grundlegendes Verständnis von Diskursstrukturen (siehe 4.2.1).

4.2.11 Einschübe

Einschübe sind Reaktionen auf Fragen, die zunächst ein anderes Thema in den Diskurs einführen, bevor die ursprüngliche Frage dann beantwortet wird:

A „Ich suche Literatur zum Thema *Morphologie*.“

B „Wissen Sie, in welchem fachlichen Kontext Sie zu *Morphologie* Literatur suchen?“

A „Muß ich das jetzt schon wissen?“

B „Nein, ich kann zu allen fachlichen Kontexten Literatur liefern.“

A „Ich suche Literatur zum Thema *Morphologie* in der Geologie.“

Alle Versuche, solche Phänomene im generellen Fall ohne Rückgriff auf zugrundeliegende allgemeine Diskursstrukturen wie Ziele und Motive der Diskursteilnehmer zu behandeln, sind zum Scheitern verurteilt, weil es keine endliche Liste solcher Einschübe gibt. Gleichzeitig handelt es sich um ein alltägliches Phänomen im Dialog zwischen Menschen, das auch im Mensch-Maschine-Dialog auftreten kann. Einschübe sind ein Beispiel für (verschachtelte) Diskursstrukturen, die durch Auswertung der Diskursrepräsentation und (im Sinne von Grosz und Sidner (1986)) Modellierung der Benutzerziele, der aktuell salienten Objekte im Diskursverlauf sowie der Struktur der Diskurs(abschnitt)intentionen erkennbar und behandelbar sind (siehe auch 4.2.1).

4.2.12 Conversational fillers

Conversational fillers sind Füllwörter wie *ah*, *yes*, *ok*, *hm* usw. Sie treten normalerweise nur in gesprochener Sprache auf und haben im wesentlichen zwei Funktionen. Auf Seiten des Sprechers ermöglichen sie, ohne konkrete inhaltliche Äußerungen zu tätigen, den turn zu halten. Conversational fillers signalisieren also den Wunsch, die Initiative im Dialog weiter zu halten. Auf Seiten des Hörers haben sie kohäsionsstiftende Funktion, indem sie bestätigen, rückversichern und Aufmerksamkeit anzeigen. Conversational fillers sollten vor diesem Hintergrund also nur unter zwei Voraussetzungen für Dialogsysteme interessant sein:

1. Es wird überhaupt gesprochene Sprache verarbeitet.
2. Der Dialog ist hinreichend frei, so daß turn taking realistisch in Frage kommt.

Beides ist für derzeit existierende Systeme nicht zutreffend. Auch im Vorzeigeprojekt für gesprochene Spontansprache, VERBMOBIL, verhält sich die Maschine völlig passiv: Erst auf Tastendruck wird sie eingeschaltet und ihr damit der turn übergeben (vgl. S. 106).

Dennoch sollte es möglich sein, conversational fillers auch in aktuell eingesetzten Dialogsystemen als Bestätigung einzusetzen. Dies wäre genau dann vorstellbar, wenn, wie für das turn taking (siehe 4.2.10) angeführt, komplexe Benutzereingaben inkrementell verarbeitet werden könnten. Dann wäre beispielsweise für jedes erfolgreich vom System erkannte subgoal ein bestätigendes *hm* oder *yes* denkbar. Ein solcher Vorschlag ist hier nur im Rahmen der Diskussion möglicher linguistischer Phänomene zu sehen. Ob es für den Benutzer auch tatsächlich angenehm und hilfreich ist, mit einem System zu arbeiten, das conversational fillers generiert, ist eine ganz andere Frage (siehe dazu auch Kapitel 6).

4.2.13 Ambiguität

Natürliche Sprache ist wesentlich gekennzeichnet durch Ambiguität, entsprechend sind ambige Ausdrücke und Konstruktionen auch an der Mensch-Maschine-Schnittstellen häufig anzutreffen. Beispiele sind Skopusambiguitäten (siehe 4.2.15), die ambige Verwendung von Konnektiven (siehe 4.2.14) sowie lexikalische (Satz (113)) oder syntaktische Ambiguitäten (Beispiel (114)):

(113) „Ich suche Bücher über den Sekretär von Gerhard Schröder.“

(114) „Ich suche Bücher über Computerlinguistik mit einem roten Einband.“

In beiden Fällen ist die Eingabe aus Sicht eines Linguisten und damit aus Sicht der Maschine (die ja über ein durch linguistische Theorien geprägtes „Sprachverständnis“ verfügt) ambig. In (113) ist unklar, ob es um ein Möbelstück oder einen Angestellten geht und in (114), einem klassischen Beispiel für PP-Attachment oder Modifikator-Ambiguität, ist unklar, was durch die Präpositionalphrase „mit einem roten Einband“ genau modifiziert wird. Für einen Menschen kommt die Interpretation, daß die „Computerlinguistik“ einen roten Einband haben könnte, natürlich überhaupt nicht in Frage. Dieser Schluß wird aber unterbewußt gezogen und beruht auf Wissen über die Welt, daß nämlich Bücher einen farbigen Einband haben können, Wissenschaften aber nicht. Da Menschen über ein sehr viel reichhaltigeres Weltwissen verfügen als Computer, sind in der zwischenmenschlichen Kommunikation die meisten der, objektiv betrachtet, ambigen Ausdrücke, aus Sicht der Gesprächspartner in der Verwendung nicht ambig: Entweder kommen den Kommunikationspartnern die Lesarten, weil aus ihrer Sicht abwegig, überhaupt nicht in den Sinn („Peter überfällt eine Bank“ mit „Bank“

im Sinne eines wetterfesten Sitzmöbels), oder aber sie sind so sehr durch den Kontext des Gespraches festgelegt, da bestimmte, normalerweise leicht zugangliche Alternativen ausgeschlossen sind (zu Kontext, siehe auch 4.2.9).

Aufgrund der Schwierigkeiten, Computer mit ausreichend Wissen ber die Welt auszustatten (siehe auch S. 105), werden auch in absehbarer Zeit Computer nicht in der Lage sein, wie Menschen Ambiguitaten durch Erfahrung („Geldinstitute werden berfallen, keine Mbel“) einschranken zu knnen. Allerdings sind Computer (eine entsprechend anspruchsvolle Eingabeverarbeitung vorausgesetzt) sehr wohl in der Lage, Ambiguitaten zu erkennen und dem Benutzer zu melden.

Dies kann aber durchaus problematisch sein, da Computer aufgrund des mangelnden Weltwissens auch dort Ambiguitaten erkennen, wo Menschen durch Erfahrung oder Kontext langst (unterbewut) disambiguiert haben. Da zudem eine adaquate Explikation aufwendig ist⁵⁶ und die Explikation von Lesarten, die Menschen (aufgrund ihres Weltwissens) fr vllig ausgeschlossen halten, strend wirken kann (siehe zu diesem Punkt auch 4.2.15), kann die Erkennung von Ambiguitaten in der Eingabe durch die Maschine hinsichtlich der Nutzerzufriedenheit unter Umstanden kontraproduktiv sein. Eine Alternative zur Explikation stellt die parallele Bearbeitung aller Lesarten und die separate Darstellung der verschiedenen Ergebnisse dar. Dies kann aber zu langen Verarbeitungszeiten fhren und sich daher wiederum negativ auf die Nutzerzufriedenheit auswirken.

Eine Mglichkeit fr den Computer, um vor allem mit lexikalischen Ambiguitaten umzugehen, ist die Verwendung von Kontextinformationen. Wenn die bisherige Interaktion mit dem Benutzer im Bereich von Finanzen, Wirtschaft und Geld stattgefunden hat, dann kann das System, wenn es vor der Entscheidung steht, welche Bedeutung es „Bank“ zuweisen soll, diese Kontextinformation als Entscheidungsgrundlage nehmen (siehe auch 4.2.9). Dies wird umso besser funktionieren, wie zuknftige Systeme in die Lage versetzt werden, einem Diskurs zu folgen und diesen in Echtzeit zu segmentieren (siehe 4.2.1). Dies ist ntig um erkennen zu knnen, welche Objekte gerade prominent sind⁵⁷ und ob das Thema („Finanzwesen“) immer noch behandelt wird, oder ob ein Wechsel in ein neues Diskurssegment und zu einem neuen Thema stattgefunden hat.

Die Folgen einer ambigen Anfrage sind je nach System unterschiedlich. In OSIRIS ist es derzeit vollkommen egal, welche Lesart in (114) gemeint ist, da die Semantik der Prapositionen radikal vereinfacht ist (siehe 3.2.8). Erst fr eine deutlich elaboriertere Semantik wrde es daher einen Unterschied machen, wie (114) interpretiert wird.

4.2.14 Disjunktive Verwendung von *und*

Die Konjunktion „und“ wird umgangssprachlich, aber auch an Mensch-Maschine-Schnittstellen hufig mit disjunktiver Bedeutung verwendet (Das-Gupta, 1987; Ogden

⁵⁶Siehe z.B. Wahlster et al. (1983) zur berbeantwortung von Ja/Nein-Fragen.

⁵⁷Vgl. Grosz (1986) zur Reprasentation und Verwendung von Fokus in einem dialogverstehenden System.

und Kaplan, 1986). So kann eine Anfrage wie

(115) „Ich suche Literatur über Frankreich und Spanien.“

nicht nur als Frage nach Literatur über beide Länder gemeinsam, sondern auch als Frage nach Literatur über Frankreich und nach Literatur über Spanien gemeint sein. Diese Interpretation wird von den Suchenden explizit bestätigt und macht für Anfragen wie (116) auch Sinn, da es nur in Ausnahmefällen möglich sein wird, in zwei Städten gleichzeitig zu leben:

(116) „Wieviele Menschen leben in Hamburg und London ?“

Daß die Konjunktion „und“ durchaus häufig an Mensch-Maschine-Schnittstellen disjunktiv verwendet wird, zeigen neben den erwähnten auch eigene Untersuchungen, die für die Schnittstelle des OSIRIS Systems durchgeführt wurden (Moranz, 1998). Dieses Phänomen wurde auch an herkömmlichen Bibliothekskatalogen beobachtet (vgl. Glöckner-Rist et al. (1989), siehe auch S. 21), die keine natürlichsprachlichen Eingaben akzeptieren, sondern auf Booleschen Operatoren basierende Ausdrücke einer Kommandosprache erwarten. Diese Beobachtung ist vermutlich eher mit den dort bereits angesprochenen Schwierigkeiten zu erklären (vgl. S. 21), die Boolesche Ausdrücke für das menschliche Verständnis darstellen.

Dabei ist der Übergang zu echten Ambiguitäten fließend: Was in (116) noch durch Weltwissen oder constraints über dem Datenbestand entschieden werden kann, ist in anderen Fällen aber eine echte Ambiguität:

(117) „Wieviele Männer und Frauen mit einem Jahresgehalt von mehr als 500.000 DM leben in Hamburg ?“

Für diese Anfrage gibt es tatsächlich *beide* Interpretationen: Die in der nach der Anzahl der Männer (unabhängig von deren Einkommen) in Hamburg und nach der Anzahl der Frauen mit einem Jahresgehalt von mehr als 500.000 DM gefragt wird und die, in der nach der Anzahl der Männern mit einem Jahresgehalt von mehr als 500.000 DM und nach der Anzahl der Frauen mit einem Jahresgehalt von mehr als 500.000 DM gefragt wird.

Ogden und Kaplan (1986) stellen in einem Experiment an einer (simulierten) NL-Schnittstelle fest, daß die Nutzer „and“ nicht wahrheitsfunktional verwenden können: Um die Vereinigungsmenge zu erhalten wurde in 30 % „and“ verwendet, während für die Schnittmenge ausschließlich⁵⁸ „and“ verwendet wurde.

Das Fazit, das Ogden und Kaplan (1986) ziehen lautet also: Eine NL-Schnittstelle kann das natürlichsprachliche „or“ sicher als logisches OR interpretieren, muß aber angesichts eines natürlichsprachlichen „and“ auf zusätzliche Informationen zurückgreifen

⁵⁸In weniger als 1 % der Fälle wurde „or“ verwendet, was Ogden und Kaplan (1986) mit explorativen Systemanfragen einzelner, besonders kreativer Personen erklären.

können, um zu einer angemessenen Interpretation zu kommen. Die logische Interpretation von „and“ muß dann natürlich auch dem Nutzer mitgeteilt werden.

Als Heuristik, wie ein natürlichsprachliches „and“ zu interpretieren ist, schlagen Ogden und Kaplan (1986) vor:

Interpretiere ein natürlichsprachliches „and“ als logisches OR, es sei denn

1. „and which“ wurde benutzt („The donkeys living in a stable and which came from Ireland“).
2. die Singularform des fehlenden Elementes wurde benutzt („The student majoring in history and math“).
3. „both“ wurde benutzt („The students majoring in both math and history“).
4. ein Relativpronomen geht dem „and“ im Satz voraus („The students that major in math and history“).

4.2.15 Quantoren

Die Verwendung von Quantoren wie „ein“, „jeder“ oder „alle“ an einer natürlichsprachlichen Schnittstelle kann zu verschiedenen Problemen führen. Zum einen tritt bei komplexen Sätzen mit mehreren Quantoren das Problem der Skopusambiguität auf:

(118) „Ich suche alle Bücher über einige Länder in denen es genau eine Amtssprache gibt.“

Für eine Anfrage wie (118) bedeutet das, daß es verschiedene Möglichkeiten gibt zu interpretieren, wieviele verschiedene Amtssprachen denn nun gemeint sind:

(119) $\exists y \exists! x: (\text{land}(y) \wedge \text{ist_amtssprache_von}(x,y))$

(120) $\exists! x \exists y: (\text{land}(y) \wedge \text{ist_amtssprache_von}(x,y))$

Je nach Interpretation sind natürlich andere Länder angesprochen und damit andere Bücher erfragt.

Skopusambiguitäten sind gut untersuchte und wohlverstandene Phänomene, die auf verschiedene Arten am Mensch-Maschine-Interface behandelt werden können.

Einerseits können die erkannten Interpretationsmöglichkeiten explizit gemacht werden und entweder vor dem Retrievalvorgang dem Benutzer zur Auswahl gestellt oder aber in mehrere, klar voneinander getrennt präsentierte Anfrageergebnisse umgesetzt werden. Problematisch an dieser Vorgehensweise ist die adäquate Explikation der Ambiguität. Den meisten Benutzern wird es nicht klar sein, daß ihre Anfrage ambig ist

und das System muß über elaborierte Methoden wie z.B. eine Textgenerierungskomponente verfügen, um die festgestellte Ambiguität z.B. natürlichsprachlich zu verdeutlichen (zur Antwortgenerierung vgl. z.B. (Luria, 1982)). Trotzdem kann auch eine allgemein verständliche Explikation zu einer Überlastung des Benutzers führen, wenn nämlich Ambiguitäten erkannt und expliziert werden, die für Menschen aufgrund ihres Weltwissens ausgeschlossen sind (vgl. 4.2.13) oder die nur aufgrund theorieinterner Annahmen auf Seiten der Eingabeschnittstelle existieren, d.h. auf strukturellen Merkmalen beruhen (in einigen Theorien z.B. auf Annahmen zur Tiefenstruktur), die für den Benutzer nicht nachvollziehbar sind.

Andererseits ist es denkbar, aus der Menge der Interpretationsmöglichkeiten mittels Heuristiken die wahrscheinlichste auszuwählen, zu prozessieren und (mit der Möglichkeit, nachträglich andere Interpretationsmöglichkeiten auszuwählen und zu prozessieren) zu präsentieren. Solche Heuristiken können z.B. die Orientierung an der Oberflächenstruktur sein, d.h. die Übertragung der linearen Folge der Quantoren aus dem tatsächlich eingegebenen Satz auf die Skopusreichweite (Perrault und Grosz, 1988). Für das Beispiel (118) würde so die Interpretation (119) gewählt. Eine andere, ebenfalls von Perrault und Grosz (1988) vorgestellte Heuristik besteht in der Vergabe von Gewichten für einzelne Quantoren. Würden die Quantoren \forall , \exists und $\exists!$ bspw. in absteigender Folge gewichtet und würde ein größeres Gewicht einen weiteren Skopus bedeuten, so würde für (118) wiederum die Interpretation (119) präferiert. Offen bleibt aber die Angemessenheit der konkreten Gewichtung eines Quantors.

Ein anderes Problem ergibt sich aus der Abbildung natürlichsprachlicher Quantoren auf logische Quantoren (siehe auch 4.2.14 für einen ähnlichen Fall bei den Konnektiven). Für Standard-Quantoren wie „alle“, „einige“ oder „einer“ ist diese Abbildung unkontrovers, jedoch ist für Quantoren wie „die meisten“, „viele“, „manche“, „wenige“, „fast keiner“ und andere diese Abbildung nicht so eindeutig.⁵⁹

4.2.16 Komposita

Wie in 3.2.6 bereits ausführlich am Beispiel von OSIRIS beschrieben, ist die Behandlung von Komposita über die bloße Feststellung einer unbestimmten Relation zwischen den im Kompositum erkannten Teilen hinaus schwierig. Dies wurde anhand einiger Beispiele aus dem Deutschen („Steinlaus“, „Mädchenschule“) demonstriert. Eine weitere Schwierigkeit besteht zudem oft in der eindeutigen Identifikation der Teile eines Kompositums. So kann das Kompositum „Planschbecken“ sowohl in „Plansch-becken“ als auch in „Planschb-ecken“ zerlegt werden. Für den Fall, daß das System korrekt von „Plansch-“ auf „planschen“ schließen kann und damit (im Gegensatz zu „Planschb-“) einen Lexikoneintrag findet, kann sich das System auf die Heuristik stützen, daß nur die Zerlegungen berücksichtigt werden, die vollständig aus dem Lexikon bekannten Teilen bestehen. Da ein Lexikon in der Praxis nie vollständig

⁵⁹Zum Einsatz generalisierter Quantoren an einem natürlichsprachlichen Interface vgl. Speelman (1993).

sein wird, ergeben sich für die genannte Heuristik Probleme, wenn das Verb „planschen“ nicht bekannt ist, d.h. die Entscheidung zwischen zwei Zerlegungen zu treffen ist, die beide auf dem System unbekanntem Wörtern basieren. Auch an dieser Stelle können Heuristiken helfen: So kann z.B. für ein Kompositum (unter der Annahme, es bestehe nur aus Bestandteilen der deutschen Sprache) ausgeschlossen werden, daß der linke Rand aus einem Teil besteht, der auf „schb“ endet, da diese Zeichenkombination (im Deutschen) am Wortende nicht auftreten kann. Nicht berücksichtigt wird dabei allerdings die Möglichkeit, daß es sich bei „Planschb“ um einen Eigennamen handeln könnte (der sehr wohl auf „schb“ enden kann). Noch schwieriger zu entscheiden sind alternative Zerlegungen, die nur aus dem Lexikon bekannten Teilen bestehen, wie z.B. „Stau-becken“ vs. „Staub-ecken“. Das isolierte Wort ist auch für den menschlichen Hörer vollständig ambig, allerdings wird ein Mensch im Kontext von „Filterbeutel“, „Saugkraft“ und „Kabelaufwicklung“ vermutlich die mögliche Zerlegung „Stau-becken“ gar nicht bewußt in Erwägung ziehen.

Es bleibt also festzustellen, daß die Analyse der Diskursstrukturen (vgl. 4.2.1) und damit das bessere Verständnis des Kontextes einer Äußerung (vgl. 4.2.9) die möglichen Ambiguitäten bei der Zerlegung von Komposita wohl reduzieren, das Problem aber nicht vollständig lösen kann (da eine Zerlegung nur deshalb als auf unbekanntem Teilen beruhend verworfen wird, da das System nicht über ein absolut vollständiges Lexikon verfügt). Die genannten Schwierigkeiten sind nicht auf das Deutsche beschränkt, sondern finden sich auch im Englischen: Perrault und Grosz (1988) geben als Beispiel „national park ranger station equipment procurement form“, ein Kompositum, dessen interne Bestandteile durch die (im Englischen obligatorische) Getrennschreibung gut erkennbar sind, dessen interne Relationen aber (wie im Deutschen) nicht unbedingt eindeutig zu erkennen sind. Auch beschreibt „Stanford Research Institute“ vermutlich ein Forschungsinstitut in Stanford, „Computer Research Institute“ aber ein Institut für Computerforschung.

Komposita gehören sowohl zur Schriftsprache wie zur Umgangssprache, sie werden oftmals spontan gebildet und sind daher an natürlichsprachlichen Schnittstellen regelmäßig zu erwarten.

4.2.17 Fazit

Ambiguitäten stellen auch für Menschen ein Verständnisproblem dar, werden aber meist über Weltwissen und Kontext soweit eingeschränkt, daß sie für den menschlichen Sprecher unterbewußt disambiguiert werden. An NL-Schnittstellen besteht also das Problem, daß das System mehr Ambiguitäten erkennt als der Mensch und diese Ambiguitäten derzeit in vielen Fällen nicht selbständig auflösen kann. Erschwerend kommt hinzu, daß das System zur Entscheidung nicht unbedingt auf die Hilfe des Benutzers bauen kann, da viele Ambiguitäten entweder für den menschlichen Hörer aufgrund von Weltwissen und Kontext plausibel ausgeschlossen sind oder aber nur aufgrund der unzulänglichen linguistischen Theoriebildung überhaupt vom System produziert werden.

Die Abbildung von non-standard Quantoren auf logische Quantoren, d.h. ihre genaue Semantik, ist noch ungeklärt. Daher sind zur Zeit keine robusten, automatischen Verfahren zur Analyse solcher Quantoren verfügbar. Ähnliches gilt für Ellipsen: Ellipsen sind theoretisch gut untersucht und verstanden, die Umsetzung in maschinentaugliche Modelle ist aber stark abhängig von Fortschritten im Bereich der Diskursrepräsentation und der Anaphernauffösung, da die (zum Verständnis notwendige) Rekonstruktion vieler Ellipsen eine tiefere Analyse des Satzes (und nicht nur das Kopieren von Satzteilen) erfordert. Der Zusammenhang mit Ambiguitäten ergibt sich aus der nicht immer eindeutigen Ergänzungsmöglichkeit des elliptischen Satzes.

Phänomene die auf einem besseren Verständnis von Diskursstrukturen (siehe 4.2.1) aufbauen sind Anaphern, Präsuppositionen, Einschübe usw. Hier sind Fortschritte nicht nur durch eine Weiterentwicklung der Theorie möglich, sondern auch durch eine bessere Umsetzung vorhandener Theorien in (experimentelle) Systeme. Neben der Entwicklung und dem Einsatz genereller Theorien zur Struktur von Dialogen kommen auch statistische Verfahren in Betracht, die auf Dialogaktkorpora arbeiten. Zur Gewinnung und Auswertung solcher Dialogaktsammlungen siehe Maier (1997); Jekat et al. (1997).

Phänomene wie Implikaturen, (indirekte) Sprechakte und indirekte Reaktionen erscheinen derzeit im allgemeinen Fall schwer lösbar. Systeme, die z.B. auf der Grundlage eines Nutzermodells (siehe 5.2.3) Annahmen über die Pläne und Ziele des Fragenden machen und deshalb sowohl auf indirekte Äußerungen wie auch auf Fragen kooperativ reagieren können, wären für den Informationssuchenden ein großer Gewinn. Um sehr allgemein formulierte Aussagen wie z.B. die Grice'schen Konversationsmaximen maschinell umsetzen zu können, ist aber zunächst eine flächendeckende Möglichkeit zur Repräsentation der Ziele und Pläne eines Nutzers nötig, zusammen mit einem Inferenzmechanismus, um aus diesen Informationen die nötigen Schlüsse für das Systemverhalten ziehen zu können. Wirklich umfassende Ansätze stehen hier noch aus. Weiterhin müssen auf Seiten der Systeme für ein kooperatives Verhalten auch die nötigen Informationen über die Domäne vorhanden sein: Allein die Erkennung der Ziele des Nutzers hilft dem System wenig, es muß auch in der Lage sein, die der Zielerreichung dienenden Informationen und Objekte zu identifizieren. Im Fall eines Auskunftssystems bedeutet dies, daß es letztlich nicht reicht, wenn das System in der Lage ist, Sprecherintentionen wie z.B. den Besuch einer lokalen Zweigstelle einer Bibliothek zu erkennen, aber über zuwenig Informationen im Datenbestand verfügt, um mit Auskünften zur Lage, den Öffnungszeiten usw. der Zweigstelle hilfreich tätig zu werden.

Ein grundsätzliches Problem bei den an einer Dialogschnittstelle zu erwartenden linguistischen Phänomenen entsteht, wenn Menschen Schwierigkeiten mit der (logisch) korrekten Verwendung bestimmter Ausdrücke (z.B. „und“ und „oder“, siehe 4.2.14) haben. In diesen Fällen kann die Maschine die menschlichen Benutzer schlecht modellieren, da es sich nicht um systematische, sondern lediglich um häufig auftretende Phänomene handelt, die zudem (auch nach dem Empfinden menschlicher Experten)

fehlerhaft sind. Hier sind Heuristiken nötig, die für die aktuelle Eingabe versuchen, die wahrscheinlichste Interpretation zu identifizieren und diese interaktiv mit dem Nutzer zu verifizieren.

Zusammenfassend läßt sich sagen, daß weitere Fortschritte insbesondere beim Verständnis von Diskursstrukturen und bei der Analyse des Äußerungskontextes nötig sind, um das Verhalten von Dialogschnittstellen in Hinblick auf natürlichsprachliche Eingaben, die einzelne, unzusammenhängende Sätze übersteigen, zu verbessern. Allerdings können umfassendere Theorien zur Dialogmodellierung noch nicht direkt voneinander profitieren, da in wesentlichen Punkten noch Uneinigkeit über Ziele, Methoden und Evaluationsverfahren bei der Dialogmodellierung herrscht (Cohen, 1996). Pragmatische Phänomene spielen zwar sofort in eine solche Erweiterung des Eingabeumfangs hinein, führen aber sofort auch zu einer drastischen Erhöhung des Aufwandes bei der Domänenrepräsentation. Die Schwierigkeiten, in ausreichendem Maße Weltwissen für die angemessene Reaktion auf die erkannten Nutzerbedürfnisse zur Verfügung zu stellen, wurden an anderer Stelle aber bereits angesprochen, so daß eine Erweiterung der Möglichkeiten bestehender Systeme wie OSIRIS in Hinblick auf pragmatische Phänomene derzeit kurzfristig nicht möglich erscheint.

5 Verbesserungen und Erweiterungen von OSIRIS

Obwohl OSIRIS deutliche Fortschritte gegenüber der in 2.3 und 3.5 genannten Kritik an Dialogschnittstellen bietet, besteht in vielen Punkten noch Erweiterungs- und Verbesserungsbedarf. Diese Punkte sollen hier unterschieden werden nach der Zugehörigkeit zur natürlichsprachlichen Eingabeverarbeitung einerseits und weiteren Systemverbesserungsmöglichkeiten andererseits.

5.1 Erweiterung der Möglichkeiten der Eingabeverarbeitung

Um die Chancen einer grundlegenden Verbesserung von im weitesten Sinne informationsverarbeitenden Systemen abschätzen zu können, lohnt ein weiterer Blick auf FhG-ISI (1998).

Den in 2.1 geschilderten Aussichten auf eine drastische Steigerung des Bedarfs an Methoden zur intelligenten Handhabung von Informationen steht eine deutliche Skepsis der in FhG-ISI (1998) befragten Experten hinsichtlich grundsätzlicher Änderungen bei den verwendeten Technologien gegenüber. Auffällig ist, daß marktfähige Lösungen auf der Grundlage eines grundsätzlich anderen oder gar vollständigen Verständnisses der kognitiven Leistungen des Menschen in absehbarer Zeit bezweifelt werden.

- So sind 38 % der Befragten davon überzeugt, daß Systeme, die die Mustererkennungsleistung des Menschen (bis hin zu Karikaturen) erreichen, nie realisiert werden. Diejenigen, die eine solche Leistung für prinzipiell möglich halten, sehen deren Verwirklichung deutlich nach 2025.
- 18 % der befragten Experten halten Systeme, die Sinnzusammenhänge wie Metaphern, Synonyme, Argumentationsmuster usw. aktiv erkennen und erklären können, für nie realisierbar.
- Große Skepsis besteht auch angesichts der Vorstellung, ein Übersetzungsgerät könne Dokumente in Büros automatisch und ohne Bedarf an menschlicher Nachbearbeitung übersetzen: 17 % halten dieses Gerät für niemals realisierbar.
- Ebenfalls 17 % der Experten hält die Entwicklung von Anlagen für unmöglich, die Zusammenfassungen oder Auszüge aus Büchern oder Dokumenten automatisch und in beliebiger Länge durchführen.
- Und schließlich sind 38 % davon überzeugt, daß Informationsdatenbanken, die durch automatisches Lernen ihr Wissen ohne Mißverständnisse ordnen, nicht realisiert werden können.

Es besteht also eine deutliche Skepsis, innerhalb der kommenden 25 Jahre zu Lösungen zu gelangen, die auf fundamentalen Durchbrüchen im Verständnis der kognitiven Leistungen des Menschen beruhen. Damit wird eine Untersuchung der Entwicklungsmöglichkeiten vorhandener Technologie umso wichtiger. Für Dialogschnittstellen wie

für die Suche in natürlichsprachlichen Texten bedeutet dies, daß auf der Basis heute verfügbarer Technologie (im wesentlichen im Bereich der syntaktischen Analyse) Verbesserungen erreicht werden müssen. Einen Überblick über die linguistischen Phänomene bei der Suche in natürlichsprachlichen Texten geben Ruge und Goeser (1998) – die ablehnenden Reaktionen von Ladewig (1998a) und Ladewig (1998b) zeigen in diesem Zusammenhang, wie wenig Problembewußtsein oft im Bereich des Information Retrieval anzutreffen ist. Umgekehrt heißt dies aber auch, daß die Möglichkeiten natürlichsprachlicher Analyse oft ungenutzt bleiben und sich die Leistung bestehender Systeme für Textretrieval oft mit einfachen Mitteln steigern lassen.

5.1.1 Erweiterung der OSIRIS-Eingabemöglichkeiten

Eine Erweiterung der Eingabemöglichkeiten auf der Basis der bisher eingesetzten Techniken macht also Sinn – sie könnte konkret durch eine Reduktion der Vorgabe *Ich suche Literatur zum Thema ...* auf der Eingabemaske erreicht werden. Diese Vorgabe hat zur Zeit in OSIRIS die Aufgabe, die mögliche Komplexität natürlichsprachlicher Benutzeranfragen zu reduzieren. Dies war sinnvoll, weil OSIRIS eine klar definierte Aufgabe (Suche nach Literatur zu einem bestimmten Thema) hatte. Als Vorteil wurde in 3.1 der deutlich eingeschränkte syntaktische Analyseaufwand genannt. Im Zuge einer Weiterentwicklung von OSIRIS macht die Reduktion dieser Vorgabe aus zwei Gründen Sinn:

1. Aus den Nutzerbefragungen und Evaluationen (siehe 6.1) ist bekannt, daß auch die (im Vergleich zu einem konventionellen OPAC deutlich einfachere) Schnittstelle von OSIRIS in den Unterscheidungen zwischen den Suchaspekten „Themen“, „Autoren“, „bestimmte Werke“, „Zeitschriften“ usw. den Nutzern Schwierigkeiten bereitet. Auch für das bestehende System muß also eine weitere Vereinfachung der Schnittstelle erreicht werden.
2. Die Integration weiterer heterogener Datenbestände (s. a. 5.2.1) führt an herkömmlichen Schnittstellen zu einer stark vergrößerten Menge an Auswahlmöglichkeiten (als extremes Beispiel siehe Abbildung 9, S. 20). Ohne eine deutlich vereinfachte Schnittstelle für den Zugriff auf die verschiedenen Datenarten werden die zusätzlichen Möglichkeiten für den Suchenden schnell nutzlos, wie Untersuchungen zum OPAC zeigen (Recker et al., 1996): Eine Erweiterung der Suchmöglichkeiten, wie sie angesichts der im System prinzipiell zur Verfügung stehenden Informationen für eine genauere Recherche wünschenswert wäre, bedeutet zugleich eine Steigerung der Komplexität der Oberfläche. „Eine unkritische Ausweitung der Suchwege und -kriterien übersteigt jedoch schnell das Differenzierungsvermögen und auch die Erwartungen des Benutzers. [...] Funktionen des OPAC, die ein 'sophisticated searching' explizit ermöglichen, werden selten genutzt und ausgeschöpft.“ (Dreis, 1994).

Die nicht-linguistischen Erweiterungsmöglichkeiten von OSIRIS wie z.B. die Integration weiterer Datenbestände werden in 5.2 thematisiert, so daß im Folgenden die Vereinfachung der Schnittstelle in Hinblick auf die bestehenden Suchaspekte beschrieben wird.

Ohne grundlegende Neuentwicklungen kann in OSIRIS derzeit grundsätzlich nach den folgenden Aspekten gesucht werden:

- Autorennamen, d.h. es wird nach der Menge der Werke gesucht, die von den genannten Autoren stammen.
- Werke zu einem bestimmten sachlichen Thema, z.B. „Abfallentsorgung“.
- Werke über eine Person, z.B. über Duke Ellington – eine Unterart der thematischen Recherche.
- Werke zu einem bestimmten Thema, das durch zeitliche oder geographische Angaben weiter eingeschränkt wird, z.B. „Abfallentsorgung in Italien“, „Buchdruck im Mittelalter“, „moderne Literatur nach 1945“, „Gesundheitsversorgung während der Kulturrevolution“ usw.
- Werke zu einem bestimmten Thema und mit bestimmten formalen Merkmalen, z.B. „mit Graphiken“, „mit Noten“, „aus dem Suhrkamp Verlag“ usw.
- Bestimmte Werke, identifiziert nach z.B. ISBN-Nummer, Signatur, Kombinationen anderer Merkmale wie Autor und Titel etc.
- Veröffentlichungen zu bestimmten Konferenzen.
- Objekte einer bestimmten Materialart, z.B. Microfiche, Zeitschrift, Tonträger, Video usw.

Die Aussage, daß nach diesen Aspekten grundsätzlich gesucht werden könne, bezieht sich nur auf die linguistische Eingabeverarbeitung, nicht aber auf die zugrundeliegende Suche in den Daten. Wenn die Auszeichnung der Werke in der Bibliothek in einem Raster erfolgt, das für den zeitlichen Bezug lediglich „Alttertum“, „Mittelalter“ und „Neuzeit“ vorsieht, kann eine Suchanfrage nach „Literatur zwischen den Weltkriegen“ auch bei aufwendiger Eingabeverarbeitung und Rückgriff auf Weltwissen (Was sind „die Weltkriege“ und was ist die durch „zwischen“ beschriebene Zeitspanne ?) nicht sinnvoll beantwortet werden. Dennoch ist es natürlich wünschenswert, die Eingabeanalyse des Systems hinsichtlich solcher Phänomene deutlich leistungsfähiger als die Retrievalkomponente auszulegen, um im Zweifelsfall das Suchinteresse des Benutzers zu Erkennen und angemessen auf die systeminternen Begrenzungen bei der Repräsentation der zeitlichen Aspekte hinweisen zu können. Wo allerdings die Grenze zu ziehen ist, bis zu der die Eingabeanalyse die Retrievalfähigkeiten des Systems überschreiten soll und wie weit dies angesichts der starken Verwendung von Weltwissen zu leisten ist, ist eine offene Frage.

Eine einheitliche Schnittstelle zu allen genannten Suchaspekten könnte durch die Vorgabe von *Ich suche ...* erreicht werden: Da die grundlegende Benutzerintention⁶⁰ bei der Interaktion mit OSIRIS nach wie vor festgeschrieben ist, ist die Erwartung an den Benutzer, diese Phrase zu ergänzen, zulässig. Durch die Verkürzung der Vorgabe muß die Eingabeanalyse nun den vom Benutzer gemeinten Suchaspekt identifizieren. Eine Möglichkeit, wie dies geschehen kann, ohne daß flächendeckende, tiefe Analyse der natürlichsprachlichen Eingaben erforderlich wäre, sind Metaregeln. Eine solche Metaregel versucht anhand der Struktur der Eingabe, einer ontologischen Analyse der verwendeten Nomen sowie einer Reihe von Schlüsselwörtern den vom Benutzer gemeinten Suchaspekt zu erkennen. Beispiele für solche Metaregeln können sein:

Autorennamen „Ich suche X von Y“, wobei X nicht ein bestimmtes Objekt wie z.B. „Homo Faber“ oder „Dynamic Patterns“ ist, sondern entweder ein Ausdruck wie „etwas“, „alles“ usw., oder ein Nomen, das ontologisch ein Objekt wie 'Schriftstück', 'Tonträger', 'Datenträger' oder ähnliches bezeichnet. Die Erkennung dieses durch X bezeichneten Objektes kann dabei durch eine Mischung aus endlicher Wortliste („Bücher“, „Werke“, „Veröffentlichungen“ ...) und Ontologie-basierter Analyse erfolgen. Y sollte im besten Fall als im Datenbestand bekannte Person erkannt werden, im schlechteren Fall als ein im Lexikon als Eigenname markierter Ausdruck. Y darf dabei zusätzlich zum Namen auch einen definiten Artikel enthalten wie in „Ich suche die Werke des Marquis de Sade“.

Für diese Regel werden einige Varianten nötig sein, wie z.B. „Ich suche von Y X“ oder „Ich suche Ys X“, wobei für X und Y die genannten Eigenschaften gelten. Es wird eventuell vorkommen, daß Benutzer ihre Eingabe so verkürzen, daß sie nur den Namen des Autors, dessen Werke sie suchen, eingeben. So könnte „Ich suche Thomas Bernhard“ gemeint sein im Sinne von „Ich suche alles von Thomas Bernhard“. Dies könnte eine Metaregel „Ich suche Y“ erkennen, wobei Y eindeutig als Eigenname identifizierbar sein muß.

Werke zu einem bestimmten sachlichen Thema „Ich suche X über Y“, wobei Y eine beliebige Nominalphrase sein kann und für X das in *Autorennamen* Gesagte gilt. Als Variante ist auch „Ich suche X zu/zur/zum Y“ möglich, wie in „Ich suche Bücher zur Entwicklung der Vereinigten Staaten von Amerika“ oder „Ich suche Bücher zu Stalin“. Wird Y als Eigenname erkannt, dann handelt es sich um die Suche nach Büchern über eine Person. Dem Namen kann ein definitiver Artikel vorangehen („die Bachmann“). Eine in die Nominalphrase Y eingebettete Präpositionalphrase kann eine weitere Einschränkung des Themas sein. Ob diese Einschränkung zeitlich oder geographisch zu verstehen ist, kann in einigen Fällen an der Präposition („zur Zeit“, „nach“) erkannt werden, in anderen Fällen aber muß das Nomen ebenfalls betrachtet werden („im Mittelalter“ vs. „im Tessin“).

Eine solche Einschränkung des Themas kann allerdings auch über ein geographisches

⁶⁰Es geht um die *Suche* nach Objekten, nicht um Ratschläge, allgemeine Auskünfte usw.

oder zeitliches Adjektiv vorgenommen werden („Ich suche Bücher zum chinesischen Kohletagebau“, „Ich suche Artikel zur mittelalterlichen Logik“), so daß die genannte Metaregel „Ich suche X zu/zur/zum Y“ auch mit einer internen Struktur für Y der Form $NP \rightarrow AP N$ existieren kann.

Thematische Suche mit formalen Merkmalen „Ich suche X von/über/zu Y mit/aus/ohne ... Z“ wobei für X und Y das für thematische Recherche Gesagte gilt und Z ein beliebiges formales Merkmal sein kann wie in „Ich suche Bücher über Duke Ellington mit Noten“ oder „Ich suche Bücher von Arno Schmidt aus dem Stahlberg Verlag“. Das allgemeine Problem an dieser Stelle ist die Ambiguität der Präpositionalphrase „aus dem Stahlberg Verlag“ (PP-attachment), die durch einen von der Metaregel unabhängigen Mechanismus aufgelöst werden muß: Werden Bücher aus dem Stahlberg Verlag gesucht, die von Arno Schmidt geschrieben wurden, oder stammt Arno Schmidt aus dem Stahlberg Verlag? Durch die Vielfalt der für den vorhandenen Datenbestand aus Büchern, Zeitschriften, Kongreßberichten und Tonträgern möglichen formalen Merkmale wird jede disambiguierte Präpositionalphrase über Y hinaus als einschränkendes formales Merkmal angenommen werden müssen.

Bestimmte Werke Bestimmte Werke, identifiziert nach z.B. ISBN-Nummer, Signatur oder Kombinationen anderer Merkmale wie Autor, Titel usw. können mit folgender Metaregel identifiziert werden: „Ich suche X von Y“ wobei X ein Buchtitel o.ä. ist und Y ein Eigenname wie in „Ich suche 'Homo Faber' von Max Frisch“. Dies ist aber nur eine von vielen verschiedenen Möglichkeiten, bestimmte Werke zu suchen. Deshalb wird auch eine sehr allgemeine Metaregel wie „Ich suche X“ nötig sein, in der X einzig dadurch gekennzeichnet ist, daß es (auch apositorisch gereichte) formale Merkmale wie ISBN-Nummer, Signatur oder ähnliche auf ein Objekt ausgerichtete Merkmale enthält.

Bei der Unterscheidung zwischen der Suche nach einem bestimmten Werk und der nach einer Menge von Büchern zu einem Thema kann letztlich nur auf die Verwendung bestimmter Suchmerkmale (ISBN-Nummern etc.) geachtet werden, da eine rein strukturelle Unterscheidung zwischen einer thematischen Suche wie „Ich suche Bücher von Oliver Sacks über Neuropsychologie“ und „Ich suche Bücher von Ingeborg Bachmann über Heidegger“ nicht zu leisten ist. Daß die erste Anfrage zum Gesamtwerk des Autors führt und letztere zu genau einer Veröffentlichung, ist aus der Struktur der Anfragen nicht zu erkennen.

Proceedings „Ich suche X zu/zur Y“, wobei X vermutlich eine relativ enge Variationsbreite hat und „die Proceedings“, „die Tagungsbeiträge“, „die Konferenzveröffentlichungen“ oder ähnliche, eindeutig auf Tagungen verweisende Ausdrücke annehmen kann. Y wird hingegen die Bezeichnung der Konferenz sein und durch einfache pattern wie „ABC'97“, „Jahrestagung ...“, „Konferenz ...“ u.ä. gekennzeichnet sein. Dabei stellt die Suche nach Proceedings natürlich einen Sonderfall der „Suche nach

bestimmten Werken“ dar, der aber besonders eindeutig zu erkennen ist und daher durch besondere Regeln zu behandeln sein sollte.

Materialart Die Suche nach Objekten einer bestimmten Materialart ist kein isolierter Suchaspekt, sondern stets gekoppelt an eine der vorangegangenen Metaregeln. Die Materialart drückt sich dabei meist sofort im Nomen aus („Ich suche Bücher/Tonträger über ...“), wenn nicht bewußt die unspezifische Form „Veröffentlichungen“ oder einfach ein Quantor („alles“, „etwas“) verwendet wird. In Fällen, in denen der Benutzer explizit mit einem materialunspezifischen Nomen („Veröffentlichungen“) sucht, kann dies als Signal im Grice’schen Sinne gewertet werden (siehe 4.2.5), verschiedene Materialarten als Ergebnis anzubieten. Die Materialart kann aber auch als präpositionaler Zusatz an eine zunächst materialunspezifisch formulierte Anfrage angefügt werden: „Ich suche alles von Lutoslawski auf CD“. In beiden Fällen ist es eine endliche Liste von Materialbezeichnern, die an den genannten Stellen auftreten kann.

Probleme bei der Eingabeanalyse In vielen Fälle werden Ambiguitäten der Präpositionalphrase eine eindeutige Analyse verhindern (vgl. Langer (1996)):

(121) „Ich suche Bücher über das Exil von Thomas Mann“

(122) „Ich suche Bücher über Mode im Großformat“

(123) „Ich suche Bücher über die Musik von Thomas Mann“

(124) „Ich suche Bücher über die Musik von Boris Vian“

(121) kann sowohl *Bücher des Autors Thomas Mann*, die sein Exil zum Thema haben meinen, als auch ein Werk, das sich mit dem *Exil des Thomas Mann* (der in dieser Lesart nicht notwendig ein Autor sein muß) befaßt. Die menschliche Tendenz, in (122) und (123) die Präpositionalphrase als direkte Tochter der Verbalphrase zu interpretieren, beruht auf Weltwissen. In (122) wird das Wissen über Mode und Bücher dafür sorgen, daß es Bücher in verschiedenen Formaten, z.B. *Taschenbuchformat* oder *Großformat* gibt, Mode aber nur in *Übergrößen*. Mit genügend Aufwand sind diese Inferenzen auch von einem System zu leisten, das über ein reichhaltig ausgestattetes Lexikon oder eine separate Ontologie verfügt.

(123) wird für Menschen aufgrund des Wissens disambiguiert, daß Thomas Mann ein Autor war, der wie z.B. Boris Vian vielleicht selber musiziert hat, im Gegensatz zu diesem als Musiker aber nicht bekannt wurde und daher vermutlich über Musik *geschrieben* hat. Hingegen ist Boris Vian sowohl als Autor wie auch als Trompeter in Erscheinung getreten, was (124) für den Kenner wiederum so ambig wie (121) macht. Was für den menschlichen Experten aber bereits schwierig zu interpretieren ist, kann von der Maschine nicht oder zumindest nicht flächendeckend erwartet werden.

In anderen Fällen kann die Eingabeanalyse aufgrund der Verwendung eines durch das System nicht eindeutig zu interpretierenden Nomens behindert werden:

(125) „Ich suche Bücher über chinesische Tonfiguren im mittleren Reich.“

Ob das „mittlere Reich“ nun zeitlich oder geographisch zu interpretieren ist, kann oftmals nur ein Fachwissenschaftler entscheiden. Ein Bibliothekssystem, in dem naturgemäß Objekte zu *allen* Wissensgebieten gespeichert werden (zur Problematik des Lexikons in dieser Situation siehe 3.2.3), wird in der Praxis nicht über vollständiges Wissen in allen Fachgebieten mit einer für Fachwissenschaftler befriedigenden Granularität erreichen können.

Weiterhin treten Probleme auf, wenn Benutzer die Systemfähigkeiten überschätzen und statt identifizierender nominaler Ausdrücke stark vom Weltwissen abhängige Beschreibungen eingeben, z.B. statt des eindeutigen „Triptychon von Max Frisch“ in (126) die (sachlich korrekte) Beschreibung „der letzte Roman von Max Frisch“ in (127):

(126) „Ich suche 'Triptychon' von Max Frisch.“

(127) „Ich suche den letzten Roman von Max Frisch.“

(128) „Ich suche 'Fidelio' von Beethoven.“

(129) „Ich suche Beethovens einzige Oper.“

Ein anderes Beispiel ist die eindeutig zu interpretierende Anfrage (128), die für den menschlichen Kenner aber natürlich auch als (129) formuliert werden kann.

Im Extremfall sind auch für Menschen nicht mehr eindeutig zu interpretierende Eingaben denkbar. In (130) wird die Existenz eines Preisträgers präsupponiert – es stellt sich aber auch dem menschlichen Hörer die Frage, ob der Preis 1998 überhaupt vergeben wurde, d.h. ob die Präsupposition tatsächlich Bestand hat. In (131) werden zwei definite Beschreibungen kombiniert, von denen die erste auch für Menschen nicht eindeutig zu beantworten ist (wenn „Größe“ im übertragenen Sinne und nicht als „Körpergröße“ verstanden wird):

(130) „Ich suche die Werke des Irmgard-Heilmann-Preis-Trägers 1998.“

(131) „Ich suche den letzten Roman des größten lebenden deutschen Schriftstellers.“

Dies sind allerdings keine für die vorgeschlagene Erweiterung der Eingabemöglichkeit spezifischen Probleme. Natürlich können auch bereits jetzt in OSIRIS und vergleichbaren Systemen solche nicht oder nur theoretisch behandelbaren Anfragen auftreten.

Andere Probleme treten auf, wenn verschiedene materialspezifische Eigenschaften der gesuchten Objekte zur Einschränkung der Suche verwendet werden. So gibt es für eine Tonaufnahme im Gegensatz zu einem Buch weitere beteiligte Personen, besondere Techniken oder Ausstattungsmerkmale, die erfragt werden können:

(132) „Ich suche 'Fidelio' mit Fischer-Diskau“.

(133) „Ich suche 'Fidelio' mit Dolby-Surround“.

(134) „Ich suche 'Fidelio' mit Beiheft“.

Diese zusätzlichen Eigenschaften werden mit Hilfe bestimmter Präpositionen ausgedrückt, die über Materialgrenzen hinweg aber ganz unterschiedlich verwendet werden können, so daß eine detaillierte Repräsentation der Eigenschaften der im System verzeichneten Objekte nötig ist:

(135) „Ich suche 'Homo Faber' von Frisch“.

(136) „Ich suche 'Fidelio' von Masur“.

(137) „Ich suche 'Casablanca' von Warner Brothers“.

Mit der Präposition „von“ wird in (135) für ein Buch der Autor erfragt, in (136) für eine Oper der Interpret und in (137) für einen Film der Besitzer der Verleihrechte. Schließlich sind beinahe beliebig viele Fälle denkbar, in denen ungenaue oder extrem verkürzte Eingaben der Nutzer in nur sehr schwer für die Maschine interpretierbare Anfragen münden:

(138) „Ich suche alles von Haffner zu Bismarck.“

(139) „Ich suche alles zu Haffner über Bismarck.“

(140) „Ich suche Haffner zu Bismarck.“

(141) „Ich suche von Bismarck zu Hitler.“

(142) „Ich suche das Manifest.“

In (138) wird „zu“ vermutlich im Sinne von „über“ gebraucht, jedoch ist dies auch für Menschen keine wirklich eindeutig zu interpretierende Eingabe. So können auch Werke eines adeligen Herrn *Von Haffner zu Bismarck* gemeint sein, oder es könnte sich eine Zeitangabe darin verbergen („Von Bush zu Clinton: Amerikanische Außenpolitik im Wandel“).

(139) ist sowohl als Frage nach Literatur von Dritten zum Werk von Haffner über Bismarck als auch als Frage nach dem Werk Haffners selbst zu interpretieren.

(140) könnte die verkürzte Frage nach dem Werk des Adligen *Von Haffner zu Bismarck* sein. Für einen Datenbestand, in dem auch nach weiteren Eigenschaften von Personen (z.B. ihrer Adresse) gefragt werden kann (siehe 5.2.1), würde (140) zunächst aber als Frage nach der Person selbst interpretiert, d.h. je heterogener der Datenbestand ist, desto vielfältigere Sucheingaben sind zu erwarten. (140) kann natürlich aber auch als Frage nach einem Werk von Herrn Haffner über Herrn Bismarck gemeint sein.

Die unklare und verkürzte Eingabe (141) kann sinnvoll nur als die Frage nach dem Werk von Haffner mit dem Titel „Von Bismarck zu Hitler“ interpretiert werden und (142) schließlich fragt mit verkürzter Redeweise nach dem berühmten (und deshalb eben auch verkürzt benennbaren) Werk von Marx mit dem Titel „Das Kommunistische Manifest“.

Zusammenfassend läßt sich sagen, daß je umfassender die Eingabemöglichkeiten in einem System ausfallen, umso größer natürlich auch die Möglichkeiten des Benutzers sind, für das System schwer oder gar nicht zu interpretierende Eingaben zu tätigen. Vor diesem Hintergrund wird es wichtig, daß die Interaktion des Benutzers mit dem System kein „Simultanspiel“ im Sinne der Spieltheorie ist, in dem das System einzig auf der Grundlage der isolierten Suchanfrage und mit den Mitteln des systeminternen Wissens versucht, zu einer angemessenen Reaktion zu kommen. Vielmehr ist es wünschenswert, durch Rückfragen des Systems an den Benutzer eine kooperative Situation zu schaffen, in der Ambiguitäten gemeinsam aufgelöst werden und so das System von der alleinigen Verantwortung der korrekten Auswahl aus verschiedenen Interpretationsmöglichkeiten entlastet wird (siehe 5.1.4).

5.1.2 Erweiterung der linguistischen Komponenten in OSIRIS

Bei den in 5.1.1 gezeigten Anfragen sind zusätzliche Informationen insbesondere zu Nomen notwendig, um über deren mögliche Verwendungsweise als geographische Bezeichnung oder Personennamen entscheiden zu können. Wie könnten diese Informationen nun in OSIRIS genutzt werden ?

Gegenüber der Abbildung 29 auf Seite 60 kann das Lexikon um weitere, spezifische Informationen angereichert sein, wie in Abbildung 31 illustriert:

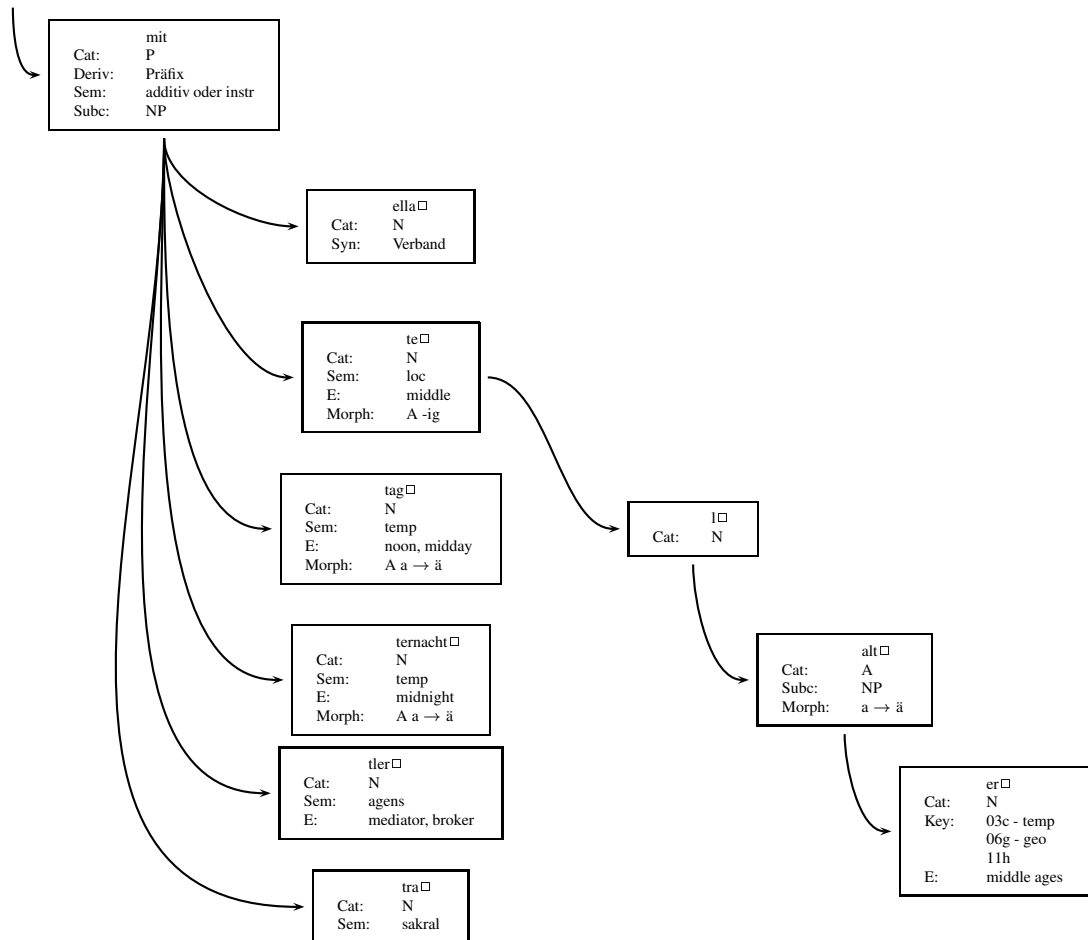


Abbildung 31: Erweiterter, annotierter Lexikonausschnitt

Neben den herkömmlichen Angaben zu Wortart, Semantik (*Sem*), morphologischen Besonderheiten (*Morph*), Synonymen (*Syn*) und möglicherweise Übersetzungen (*E*) sind auch Informationen zum Subkategorisierungsrahmen (*Subc*) sowie zu Einträgen in Schlüssel Tabellen (*Key*) möglich.

Der Subkategorisierungsrahmen gibt Auskunft über den vom Eintrag erwarteten rechten Kontext und leistet damit Entscheidungshilfe bei der Unterscheidung zwischen Argument oder fakultativer Ergänzung eines Elementes.

Mit *Key* versehene Einträge bezeichnen die Elemente, die Mitglied einer sogenannten Schlüssel tabelle der Systematik der Bibliothek sind (s. a. Abbildung 32, S. 164). Mitglieder einer Schlüssel tabelle stellen in gewisser Hinsicht Zeiger auf bestimmte Teile der Systematik dar. So kann *Kalter Krieg* in den Politikwissenschaften auf eine Notation mit Literatur zu einem bestimmten Zeitraum des 20. Jahrhunderts verweisen, während *Hildesheim* auf eine Notation mit Literatur bezüglich einer politisch oder geographisch definierten Entität zeigt.

Die zugrundeliegenden Informationen liegen für die Systematik in Tabellenform vor, sind aber nicht flächendeckend oder systematisch erstellt worden. D.h. sie spiegeln,

wie bereits das Auftreten der Klassen in der Systematik, die lokalen Bedürfnisse der Bibliothek wider. So können bspw. beim geographischen Schlüssel im Bereich „Steuerrecht“ 14 verschiedene geographische Objekte unterschieden werden, die angefangen bei „Frankreich“ über „Benelux“ und „Osteuropa“ bis hin zu „Tanzania“ und „übrige Welt“ die verschiedensten Abstraktionsstufen vereinen. Diese Aufteilung resultiert aus den unterschiedlichen Interessenschwerpunkten: Offenbar gibt es ein Interesse am französischen Steuerrecht, das sich in detaillierter Literatur zu diesem Thema niederschlägt, während das Interesse an diesem Thema in Bezug auf osteuropäische Länder eher gering ausfällt. Hingegen scheint es einen Schwerpunkt zum Steuerrecht in Tanzania zu geben, da dieses Land durch einen Schlüssel besonders hervorgehoben wird. Es ist einleuchtend, daß das Interesse in der „Kunstgeschichte“ wieder ganz anders ausfällt, so daß innerhalb einer Bibliothek für einen Aspekt kein einheitlicher Schlüssel verwendet wird, sondern einer, der in bestimmten Notationsgruppen Gültigkeit hat.

Diesen Schlüsseln gilt es eine gemeinsame Grundlage zu geben, auf die sie einheitlich abgebildet und auf der einfache Inferenzen gezogen werden können. Eine solche gemeinsame Grundlage stellt eine flächendeckende Ontologie geographischer Entitäten dar, bei der die Teil-Ganzes-Relation für Vergleichsmöglichkeiten zwischen unterschiedlichen Objekten sorgt. Eine solche Ontologie ist ohne großen manuellen Aufwand aus online verfügbaren Quellen zu erstellen und könnte wie exemplarisch in Abbildung 32 auf Seite 164 dargestellt aussehen.

Die in dieser angereicherten Schlüsseltabelle enthaltene Information kann z.T. ins Lexikon integriert werden, indem den jeweiligen Lexikoneinträgen die zu ihnen gehörigen Schlüsselkennungen angefügt werden. Die Informationen zu den Schlüsseln werden dann bei der Verarbeitung einer Suchanfrage auf der Wissensbasis berücksichtigt, z.B. wenn ein Benutzer nach „Steuerrecht in Ostafrika“ fragt und das System entweder als bestmöglichen Treffer die (speziellere) Klasse „Steuerrecht in Tanzania“ liefert oder aber eine Rückfrage startet mit dem Hinweis, eine Klasse „Steuerrecht in Ostafrika“ sei nicht verfügbar, aber spezifischere Information zum Steuerrecht in einem Teil Ostafrikas. Umgekehrt kann aufgrund möglicher Inferenzen über die Teil-von-Relation die Suchanfrage „Steuerrecht in den Niederlanden“ mit der Klasse „Steuerrecht in Benelux“ beantwortet werden, wenn es keine Klasse zum Steuerrecht in den Niederlanden gibt.

Zusätzlich können Schlüsselinformationen aber auch bereits bei der Eingabeanalyse die Identifikation von Präpositionen unterstützen. Angenommen, eine Benutzeranfrage lautete *Weinbau in Spanien*. Aus der Anfrage an sich kann nur entnommen werden, daß der Benutzer Literatur zum Thema *Weinbau* sucht, wobei *Spanien* eine durch *in* ausgedrückte Spezifikation dieses Themas ausdrückt. Wie genau aber die Semantik von *in* hier zu interpretieren ist, bleibt völlig unklar.

geo			
Polen polnisch Bereich 1: Schlüssel A	Osteuropa osteuropäisch Bereich 2: Schlüssel C	Europa europäisch	
Ungarn ungarisch Bereich 1: Schlüssel B Bereich 3: Schlüssel F	Osteuropa osteuropäisch Bereich 2: Schlüssel C	Europa europäisch	
Tanzania tanzanisch Bereich 1: Schlüssel D	Ostafrika ostafrikanisch	Afrika afrikanisch Bereich 4: Schlüssel M	
Deutschland deutsch Bereich 1: Schlüssel C	Westeuropa westeuropäisch Bereich 2: Schlüssel B	Europa europäisch	
Niederlande niederländisch Holland holländisch Bereich 1: Schlüssel E	Benelux Bereich 3: Schlüssel T	Westeuropa westeuropäisch Bereich 2: Schlüssel B	Europa europäisch
Schweiz schweizerisch eidgenössisch Helvetia helvetisch Bereich 1: Schlüssel F	Westeuropa westeuropäisch Bereich 2: Schlüssel B	Europa europäisch	

Abbildung 32: Exemplarisch angereicherte Schlüsseltabelle

Aufgrund des eingeschränkten Einsatzgebietes von OSIRIS sowie der beschränkten Möglichkeiten, automatisch Lexikoneinträge mit Semantikinformatoren zu erzeugen (vgl. z.B. Riloff und Shepherd (1997), Winiwarter (1996)), wird in OSIRIS zur Zeit Semantik nur für Präpositionen und hier auch nur einheitlich als „wird-modifiziert-durch“ verwendet. Aber auch mit einer reichhaltigeren Semantik, insbesondere auf

dem schwierigen Gebiet der Semantik räumlicher Präpositionen (Maienborn, 1990a, 1991), wäre nicht viel gewonnen, ohne weiteren Aufschluß über den Charakter von *Spanien*. Handelt es sich um eine Spezifizierung der Methode (*Weinbau in jungen Eichenfässern*), der Zeit (*Weinbau in der Frühen Neuzeit*), des Ortes (*Weinbau in der Normandie*) oder der Umstände (*Weinbau in der Krise*)? Mit einem Vermerk im Lexikon beim Eintrag für *Spanien* (Key: 09f-geo) der besagt, daß *Spanien* hinsichtlich der Schlüsselnummer 09 die Schlüsselkennung *f* erhält und dem Zusatz, daß die Schlüsselnummer 09 geographische (oder auch geopolitische) Aspekte schlüsselt, kann für die ursprüngliche Anfrage des Benutzers entschieden werden, daß *in Spanien* offenbar eine lokale Einschränkung darstellt.

Zukünftig reicherer Analyseoutput Um diese möglichen zusätzlichen Informationen in OSIRIS nutzbar zu machen, ist ein neuer, reichhaltigerer Outputcontainer als der auf Seite 71 in Abbildung 30 gezeigte nötig:

Language:	D
OriginalInput:	Abfallverbrennungsanlage
Add1:	verbrennungsanlage
Add1a:	Abfall
Add2:	anlage
Add2a:	Abfall
Add2b:	verbrennung
Spez:	emissionsfrei
SpezAdd1:	frei
SpezAdd1a:	emission
Mod:	Norditalien
ModAdd1:	italien
ModAdd1a:	Nord
Key:	italien
KeyType:	geo
KeyValue:	01g, 03h, 04H, 06j, 07k, 09o, 12h
ModRel:	part-of
Add1aSyn:	Müll
Add2aSyn:	Müll
EAdd2:	arrangement, facility, plant, grounds, investment
EAdd2a:	waste, refuse, rubbish, garbage, trash
EAdd2b:	burning, combustion, cremation, burn
ESpezAdd1a:	emission, issue
EModAdd1:	italy
EModAdd1a:	north

Abbildung 33: Zukünftig möglicher reicher Container der Analysekomponente

Die Einträge *Search* bis *ModAdd1a* sind Ergebnis der Wortformenreduktion sowie der heuristischen Kompositazerlegung (vgl. 3.2.6). Dabei wird stets der größtmögliche rechte Rand eines Kompositums verwendet, der als eigenständiges Wort dem Lexikon

bekannt ist (Einträge *Add1* und *Add2*). Der dabei abgetrennte linke Teil wird als zusätzliches Suchkriterium ebenfalls mitgeliefert (Einträge *Add1a*, *Add2a* und *Add2b*).

Die Einträge *Key*, *KeyType* und *KeyValue* in der Outputstruktur werden aus den angeereicherten Schlüsseltabellen gewonnen und den jeweiligen Lexikoneinträgen angefügt. In dem der Abbildung 33 zugrundeliegenden Beispiel wird bei der Bearbeitung des Lexikoneintrages von *Italien* erkannt, daß *Italien* in einigen Bereichen der Systematik einen Schlüssel darstellt (*Key: italien*). Weiterhin wird direkt aus dem Lexikon ausgelesen, für welchen Schlüssel *Italien* mit welcher Schlüsselkennung behandelt wird: *KeyValue: 01g, 03h, 04H, 06j, 07k, 09o, 12h*. Und schließlich ist neben der Tatsache, daß ein bestimmter Eintrag als Schlüssel fungieren kann und daß dieser Schlüssel bestimmte Kennungen bekommen kann, auch verzeichnet, welcher *Aspekt* denn überhaupt geschlüsselt wird, hier also ein geographischer Aspekt.

Der Eintrag *ModRel* ist ein erster Versuch, die mittels der Semantik von *in* erkennbare Relation zwischen *Abfallverbrennungsanlage* und *Norditalien* zu beschreiben. Aufgrund der Tatsache, daß *italien* ein Eintrag in der geographischen Schlüsseltabelle ist (*KeyType: geo*), kann auf eine *part-of* Relation geschlossen werden. Allein aus der Semantik von *in*, ohne die Information im Lexikon, die ursprünglich aus der Schlüsseltabelle stammt, wäre dieser Schluß nicht möglich.

Die beiden Einträge *Add1aSyn* und *Add2aSyn* sowie die übrigen Einträge mit englischen Übersetzungen sind Beispiele dafür, wie Informationen aus anderen Lexika ins OSIRIS-Lexikon integriert werden können.

Um die hier gezeigten, zusätzlichen Informationen in OSIRIS zu nutzen ist lediglich eine Überarbeitung der Schlüsseltabelle notwendig. Diese muß manuell vervollständigt und auf Konsistenz geprüft werden. Nach dieser (nicht allzu umfangreichen) Vorarbeit können die Daten aus dem Bibliothekskatalog in eine neue OSIRIS-Wissensbasis geladen werden, die nun anhand der Schlüsselinformation wie gezeigt eine Modifikation durch einen geographischen Aspekt erkennen kann. Auf der Grundlage einer sorgfältig ergänzten Schlüsseltabelle kann dies natürlich auch dann geschehen, wenn in der Anfrage des Benutzers ein geographischer Aspekt genannt wird, der ursprünglich in der Bibliothek nicht als Schlüssel bekannt war (z.B. „Belgien“), der aber durch die Teil-Ganzes-Relation der erweiterten Schlüsseltabelle eindeutig einem Schlüssel zugeordnet werden kann (z.B. „Benelux“).

5.1.3 Anfragebeispiele

In diesem Abschnitt werden einige Anfragen aus den OSIRIS-Log-Files kommentiert wiedergegeben. Sie sollen das bislang Gesagte unterstreichen und hätten dazu auch Beispiel für Beispiel an den entsprechenden Stellen genannt werden können. Eine gesammelte Wiedergabe bietet hingegen den Vorteil, sich einen Überblick über die verschiedenen Eingabephänomene zu machen. Aus technischen Gründen liegen keine lückenlosen Informationen über den gesamten Zeitraum des OSIRIS-Einsatzes vor, auch erfolgte die Auswahl der gezeigten Beispiele rein subjektiv. Groß- und Klein-

schreibung sowie orthographische Fehler in den Eingaben wurden originalgetreu übernommen. Alle Beispiele sind Vervollständigungen der Vorgabe *Ich suche Literatur zum Thema ...* für die thematische Recherche, die stärker formalen Suchaspekte „Autor“, „Zeitschrift“ usw. wurden nicht betrachtet. Weitere Beispiele, die als linguistische Phänomene in diesem Zusammenhang nicht interessant sind, aber besonders deutlich die Leistungsfähigkeit von OSIRIS zeigen, finden sich in 3.3.

Die folgende Liste an **Eigennamen** beweist zunächst einmal mehr die Notwendigkeit einer besonderen Behandlung komplexer Eigennamen in einem Bibliothekssystem (siehe auch 3.2.9):

- (143) Ephialtes, Pippin III., Jelinek, Gustav Stresemann, Karl der Große, Adam von Trott, Pamela Anderson, Mozart, benjamin britten, ovid, schimanski, carl maria von weber, JAKOB MICHAEL REINHOLD LENZ, dvorak

Welche Intention der Suchende aber tatsächlich bei der Suche hatte, läßt sich nicht mehr ermitteln: OSIRIS interpretiert die Eingabe eines Eigennamens als Gegenstand der thematischen Suche natürlich als Aufforderung, Literatur *über* eine bestimmte Person zu suchen, jedoch wissen wir aus den Untersuchungen zu OSIRIS (vgl. 6.1), daß Benutzer die Aufteilung der Suchaspekte oft nicht verstanden und einfach an einer beliebigen Stelle die unterschiedlichsten Eingaben tätigten. Durch eine Vereinheitlichung der Schnittstelle, wie in 5.1.1 vorgeschlagen, würde dieses Problem gelöst – allerdings um den Preis, daß andere, in OSIRIS derzeit klar zu erkennende Suchaspekte nun aufwendiger analysiert werden müßten.

Von besonderem Interesse sind die zahlreichen Interaktionssequenzen, die den oft explorativen Charakter der Benutzereingaben deutlich machen (in (144) eine Sequenz von drei aufeinanderfolgenden Anfragen eines Benutzers):

- (144) Dick + Doof
Laurel + Hardy
Stan Laurel Und Oliver Hardy

Dabei handelt es sich vermutlich um eine (nachträglich nicht weiter zu analysierende) Mischung aus

- für den Suchenden unbefriedigenden Systemreaktionen, die zu weiteren Anfragen führen.
- Überprüfung des ungewohnten Konzeptes einer automatisch expandierten und interpretierten Suchanfrage, das den Anspruch erhebt, Reformulierungen (wie sie an klassischen Systemen üblich und notwendig sind) überflüssig zu machen.
- spielerischer Erforschung der Systemmöglichkeiten.

Dabei zeigt sich auch, daß Suchende durchaus die in 5.1.1 postulierten Eingabeformen benutzen, allerdings in OSIRIS noch an einem System, das auf diese Eingaben gar nicht ausgelegt ist und ganz andere Möglichkeiten der Suche bietet:

(145) Dold

Dold und Mathematik
 Albrecht Dold als Autor in der Mathematik
 dold als autor mathematik

Die Anfrage-Sequenz in (145) wäre statt der elaborierten Reformulierungen für die *thematische Suche* ein einfacher Wechsel zu *Autoren* sinnvoller gewesen.

Als Beispiele für die **Einschränkung eines Themas auf einen zeitlichen Aspekt** finden sich in den Log-Files die in (146) genannten Eingaben:

(146) die Zerschlagung der SPD 1933
 situation der frau in den fuenfziger jahren
 Gesundheit im zweiten Weltkrieg
 hoerspiel ab 1945
 FLUGBLATTPROPAGANDA IM 2. WELTKRIEG

Dabei wurden die zeilenweise getrennten Eingaben in (146) von verschiedenen Nutzern getätigt und stehen in keiner Beziehung untereinander.

In (147) sind Beispiele für die **Einschränkung eines bestimmten Themas auf einen geographischen Aspekt** mit Hilfe einer Präpositionalphrase aufgeführt:

(147) regionalplanung in osnabrück
 Asylpolitik in der BRD
 die Schlacht im Teutoburger Wald
 charta von athen
 Ghettoproblem in Amerika
 Textilhändler aus dem Münsterland
 Burgen im Münsterland
 ausgrabungen in italien
 Kinderfernsehen in der DDR
 Ratsherren der Stadt Braunschweig
 Turm zu Babel

(148) der schlanke japanische Staat
 der japanische Staat

(149) Schengener Abkommen

Dabei sind die Eingaben in (148) Beispiele für die Formulierung eines geographischen Aspektes mit Hilfe eines Adjektivs. Ob letztlich auch (149) hierher gehört, ist diskussionswürdig, da man den geographischen Aspekt im Ausdruck „Schengener Abkommen“ auch als besondere Bezeichnung eines singulären politischen Ereignisses interpretieren kann. In dieser Hinsicht wären das „Schengener Abkommen“, das „Abkommen von Camp David“, der „Maastrichter Vertrag“, „die Schlußakte von Helsinki“ usw. keine Themen mit geographischen Aspekten, sondern lexikalisierte Bezeichnungen für andernfalls abstrakt gebliebene Verwaltungsakte. Da es sich bei „Abkommen“, „Vertrag“ usw. um sehr abstrakte Ausdrücke handelt ist es sinnvoll, z.B. „Schengener Abkommen“ als Eigennamen und nicht als die Einschränkung eines Themas um einen geographischen Aspekt zu betrachten.

Ein Problem stellen Fälle dar, die auf den ersten Blick wie eine Einschränkung aussehen:

(150) Das Volkslied bei Herder

Die mit (150) gesuchte Literatur soll aber vermutlich Herders Auffassung oder Beitrag zum Volkslied behandeln, d.h. es sind Bücher von Herder zum Volkslied gefragt wie auch Bücher von Dritten über das Verhältnis von Herder zum Volkslied.

Zu der bereits mehrfach in dieser Arbeit angesprochenen Möglichkeit, Boolesche Operatoren und Trunkierung an Dialogschnittstellen zu verwenden (vgl. u.a. 2.1.2 und 2.2), finden sich im OSIRIS-Log-File ebenfalls Beispiele. Obwohl versucht wurde, dem Benutzer durch Erklärungen, Beispiele und nicht zuletzt durch die Vorgabe von *Ich suche Literatur zum Thema ...* deutlich zu machen, daß eine natürlichsprachliche Eingabe und keine Stichwortsuche von ihm erwartet wird, finden sich dennoch häufig reine Stichwörter in den Log-Files:

- (151) Infektionskrankheiten, parasitäre Krankheiten
 Schneewittchen Märchendeutung
 internet kommunikation marketing
 unterrichtsverlauf motivation

(151) gibt die verschiedenen Anfragen unterschiedlicher Nutzer wieder, die Eingaben in (152) wurden in einer Sequenz von nur einem Nutzer eingegeben:

- (152) seele
 seele frauen
 seele frau weiblichkeit
 psychologie der frau seele
 frau seele
 frauenseele maennerseele

Trunkierungsversuche finden sich ebenfalls. Hier die Anfragesequenz eines Suchenden:

- (153) umweltkrankheit?
 umweltkrank
 umweltkrank?
 umwelt?
 umweltkrank?

Auch die in 2.1.2 zitierte These, daß Benutzer **Boolesche Operatoren** logisch falsch verwenden, d.h. `and` auch im Sinne von `or` verwenden, läßt sich mit etwas gutem Willen aus den Daten ersehen. Daß diese Interpretation nicht sicher ist, sondern den Willen des Betrachters, das in Frage stehende Phänomen in den Daten zu sehen voraussetzt, liegt daran, daß das, was der Suchende tatsächlich meinte, nicht mehr zu rekonstruieren ist:

- (154) frauen und computer
 (155) hypertext und hypermedia

Es erscheint aber nicht abwegig, daß mit (154) nach Literatur zum Umgang von Frauen mit Computern gesucht wird, daß es auf jeden Fall um den Zusammenhang zwischen Frauen und Computern geht. Dies würde im Booleschen Sinne auf ein logisches `and` als Interpretation für das natürlichsprachliche „und“ hinauslaufen. Hingegen erscheint das natürlichsprachliche „und“ in (155) eher als logisches `or` zu verstehen zu sein, weil die beiden Begriffe „hypertext“ und „hypermedia“ so eng miteinander verwandt sind, daß eine `and`-Verknüpfung unwahrscheinlich erscheint. Dies ist aber, wie schon erwähnt, Interpretation des Betrachters.

Hingegen ist (156) ganz eindeutig der Versuch, mit einem Negationsoperator an der natürlichsprachlichen Schnittstelle zu arbeiten:

- (156) Zinsänderungsrisiken not Wechselkursrisiken

Andere Eingaben deuten auf den Versuch der Verwendung von **elliptischen Ausdrücken** hin. Die Beispiele in (157) entstammen einer Sequenz von Benutzereingaben, wobei die zweite, dritte und vierte Eingabe offenbar jeweils elliptische Nachbesserungen am ursprünglich eingegebenen *Algebra in einer Variablen* sein sollen. So etwas ist in der derzeitigen OSIRIS-Version natürlich nicht zu verarbeiten:

- (157) Algebra in einer Variablen
 einer variablen
 einer veränderlichen
 in einer veränderlichen
 algebra in one
 algebra in variable
 polynome in einer variablen

Oft drängt sich der Eindruck bei der Durchsicht der Log-Files auf, daß die Suchenden die Fähigkeiten des Systems überschätzen:

(158) Welche Wörterbücher zur Medizin sind vorhanden ?

(159) Wie weit ist mein Kind entwickelt?

(160) Wie funktioniert der Zelltod ?

Dabei ist zu beachten, daß die Beispiele (158) bis (160) von unterschiedlichen Benutzern stammen und in keinem Zusammenhang zueinander stehen.

Andererseits sind gerade in den Eingabe-Sequenzen explorative, z.T. fast schon verspielte Muster zu erkennen, die die Frage aufwerfen, ob die Benutzer sich überhaupt um die Systemfähigkeiten kümmern, oder ob sie nicht vielmehr einfach Dinge eingeben, die ihnen in den Sinn kommen (siehe auch 2.2):

(161) freundschaft bei kindern
freundschaft und kind
freunschaft bei kindern
freunschaft bei kindern
freundschaft bei kindern
freundschaft und kind
feindschaft und kind
feindschaft
feindschaft und kind
kind und kegel
kegel

oder auch

(162) erkältungskrankheiten
erkaeltung
husten
schnupfen
grippe
grippale infekte

Bei anderen Anfrage-Sequenzen gewinnt man den Eindruck, daß der explorative Charakter eher einer gewissen Hilflosigkeit entspringt. Die systematische Variation der Eingabe in (163) bei konsequenter Vermeidung syntaktisch wohlgeformter Nominalausdrücke deutet auf ein Mißverständnis über die Beschaffenheit sinnvoller Eingaben in OSIRIS hin:

- (163) Analyse Lebensmittel Qualitätsmanagement Lebensmittel
 Qualitätsmanagement
 Qualitätsmanagement
 Qualitätsmanagement Lebensmittel
 Qualitätsmanagement
 Qualitätsmanagement Hygiene
 Hygiene Kontrolle
 Qualitätskontrolle
 Lebensmittel
 Lebensmittel Qualität

Bei vielen Anfragen fragt man sich, ob die Benutzer tatsächlich erwarten, eine relevante Auskunft für ihre Frage am Online-Katalog einer Universitätsbibliothek zu bekommen. In (164) ist zu diesem Punkt die Anfragesequenz eines einzigen Benutzers wiedergegeben:

- (164) babyausstattung
 baby
 einzelhandel für kinderbedarf
 Einzelhandelsbetrieb
 Einzelhandel

Bei anderen Beispielen wird aber sehr schnell deutlich, daß Benutzer wohlgeformte natürlichsprachliche Anfragen formulieren, die aufgrund eines besonderen Stils die Eingabeanalyse vor Probleme stellen:

- (165) Europa im Blick der Statistik

5.1.4 Benutzerinteraktivität

Interaktion mit dem Benutzer, d.h. die explizite Rückfrage des Systems in unklaren Situationen, wird aus zwei Gründen unabdingbar sein für eine Weiterentwicklung von OSIRIS oder eines mit OSIRIS vergleichbaren Dialogsystems.

Die Auflösung von Ambiguitäten in der Eingabe (seien diese „echt“ im linguistischen Sinne oder nur aufgrund der Einschränkungen der Eingabeanalyse des Systems entstanden, siehe 4.2) ist z.T. nur durch Rückfrage beim Benutzer möglich. Dabei führt die Minimierung der Eingaberestriktionen (wie in 5.1.1 beschrieben) bei der Identifikation des Suchaspektes zu zusätzlichen Ambiguitäten (z.B. beim PP-Attachment), die nur gemeinsam mit dem Benutzer zu lösen sind. Auch wird die größere Freiheit an der Eingabeschnittstelle in manchen Fällen zu für die Maschine schwer interpretierbaren deskriptiven Nominalphrasen führen („der diesjährige Literaturnobelpreisträger“ vs. „Günther Grass“). Eine Möglichkeit des Systems, auf diese Ausdrücke zu reagieren, könnte in einer Metaregel bestehen, die Anfragen wie (130) und (131) (hier wiedergegeben als (166) und (167)) als Fragen nach allen bzw. einem bestimmten Werk eines nur durch eine deskriptive Nominalphrase bezeichneten Autoren erkennt.

(166) „Ich suche die Werke des Irmgard-Heilmann-Preis-Trägers 1998.“

(167) „Ich suche den letzten Roman des größten lebenden deutschen Schriftstellers.“

Auf diese beschreibenden Teile könnte dann im Stile von „Eliza“ (Weizenbaum, 1967) mit einer Rückfrage reagiert werden, im Fall von (166) z.B. mit „Bitte nennen Sie mir den Namen des Irmgard-Heilmann-Preis-Trägers 1998“ oder für (167) mit „Ich kenne sehr viele Schriftsteller, bitte nennen Sie mir den Namen des größten lebenden deutschen Schriftstellers“. Wie jedoch bereits erwähnt (siehe 4.2.7), sind partielle und ad hoc Erweiterungen der Systemfähigkeiten in Bezug auf solche Eingaben kontraproduktiv, weil sie beim Benutzer den Eindruck erwecken können, das System könne tatsächlich die nötigen Inferenzen leisten und habe eine Vorstellung davon, was ein „großer deutscher lebender Schriftsteller“ ist (siehe auch 7.2). Eine Rückfrage sollte daher viel eher deutlich machen, daß das System erkannt hat, daß ein Objekt oder eine Person identifiziert werden soll. Darüber hinaus muß die Rückfrage aber auch deutlich machen, daß das System keine Überlegungen zum Inhalt dieser Beschreibungen angestellt hat, sondern ganz grundsätzlich die gewünschte Identifikation nicht leisten kann: „Bitte nennen Sie eine eindeutige Bezeichnung oder einen Namen für Ihren Ausdruck 'der größte lebende deutsche Schriftsteller'“.

Die in 5.1.1 beschriebenen unterschiedlichen Eigenschaften der verschiedenen Objekttypen (Verfasser und Verlag für Bücher, Komponist, Interpret, ggfs. Mitwirkende und Plattenfirma für Musikstücke etc.) und deren natürlichsprachliche Benennung sorgen bei zunehmendem Ausbau des Datenbestandes für steigende Komplexität bei der Identifikation des Suchaspektes (siehe (135) bis (137), S. 160). So kann die Anfrage

(168) „Ich suche 'King Lear' von XY“

in einem System, das nicht nur nach Büchern, sondern auch nach Filmen, Musik und damit eben auch Verfilmungen, Theateraufnahmen, Fernsehaufzeichnungen, Vertonungen, Filmmusiken etc. sucht, nur dann eindeutig beantwortet werden, wenn in (168) statt „XY“ z.B. „Shakespeare“ genannt würde: Wenn dem System „Shakespeare“ als Autor von „King Lear“ bekannt ist, können die verschiedenen Druckfassungen präsentiert und die Suche in weiteren Materialarten angeboten werden. Wird hingegen statt „XY“ in (168) ein dem System unbekannter Ausdruck, der vielleicht nicht einmal als Eigenname zu erkennen ist, verwendet, dann wäre eine Rückfrage des Systems wie in (169) der kommentarlosen Auflistung aller mit „King Lear“ betitelten Objekte vorzuziehen:

(169) „Ich kann das gesuchte Objekt leider nicht eindeutig identifizieren. Suchen Sie im Bereich *Drama, Aufnahmen von Theateraufführungen, Verfilmungen, Vertonungen* oder *Sekundärliteratur* ? Oder möchten Sie *alle* Ergebnisse sehen ?“

Im Fall, daß das System für eine Suchanfrage keine oder nur eine dem Benutzer unzureichend erscheinende Antwort findet, sollte das System versuchen, durch Rücksprache mit dem Benutzer eine modifizierte Anfrage zu formulieren. Eine Möglichkeit dazu bietet eine Ergänzungskomponente zum Lexikon, das sogenannte *Synonymlexikon*. Hieraus könnte das System dem Benutzer Vorschläge zu einer modifizierten Anfrage generieren, die es noch nicht automatisch selber durchgeführt hat. Eine solche Rückfrage ist sinnvoll, da die Entscheidung, welcher Ausdruck tatsächlich synonym zum ursprünglich verwendeten Ausdruck ist, nicht immer automatisch getroffen werden kann. Auch das Angebot an den Benutzer, eine *Online-Enzyklopädie* zu konsultieren, um im Anschluß eine veränderte Anfrage formulieren zu können, ist eine sinnvolle Rückmeldung zur Verbesserung der Rechercheleistung. Die dabei vom Benutzer explizit hergestellten Bezüge zwischen alter Formulierung und neuer Formulierung sollten dabei vom System gelernt werden (z.B. in Form eines benutzerspezifischen Synonymlexikons).

Eine Rückfrage oder Rückmeldung muß aber nicht immer eine Frage sein, sondern kann (in Anlehnung an die in 4.2.5 angesprochene hilfreiche Überbeantwortung von Fragen gemäß Grice'scher Maximen) auch wertvolle Zusatzinformation zur Verfügung stellen: Für eine eindeutig zu beantwortende Suchanfrage kann zusätzlich auf einen Experten vor Ort verwiesen werden, eine Fortsetzung der Suche in anderen Datentypen (Musik zum Buch, Verfilmung zum Werk) kann als lohnenswert angeboten werden usw.

Ebenfalls durch Rückmeldungen zu verbessern ist die in 3.2.7 bereits angesprochene Eingabefehlerkorrektur. Einerseits wird die Situation des Benutzers verbessert, wenn er vor einer vom System durchgeführten Korrektur explizit gefragt wird (siehe auch 6.2). Andererseits werden durch Rückfragen auch „riskantere“ Ersetzungen möglich (siehe 4.2.9), die explizit vom Benutzer nicht nur bestätigt, sondern auch wirklich inhaltlich kontrolliert werden müssen.

Solche Strategien zur Prüfung der Eingabe und zur Dialogführung mit dem Ziel, gescheiterte oder unzulängliche Recherchen interaktiv zu verbessern, tragen wesentlich zur frustrationsarmen Benutzung und zum „intelligenten“ Erscheinungsbild der Schnittstelle bei.

Die Möglichkeit von Rückfragen oder Rückmeldungen des Systems an den Suchenden hat aber wiederum Folgen für die Komplexität der Interaktion: Ein Diskurs über die vom System gefundenen Objekte führt, wenn es sich bspw. um graphische Elemente am Bildschirm oder um gefundene Bildobjekte handelt, sofort zur Verwendung deiktischer Ausdrücke, wie in 4.2.8 ausgeführt.

5.1.5 Fazit

Die Untersuchung der Erweiterung der Eingabemöglichkeiten eines Systems wie OSIRIS ist auf der Basis der bisher eingesetzten Techniken sinnvoll: Dies wurde zu Beginn von 5.1 festgestellt. Eine solche Erweiterung mit dem Ziel, eine einheitliche Schnitt-

stelle zu allen vorhandenen Suchaspekten zu erhalten, könnte durch eine Reduktion der Vorgabe *Ich suche Literatur zum Thema ...* auf der Eingabemaske hin zu *Ich suche ...* erreicht werden (5.1.1). Die durch Verkürzung der Vorgabe erschwerte Identifikation des intendierten Suchaspektes kann durch Metaregeln geleistet werden.

Grundsätzlich ist es dabei wünschenswert, wenn die Möglichkeiten der Eingabeanalyse die Möglichkeiten der Retrievalkomponente übersteigen. Damit kann das System auf solche Fragen angemessen reagieren, die anhand der Datenlage nicht zu beantworten sind. Schwierig ist, es zu bestimmen, bis zu welchem Grad die Eingabeanalyse die Retrievalfähigkeiten des Systems überschreiten soll und wie weit dies angesichts der starken Verwendung von Weltwissen zu leisten ist.

Je umfassender die Eingabemöglichkeiten in einem System ausfallen und je heterogener der zugrundeliegende Datenbestand ist, umso größer sind die Möglichkeiten, daß Suchende Eingaben tätigen, die für das System nur schwer oder gar nicht zu interpretieren sind. Diese durch die Erweiterung der Eingabemöglichkeiten entstehenden Schwierigkeiten können genauso wie bereits bestehende Probleme (z.B. die Eingabefehlerkorrektur) durch Rückfragen und Rückmeldungen des Systems an den Benutzer gelöst werden. Die dadurch entstehende Interaktion mit dem Suchenden erhöht die Nutzerzufriedenheit, da dieses Vorgehen zur Verbesserung der Rechercheergebnisse, der einfacheren Bedienbarkeit und zum „intelligenten“ Erscheinungsbild der Schnittstelle beiträgt.

5.2 Weitere Verbesserungen von OSIRIS

Wie in 5.1 angesprochen, ist die Erweiterung der natürlichsprachlichen Eingabemöglichkeiten in OSIRIS im Zusammenhang mit anderen Verbesserungen wie z.B. einer Erweiterung des zugrundeliegenden Datenbestandes, einer Verbesserung der Suchverfahren sowie anderer, nicht-linguistischer Verbesserungen der Eingabeschnittstelle zu sehen. Ziel dieser Verbesserungen ist dabei nicht der punktuelle Ausbau des Systems, sondern die Steigerung der Nutzerzufriedenheit insgesamt durch aufeinander abgestimmte Veränderungen (zur Evaluation von Dialogsystemen siehe 6.1 und auch (Saracevic, 1995)).

5.2.1 Integration weiterer (heterogener) Datenbestände

Die Integration neuer Datenbestände in ein System wie OSIRIS ist wichtig und notwendig, da OSIRIS (als Bibliothekssystem) sich den neuen Anforderungen an die Rolle der wissenschaftlichen Bibliothek anpassen muß. Durch die zunehmende Vernetzung (nicht nur) der Wissenschaftler, die (auch aus Kostengründen) wachsende Bedeutung der Online-Publikationen und die Bemühungen der Bibliotheken, den Rang als zentrale Vermittlungsstelle zu *allen* relevanten Informationsarten zu wahren (und für andere Berufsgruppen auszubauen), wird das Angebot an den Suchenden, nicht nur im Bibliothekskatalog, sondern auch z.B. in klassifizierten Datenbanken oder in erschlossenen Teilen des WWW zu suchen, notwendig (vgl. Birmingham et al. (1994); Griffiths und Kertis (1994); Mark Pejtersen (1998)). Hinzu kommt, daß viele Beiträge zu Konferenzen, aber auch Projektberichte und Zeitschriftenartikel bereits als Vorabdruck (preprint) ins Netz gestellt werden – zudem stellen zahlreiche Zeitschriften (auch) aus Kostengründen auf eine online Erscheinungsweise um, und Dissertationen werden verstärkt im Internet veröffentlicht.

Der Trend zu immer stärkerer Vernetzung auch entlegendster Rechner hat zur Folge, daß zunehmend nach im Netz verteilten, heterogenen Daten wie Texten, Bildern, Audiodaten oder Einträgen in Datenbanken gesucht wird. Drei Eigenschaften sind dabei für Systeme wichtig, die zukünftig erfolgreich in verteilten, heterogenen Daten suchen sollen: *Skalierbarkeit*, *Erweiterbarkeit* und *Adaptierbarkeit*.

Skalierbarkeit beschreibt die Fähigkeit des Systems, auch mit wachsenden Datenmengen noch problemlos umgehen zu können – die Eigenschaft, skalierbar zu sein, kann ein System dabei auf ganz unterschiedliche Arten erreichen, in manchen Einsatzgebieten reicht es, die Hardwareplattform auszubauen, in anderen wird es nötig sein, die Aufgaben aufzuteilen, indem z.B. eine Metasuchmaschine die Ergebnisse verschiedener spezialisierter Suchmaschinen integriert.

Erweiterbarkeit ist wichtig, weil ständig neue Datentypen, Quellen, Protokolle und Oberflächen hinzukommen, die für das System behandelbar sein sollen. Dabei wäre in einer idealen Lösung das System selbst in der Lage, solche Änderungen oder Neuerungen zu erkennen und sich darauf einzustellen. Was in Hinblick auf den seit langem andauernden Konflikt zwischen Metasuchmaschinen und den von ihnen be-

nutzten Suchmaschinen ein Politikum ist, könnte technisch über ein eigenes Datenmodell und eine eigene Abfragesprache für einen Mediator realisiert werden, um die Datenmodelle und Abfragesprachen der anderen Systeme zu subsumieren.

*Adaptierbarkeit*⁶¹ ist wichtig, um das System auf einzelne Teilbereiche hin zu optimieren. Dabei steht die Adaptierbarkeit in gewisser Hinsicht im Widerspruch zur Erweiterbarkeit. Während es bei der Erweiterbarkeit darum ging, potentiell möglichst große Bereiche der zu bearbeitenden Domäne behandeln zu können, steht hinter der Adaptierbarkeit die Erkenntnis, daß es unrealistisch ist, mit einem generischen System quer zu allen Inhalten sämtliche Suchanfragen gleich gut behandeln zu wollen. Daher ist eine erweiterbare Architektur gefragt, in die spezialisierte Module integriert werden können, die durch besondere Behandlung spezieller Domänen das Ergebnis des Gesamtsystems verbessern sollen (s. a. 5.2.2 zu User Interface Agents). Diese Lösung, die auf den ersten Blick einleuchtet, hat aber den entscheidenden Nachteil, daß nun die systeminterne Kommunikation geregelt werden muß: Es wird eine Komponente nötig sein, die die Verwaltung des Zusammenspiels der verschiedenen Module und die Entscheidung über die Grenzen der einzelnen Domänen übernimmt. Diese Komponente wird damit auch über den Aufgabenbereich der Spezialmodule entscheiden müssen und die Suchergebnisse in Abhängigkeit zum Modul, aus dem sie stammen, bewerten und integrieren müssen.

Im Folgenden werden nun einige mögliche Erweiterungen des Datenbestandes und deren spezifische Eigenschaften vorgestellt:

Personen Die Namen der an der Universität beschäftigten Wissenschaftler stellen eine vielversprechende Recherchedomäne dar. Diese Personennamen, zusammen mit Angaben über die jeweiligen Interessens- und Fachgebiete, liegen implizit z.B. in Forschungsführern oder universitären Veranstaltungsverzeichnissen bereits vor. Darüber hinaus ist die Integration von Personennamen aus den Datenbanken der Kammern und Verbände denkbar, so daß zum Ergebnis einer thematischen Recherche in OSIRIS das Angebot einer weiterführenden Suche nach menschlichen Experten, die in der Region ansässig sind, möglich wird. An einer Schnittstelle wie in 5.1 beschrieben wäre dann auch eine Anfrage wie „Ich suche einen Experten für ...“ möglich, die auf den gespeicherten Personenbezeichnungen, bestehend aus Eigennamen („Peter Meyer“), definiten Beschreibungen („Vorsitzender der Handwerkskammer Bremen“) und klassifikatorischen Merkmalen („Schichtleimtechnik“), prozessiert würde. Ergebnis einer solchen Anfrage wären einerseits Name und Kontaktadresse einer solchen Person, andererseits aber auch Institutionen, an denen Personen mit den in Frage stehenden Eigenschaften vermutet werden können. So können z.B. die Mitarbeiter eines universitären Instituts aufgrund des Forschungsprofils interessant sein, so daß das System die Institutsadresse als Antwort gibt. Die Informationen zu den Personen können viel-

⁶¹Hier nicht im Sinne einer Adaptivität des Systems an den Benutzer gemeint, sondern ausschließlich als Parameter neben Skalierbarkeit und Erweiterbarkeit in Hinblick auf den Datenbestand. Zu Adaptivität als Anpassung der Schnittstelle an den Benutzer, s. S. 202.

fach aus bereits bestehenden, zu integrierenden Datenbeständen gewonnen werden.⁶² In anderen Fällen können Profile für z.B. die beteiligten Wissenschaftler mittels der über sie bereits vorhandenen Informationen gewonnen werden. Zu den bereits vorhandenen Informationen gehören die klassifikatorische Schwerpunktbildung ihrer lokal meist komplett verfügbaren Publikationen, die Zugehörigkeit zu bestimmten universitären Institutionen, die Denomination ihrer Stelle und die thematische Ausrichtung ihrer Lehrtätigkeit laut Veranstaltungsverzeichnis.

Herfurth et al. (1993) stellen einen ambitionierten Ansatz zur Auswertung des in bibliographischen Datenbanken implizit enthaltenen Kontextwissens vor. Ziel ist es, für Suchanfragen von Benutzern Aussagen wie „Ein Repräsentant auf dem Gebiet der Systemtheorie ist Luhmann“ auf der Grundlage der Daten in einer bibliographischen Datenbank zu generieren. Dabei ist es nicht das Ziel des Systems, zur Suchanfrage passende Dokumente zu finden, sondern es sollen ausschließlich im erfragten Fachgebiet einschlägige Personen gefunden werden. Dazu wird ein vom Suchenden natürlich-sprachlich mit Hilfe einer Nominalphrase formuliertes Suchthema („Politische Einstellungen von bildungsbürgerlichen Parlamentsabgeordneten 1830-1933“) auf nicht näher beschriebene Art und Weise in Konzepte zerlegt („Bildungsbürger“, „politische Einstellung“, „Parlament, Abgeordnet“ etc.), zu denen dann in der Datenbasis einschlägige Dokumente gesucht werden. Diese Konzepte werden durch Befragung lexikalischer Quellen expandiert, und die Expansion wird dem Suchenden zur Bestätigung vorgelegt.

Obwohl Herfurth et al. (1993) explizit darauf hinweisen, daß sie „nur“ für eine Suchanfrage relevante Personen, nicht jedoch relevante Dokumente finden wollen, ist ein notwendiger Zwischenschritt bei der Bestimmung der Experten aber gerade die Suche nach Dokumenten zum Thema, deren Autoren dann als mögliche Experten in Erwägung gezogen werden. Wird eine Anfrage in drei Konzepte zerlegt, dann wird für jedes (expandierte) Konzept die Menge der relevanten Dokumente gesucht und die Autoren dieser Dokumente als Experten für das Konzept betrachtet. Anschließend wird der Schnitt über den Dokumentenmengen gebildet, in der Hoffnung, ein oder mehrere Dokumente zu finden, die für alle drei Konzepte relevant sind. Der Autor eines solchen Dokumentes wird dann aufgrund seiner Veröffentlichungen als Experte für das Thema der Benutzeranfrage betrachtet.

Existiert kein solches Dokument und damit kein Experte, der durch mindestens eine Veröffentlichung sein Expertentum für die gesamte Anfrage unter Beweis gestellt hat, wird versucht, eine suboptimale Lösung durch Zerlegung zu finden: Die beste Zerlegung wäre ein Autor, der zwei Dokumente veröffentlicht hat, nämlich eines, das für die Konzepte eins und zwei relevant ist und ein anderes Dokument, das für die Konzepte eins und drei relevant ist. Scheitern alle Versuche, für die Vereinigung der Konzepte einen Experten zu finden, wird das iterative Problemlösungsverfahren auf der Ebene der Konzeptmengen fortgesetzt: D.h. es wird versucht, einen Experten

⁶²Die Zahl der isolierten, parallel existierenden Fachdatenbanken an ein und derselben Institution ist meist erschreckend hoch – siehe auch 2.1.

für eine möglichst große Teilmenge der drei Konzepte auf die beschriebene Weise zu finden.

Wenn ein Aufsatz für ein (expandiertes) Konzept als einschlägig befunden und damit der Autor des Aufsatzes als Experte für das Konzept erkannt wurde und wenn dieser Aufsatz in einem Sammelband erscheint, der insgesamt wiederum einschlägig ist, und dessen Herausgeber ebenfalls als Experten erkannt wurden, dann trägt der Expertenstatus des Sammelbandbeitragsautoren zur Verstärkung des Expertenstatus der Herausgeber bei.

Herfurth et al. (1993) stellen mit ihrem Ansatz keine Alternative zu einem System wie OSIRIS vor, sondern eine mögliche Ergänzung oder Erweiterung. Denn ohne daß die Autoren dies selbst ansprechen würden, muß doch festgestellt werden, daß der vorgestellte Ansatz zur Identifikation von Experten zu einem (Such)Thema auf die vollständige und qualitativ hochwertige Erschließung der Dokumente angewiesen ist. Denn nur über die inhaltliche Relevanz eines Dokumentes kann (vermittelt über die Autorenrolle) für eine Person auf deren Expertentum geschlossen werden. Wie schwierig aber die Suche in Bibliotheksdaten nach solchen Dokumenten ist, die für eine natürlichsprachliche Konzept-Formulierung relevant sein sollen, wurde in 2.2 und 3.1 erörtert.

Für Forschungszwecke (und nicht unbedingt als Informationssystem) erscheint die Anreicherung eines Bestandes an (historischen) Personennamen um Eigenschaften wie Lebensdaten, Wohnort, Beruf, Familienstand, Kinderzahl usw. zu einem Bestand an Individuenkonzepten⁶³ sinnvoll. Aus Anwendersicht könnte ein solcher Datenbestand, der mit einer Inferenzkomponente gekoppelt wäre, für Eigennamen, bei denen die Existenz der dahinter vermuteten Person ungesichert ist („Shakespeare“, „Karl der Große“), Aussagen über die Konsistenz des Individuenkonzeptes geben. Für zwei oder mehr in der Datenbank verfügbare Individuenkonzepte könnten zudem Hypothesen über deren extensionale Identität angestellt werden („Philipp Melanchton“ und „Philipp Schwarzert“, „Romain Gary“ und „Emile Ajar“). Auch könnte auf der Grundlage von Individuenkonzepten, deren Bestandteile als gesichert angenommen werden, eine erste Abschätzung zur Verlässlichkeit neuer Quellen geschehen, wenn nämlich die Konsistenz des Individuenkonzeptes im Anschluß an dessen Erweiterung aus der neuen Quelle überprüft wird.

Auch aus theoretischen Erwägungen heraus ist eine Umsetzung des oben genannten Ansatzes vielversprechend, da der Versuch, Personen als Individuenkonzepte zu repräsentieren, d.h. über die Menge der sie beschreibenden Eigenschaften zu identifizieren, einen Beitrag zur aktuellen Debatte um den Status von Eigennamen darstellt. Mit Bezug auf Kripke (1972) werden Eigennamen derzeit von der Mehrheit der Forscher als *rigid designators* betrachtet, d.h. Eigennamen referieren in allen möglichen Welten stets auf dasselbe Individuum. Im Gegensatz dazu kann die Referenz von Individuenkonzepten (die letztlich definite Beschreibungen sind) zwischen möglichen Welten variieren. Die Annahme, daß Eigennamen *rigid designators* seien, wurde in

⁶³Individuenkonzepte sind Sammlungen von beschreibenden Merkmalen einer Person.

den vergangenen Jahren insbesondere im Bereich der Präsuppositionsforschung angegriffen (Van der Sandt und Geurts (1991), Van der Sandt (1992), vgl. auch 4.2.4). Geurts (1997b) faßt diese Entwicklungen zusammen und zeigt für den Ansatz von Kneale (1962), daß die Identifikation eines Eigennamens („Peter“) mit einer definiten Beschreibung („das Individuum mit Namen 'Peter'“) der Analyse als rigid designator empirisch überlegen ist.

Externe Datenbanken Nicht immer ist es möglich, andere Datenbestände in das vorhandene System zu integrieren, z.B. weil der Zugriff auf sie kostenpflichtig ist oder der Betreiber aus anderen Gründen kein Interesse hat, eine Integration zu unterstützen. In anderen Situation, in denen eine Integration möglich wäre, kann es dennoch sinnvoll sein, die in Frage stehenden Datenbestände eigenständig zu belassen, z.B. um zu verhindern, daß das entstehende System unstrukturiert und profillos würde.

Dennoch möchte man den Suchenden am eigenen System die Möglichkeit geben, auf eigene Rechnung (wenn es sich um einen kostenpflichtigen Zugriff handelt) oder aus eigenem Interesse (wenn es sich um einen Datenbestand zweifelhafter Qualität handelt) eine externe Datenbank aufzusuchen. Auch in den Fällen, in denen aus technischen Gründen eine Integration nicht in Frage kommt, möchte man dennoch einen Hinweis auf die Existenz dieses Datenbestandes geben. Dies wird möglich, wenn Datenbanken mit ihrem Profil als Objekte im System gespeichert werden und so zu einer thematischen Recherche nicht nur entsprechende Bücher, Filme usw. präsentiert werden, sondern zusätzlich auch ein Verweis auf eine Datenbank, die laut ihres Profiles für die Benutzeranfrage relevante Informationen enthalten könnten.

Dieses Konzept wurde in BREWIS⁶⁴ umgesetzt, wo neben der einfachen Suche in den Beständen der Bibliothek auch die Recherche nach für die Interessen des Benutzers relevanten Faktendatenbanken möglich ist. Die Klassifikation der Datenbanken geschieht dabei von Hand mit Hilfe des in der Bibliothek bereits verwendeten Klassifikationssystems, so daß Bücher und Zeitschriften zusammen mit Verweisen auf Datenbanken (nicht die Inhalte der Datenbanken !) gefunden werden. Durch die Vermittlung des Suchenden zur Datenbank selbst werden die einschlägigen rechtlichen Probleme bei der Online-Nutzung geistigen Eigentums (vgl. McKnight (1998)) vermieden – im besten Fall kann dabei die ursprüngliche Suchanfrage gleich an einer Online-Schnittstelle der Datenbank prozessiert und die Ergebnisse, wie sie die Datenbank selbst (mit allen Copyright-Merkmalen) präsentiert, in einem *frame set* eingebettet werden.

Tong und Holtzman (1994) berichten von einem Information Retrieval System, das in verschiedensten Datenbeständen bis hin zu externen Datenbanken recherchieren kann. Dazu wird eine Benutzeranfrage durch einen *query mediator* in eine Reihe einzelner, spezifischerer Anfragen übersetzt, die dann durch je einen *source mediator* in Anfragen für einzelne Datenquellen übersetzt werden. Dabei kann eine Datenquelle auch eine Faktendatenbank sein, deren Datenmodell (d.h. die Semantik der Felder) und Anfra-

⁶⁴Bremer Wirtschafts-Informationssystem, <http://brewis.suub.uni-bremen.de>

gespräche dem zugehörigen *source mediator* bekannt sind. Die Ergebnisse der Suche in den einzelnen Beständen werden durch den jeweiligen *source mediator* in eine Zwischenrepräsentation übersetzt, aus der dann der *query mediator* eine Gesamtantwort für den Suchenden generiert. Für die Anfrage werden dem Benutzer einige Beschränkungen auferlegt. So ist keine freie Eingabe möglich, sondern dem Suchenden wird nur die Möglichkeit eingeräumt, aus vorgegebenen, domänenabhängigen Konzepten diejenigen zu wählen, die sein Suchinteresse repräsentieren. Bei diesen Konzepten handelt es sich nicht um eine echte Ontologie, da nach Aussagen der Autoren keine vollständige Beschreibung der Domäne geleistet wird, sondern die Konzepte lediglich die von den Systementwicklern antizipierten Informationsbedürfnisse der Suchenden darstellen.

Barja et al. (1998) stellen einen Ansatz vor, der dem Endnutzer den Umgang mit der wachsenden Zahl unterschiedlicher Datentypen und Datenquellen erleichtern soll. Ähnlich wie Tong und Holtzman (1994) schlagen Barja et al. (1998) vor, die unterschiedlichen Datentypen unter einem Mediator zu integrieren und dem Suchenden unter einer einheitlichen Oberfläche mit einer einheitlichen Syntax die Suche zu ermöglichen. Dem Mediator kommt dabei die Aufgabe zu, für eine Benutzeranfrage Quellen mit möglicherweise relevanter Information auszuwählen, die Benutzeranfrage für die identifizierten Quellen in geeignete Anfragen zu übersetzen und die Suchergebnisse zu ordnen, zu integrieren und geeignet zu präsentieren. Um dies leisten zu können, sind sehr viele Metainformationen zu den Quellen nötig, und es stellt sich die Frage, wie diese Informationen erlangt werden können.

Zwei Beispiele für Ansätze, die auf einem Mediator basieren, sind für Barja et al. (1998) das deutsche Verbundprojekt MeDoc (Endres und Fuhr (1998) und Barth et al. (1998)) und die aus dem WWW bekannten Metasuchmaschinen (z.B. MetaGer, <http://meta.rzn.uni-hannover.de>). MeDoc ermöglicht die Suche auf wenigen, speziell für MeDoc aufbereiteten Volltexten sowie in (hochstrukturierten) bibliothekarischen Daten, so daß der MeDoc-Mediator hochspezialisiert und daher nur sehr schwer zu erweitern ist. Die bekannten Metasuchmaschinen hingegen bearbeiten zwar eine zahlenmäßig sehr viel breitere Datenbasis, aber kennen nur einen Datentyp, nämlich die HTML-Datei und für diese lediglich die Volltextretrievalanfrage.

Barja et al. (1998) referieren dazu eine Reihe Möglichkeiten, die in der Vergangenheit genannt wurden, um an die benötigten Metainformationen zu kommen. Eine Möglichkeit, die auch von Tong und Holtzman (1994) angesprochen wurde, sieht für einen intelligenteren Mediator vor, mit Hilfe des Datenmodells der zu befragenden Datenbanken einen parallelen Zugriff auf die Quellen zu ermöglichen. Wie allerdings ohne menschliche Expertise die Semantik der Datenbankfelder erkannt und für den Mediator nutzbar gemacht werden könnte, bleibt offen. Eine andere Möglichkeit ist die explizite Beschränkung der Domäne, in der das System eingesetzt wird, so daß der Mediator mit dem nötigen inhaltlichen Wissen über die Domäne versehen werden kann, um eigenständig relevante Quellen identifizieren, Suchanfragen geeignet übersetzen und Suchergebnisse präsentieren zu können. Inwieweit das dafür benötigte Wissen klar

begrenzt ist und wie es geeignet (d.h. z.B. einfach erweiterbar und wiederverwendbar) repräsentiert werden kann, bleibt ebenfalls offen.

Ein anderer, ambitionierter Ansatz sieht den Einsatz von semantischen Agenten vor, die in der Lage sind, bestimmte Arten von Quellen hinsichtlich ihrer inhaltlichen Relevanz für eine Suchanfrage einzuschätzen und auch zu befragen. Mehrere, unterschiedlich spezialisierte Agenten sollen dann mit Hilfe von Wissensrepräsentationssprachen ihre Kenntnisse über die inhaltliche Relevanz von Quellen und die Struktur der Anfrage austauschen können. Allerdings steht die Entwicklung solcher Agenten und solcher Anfragesprachen noch am Anfang (vgl. aber Rousseau et al. (1996)). Barja et al. (1998) selbst schlagen mit „Informia“⁶⁵ ein System vor, das beginnend mit handannotierten Metainformationen in die Lage versetzt werden soll, später selbst die für weitere Quellen benötigten Metainformationen zu erschließen. Diese Metainformationen umfassen Angaben über die Identität der Quelle (z.B. Name, Typ, Ort), bestimmte Service-Parameter (z.B. Kosten, Antwortzeiten, Verfügbarkeit), Inhaltsbeschreibung (zu welchen Themen liegen Informationen vor ?), Retrievalfunktionalität (z.B. natürliche Sprache, Boolesche Operatoren, Trunkierungsmöglichkeiten, Stopwortlisten) und Informationen zum Repräsentationsschema der Quelle. Mit Hilfe der Schema-Informationen ist „Informia“ erst in der Lage, die Ergebnisse unterschiedlich strukturierter Quellen miteinander zu vergleichen, da hier die Semantik der in den Quellen verwendeten Attribut-Wert-Paare sowie eventuelle hierarchische Abhängigkeiten erläutert werden. Nur mit diesen Informationen ist es aber möglich, Retrievaltechniken typgerecht einzusetzen: Nur wenn das System weiß, daß es mit einem Datumsfeld zu tun hat, kann es für drei Werte entscheiden, ob der Abstand von „Jan“ zu „Mar“ geringer ist als der zu „Sep“.

Die automatische Erschließung dieser Metainformationen wird für „Informia“ zwar für die Zukunft angekündigt, konkret umgesetzt wurde aber nur ein auf Retrieval basierender Zwischenschritt, der auf den manuell erstellten Schemata basiert. Schemata in „Informia“ sind strukturierte Texte mit Einträgen wie etwa `Buchtitel: string` oder `Autor: ListOfNames`. Diese Schemata werden in „Informia“ nun selbst als Datenquelle interpretiert, d.h. sie werden expandiert und dann indexiert. Die Expansion umfaßt dabei die morphologische Reduktion, die Dekomposition, die Anreicherung durch Übersetzung in WordNet Synsets und anderes. Das Schema einer neuen Quelle wird mit diesem Vorgehen durch Abbildung seiner Bestandteile auf bekannte Bestandteile im Index gewonnen: Enthält das Schema der neuen Quelle ein Feld `S_Adresse`, dann wird durch Links-Trunkierung eine Abbildung auf das im Index verzeichnete `Adresse` vorgenommen und so Stück für Stück versucht, das neue Schema durch Interpretation der einzelnen Felder zu gewinnen. Auch wenn dieses Vorgehen geschickt ist und in vielen Fällen mit relativ einfachen, aus dem Information Retrieval bekannten Methoden zum Erfolg führt, ist es doch bis zur automatischen Erschließung einer beliebigen neuen Quelle noch ein enormer Schritt. Um dies zu erreichen, ist es entweder nötig, so viele Quellen im Zugriff zu haben, daß es auf eine weitere, für die die

⁶⁵<http://informia.ubilab.ch>

automatische Erschließung scheitert, nicht ankommt, oder aber es müssen konsequent andere Methoden gefunden werden, die nicht nur über den Strukturvergleich zwischen Feldnamen, sondern über die Semantik der Feldbezeichner und den Wertebereich der Inhalte Annahmen zur Struktur der Quelle treffen. Weder verfügt „Informia“ aber über eine ausreichend große Zahl an Quellen, noch ist von solchen neuartigen Methoden etwas zu sehen.

Ein in der Zielrichtung ähnlicher, aber sehr viel weniger elaboriert ausgearbeiteter Entwurf wird von Chakravarthy und Haase (1995) unter dem Namen „NetSerf“ vorgestellt. „NetSerf“ hat zum Ziel, zu natürlichsprachlichen Anfragen, die ein Informationsbedürfnis ausdrücken, passende Quellen im Internet wie Online-Datenbanken oder thematisch einschlägige WWW-Seiten zu liefern. Dazu werden einschlägige Quellen zunächst im Netz identifiziert und anschließend manuell natürlichsprachlich beschrieben. Aus diesen Beschreibungen werden (wiederum manuell) die relevanten Terme gezogen und mit Hilfe einfacher ontologischer Relationen („Schweiz“ ist ein „Land“ das Teil von „Europa“ ist) sowie den entsprechenden WordNet Synsets angereichert, so daß eine Inhaltsbeschreibung der Quelle entsteht. Eine natürlichsprachliche Suchanfrage wird nun ebenfalls mit WordNet behandelt und für die so erweiterte Anfrage versucht, eine passende Beschreibung einer Quelle zu finden.

Archive und Altbestände in Bibliotheken Von Laien oft miteinander verwechselt, haben Archive andere, wenngleich zum Teil ähnliche Aufgaben und Probleme wie (wissenschaftliche) Bibliotheken. Horsman (1993) plädiert für den Einsatz von Methoden der NL-Verarbeitung in Archiven und beschreibt dabei auch eine Problemstellung, wie sie in wissenschaftlichen Bibliotheken aus der Erschließung von Altbeständen bekannt ist. Um die Kosten für die rückwärtige Erschließung möglichst gering zu halten, werden Karteikarten mit Erschließungsmerkmalen (z.B. aus dem alten Zettelkatalog) gescannt und mit Hilfe von OCR-Methoden maschinenlesbar gemacht und die erkannten Daten dem Online-Katalog hinzugefügt. Für die Situation in der wissenschaftlichen Bibliothek wurde dazu in 3.4.4 bereits das MILOS-Nachfolgeprojekt KASCADE beschrieben. Dort wurden mit Junger (1999) auch die enormen technischen Probleme einer qualitativ ausreichenden Automatisierung dieses Verfahrens angesprochen (s. a. 5.2.3). Für das sehr viel weniger anspruchsvolle Szenario der Publikation im Internet beschreiben Worring und Smeulders (1998) das Scannen von Papierdokumenten, um neben der Inhaltserschließung auch die Struktur des Dokumentes verfügbar zu machen. Da dabei aber nicht nur eine maschinenlesbare Beschreibung, sondern auch der gesamte Volltext als Scan für den Betrachter verfügbar wird, ist die Situation des Suchenden eine ganz andere als an einem Online-Bibliothekskatalog.

Die Synthese dieser Ansätze wird von Bayer et al. (1994) mit dem System OMNIS vorgestellt: OMNIS erlaubt nicht nur die Dokumentarchivierung mittels OCR, sondern integriert die Ergebnisse dieses Arbeitsschrittes in ein bestehendes OPAC-System, so daß der Suchende bereits am Computer die im Katalog gefundenen Werke begutachten kann. Allerdings bietet OMNIS über die geschilderte, nach Angaben der Autoren pra-

xistaugliche⁶⁶ Bestandserweiterung hinaus keine Verbesserung der Suchfunktionalität oder der Eingabeverarbeitung. Insbesondere wird in OMNIS an keiner Stelle natürliche Sprache verarbeitet.

Andere Objekte Neben den genannten Objekten sind weitere in der Wissensbasis einer zukünftigen OSIRIS-Version denkbar: Bilder zum Beispiel sind lohnenswerte Objekte für das Retrieval, da sie in vielen Anwendungsfeldern ebenso wichtig sind wie textuelle Objekte (Herzog et al., 1998). Das Forschungsgebiet „Bildererkennung“ steht noch am Anfang und wenn es praxisreife Lösungen hervorgebracht hat, dann für sehr spezielle, technische Anwendungsfälle wie z.B. die automatische Erkennung von Produktionsfehlern bei Garnen (Hermes, 1999; Hermes et al., 1999).

Für das Image Retrieval im allgemeinen Fall sind diese Lösungen nicht zu verwenden (siehe aber Djeraba und Bouet (1997) für einen Ansatz zum Image Retrieval). Vielversprechend ist hingegen die Kombination aus Bildern und sprachlicher Annotation, wie sie im Produktionsbetrieb bei Fernsehsendern oder in Filmstudios durchgeführt wird (Hermes und Dammeyer, 1999). Dabei werden die Sendungsmitschnitte systematisch von Menschen unter Angabe des exakten Zeitstempels kommentiert, um ein späteres Wiederauffinden von Aufnahmen zu einem bestimmten Thema zu ermöglichen. Diese Arbeit wird, im Gegensatz zu den hochspezialisierten und daher teuren Fachreferenten der Bibliotheken, in den Sendern von nicht extra zu diesem Zweck ausgebildeten Mitarbeitern erledigt. Hier wären computerlinguistische Verfahren sinnvoll, um mittels NL-Analyse (lexikalische und morphologische Reduktion, Synonymerkennung etc.) einerseits zu rigide Vorgaben durch EDV-gestützte Autorensysteme zu verhindern und andererseits das eigentliche (textgestützte) Retrieval für den Suchenden zu vereinfachen und die Suchergebnisse zu verbessern. Da hier im wesentlichen dieselben Eigenschaften der natürlichen Sprache eine Rolle spielen, wie es für die Suche auf bibliographischen Daten der Fall ist, können die in 3.1 genannten Vorteile des OSIRIS Systems (was die natürlichsprachliche Analyse angeht) auf diesen Anwendungsfall übertragen werden. Dabei würde insbesondere ein Ausbau der Semantik räumlicher Präpositionen zu einer Verbesserung des Bild-Retrievals führen, da so natürlichsprachliche Beschreibungen komplexer Situationen („Ein Tisch mit einer Vase, darin Rosen, links davon eine Kerze, über dem Tisch ein Kronleuchter.“) besser analysiert und auf Suchanfragen („Ein Bild mit Blumen *auf* einem Tisch, *neben* einer Kerze, das ganze *unter* einem Kronleuchter.“) abgebildet werden könnten.

Prinzipiell ist diese Methode der natürlichsprachlichen Annotation auch auf andere Objekte wie z.B. Musikstücke übertragbar, die so recherchierbar würden. Einen interessanten Ansatz stellen McNab et al. (1996) vor: Sie versuchen, Musiktitel über deren Inhalt und nicht nur über die textuelle Annotation recherchierbar zu machen. Dazu wird ein Ähnlichkeitsvergleich vorgenommen zwischen den automatisch aus der Musikdatei gewonnenen Kenndaten zu Lautstärke, Tonverlauf sowie dem vermutlichen Rhythmus und einer Eingabe des Benutzers an einem Mikrophon, in das die-

⁶⁶Im Gegensatz zu KASCADE, siehe Junger (1999).

ser eine Melodie singen oder summen kann (zum Gesamtsystem siehe auch Witten et al. (1998)). Dieser Ansatz könnte in Zukunft auch auf Bilder übertragbar sein: In Kombination mit Kantenerkennungsalgorithmen und der Erkennung von Farbflächen könnte eine Schnittstelle, die vom Suchenden erwartet, daß er neben einer textuellen Beschreibung des Bildes auch eine Skizze auf einem stiftsensitiven Display anfertigt, eine ähnlichkeitsbasierte Suche durchführen.

5.2.2 Verbesserungen im Bereich des Interfaces

Der erfolgreiche Umgang mit einem Computerprogramm und damit ein großer Teil der Nutzerzufriedenheit hängt von der ansprechenden und verständlichen Gestaltung der Benutzerschnittstelle ab. Oft kann eine Verbesserung der Präsentation von Informationen einen stärkeren Effekt für die Verständlichkeit und Nutzerzufriedenheit haben als Verbesserungen der zugrundeliegenden Verarbeitungskomponenten (Dagan und Church, 1994).⁶⁷ Dabei geht es nicht nur um die äußere Gestaltung der Schnittstelle, sondern auch um auf die Bedürfnisse der Suchenden zugeschnittene Schnittstellenfunktionalität wie bspw. Datenexportfunktionen. Dryer (1997) betont in diesem Zusammenhang mit Blick auf die *User Interface Agents* genannten intelligenten Erweiterungen der Schnittstelle, daß solche Erweiterungen nicht einfach addiert werden sollten. Um einen intelligenten und konsistenten Gesamteindruck der Schnittstelle zu erzeugen, sei es vielmehr nötig, die verwendeten Techniken aufeinander und auf die für den Benutzer wichtigen Aufgaben zu beziehen. Es erscheint daher wichtig, nicht nur die Erweiterung der Systemfunktionalität (im Fall von OSIRIS des Datenbestandes und der Eingabeverarbeitung) im Auge zu behalten, sondern neben der Evaluation der bestehenden Schnittstelle (siehe 6.1) auch an grundlegend erneuerte Versionen zukünftiger Schnittstellen zu denken (Miller, 1997).

Mit Betonung der Benutzerschnittstelle könnten die in 5.1 und 5.2 geschilderten Erweiterungen von OSIRIS im nächsten Schritt zu einem umfassenden und benutzerfreundlichen Bibliothekssystem führen, das sich durch eine einheitliche Oberfläche mit nur einer Eingabemöglichkeit für alle Suchaspekte, durch Interaktion mit dem Benutzer und weitere Zusatzfunktionalität innerhalb der Schnittstelle auszeichnet. Zu diesen Zusatzfunktionen gehören eine direkt integrierte Bestellfunktion für Fernleihen und Magazin-Bestellungen, die Offline-Ausgabe der Rechercheergebnisse in verschiedenen gängigen Formaten (z.B. in BIBTEX), die integrierte oder direkt angebundene Ausleihverbuchung mit automatischer Email-Warnung vor Ende der Leihfrist, ein integrierter Zugang zu einer großen Online-Enzyklopädie, die genauere graphische Lokalisierung des Standortes eines Buches in den Räumen der Bibliothek und die einfache Möglichkeit, Bücher per Email anderen Benutzern zu empfehlen, wie man dies von Online-Zeitschriften kennt, deren Artikel man einem Bekannten per Mausklick weiterempfehlen kann.

⁶⁷Daß bestimmte Phänomene aber gerade nicht durch Verbesserungen der Oberfläche gelöst werden können, ist im Verlauf dieser Arbeit hinreichend gezeigt worden.

Ebenfalls der Verbesserung des Systemverhaltens zuträglich erscheint die konsequente Nennung der in der Bibliothek verantwortlichen Angestellten mit einem Photo: Neben einer sowieso notwendigen Hilfefunktion im System könnte durch einen mit „Kontakt“ beschrifteten Button kontextabhängig das Angebot einer direkten Kontaktaufnahme mit dem verantwortlichen Mitarbeiter der Bibliothek („Der in diesem Bereich für Sie zuständige Fachreferent ist ...“) hergestellt werden.

Stein (1998) betont zudem die Wichtigkeit einer aktiven Hilfefunktion, die situationsabhängig gesprochene Sprache oder natürlichsprachlichen Text verwendet. Dabei stellt Stein in einer Benutzerstudie fest, daß gerade für Anfänger die aktive Hilfefunktion (im Gegensatz zur statischen Hilfe, die denselben Inhalt hatte und daher theoretisch ebenso hilfreich hätte sein müssen) den Umgang mit dem System sehr erleichterte, wobei für die aktive Hilfe die Version mit gesprochener Sprache der Version mit geschriebener Sprache vorgezogen wurde. Stein erklärt dies mit der geringeren Ablenkung von der eigentlichen Aufgabe des Benutzers, die durch gesprochene Sprache verglichen mit Text ausgeübt werde.

Shneiderman (1997) betont den besonderen Nutzen von *direct manipulation user interfaces*, die für naive Benutzer geeignete Schnittstellen zu Programmierertools und großen Datenbanken darstellen sollen. „Direkte Manipulation“ im Sinne dieser Schnittstellen meint die bildliche Darstellung der zu manipulierenden Objekte, die bildliche Repräsentation der an den Objekten ausführbaren reversiblen Aktionen und die sofortige, inkrementelle Darstellung der Ergebnisse dieser Aktionen.

Vor dem Hintergrund der Überlegungen von Shneiderman (1997) kann bereits jetzt eine übernächste Version der OSIRIS-Systemschnittstelle skizziert werden – d.i. die Erweiterung der zu Beginn dieses Abschnittes dargestellten kommenden OSIRIS-Systemschnittstelle. Zur Zeit und auch in der geplanten nächsten Version der Schnittstelle werden die gefundenen Werke in einer Liste oder in Form abstrakter Klassen präsentiert, die durch gängige Icons symbolisiert werden. Stattdessen könnte aber ein photorealistischer Regalabschnitt mit verschiedenen Buchrücken und gut lesbaren Klassenbezeichnung am Regalbrett gezeigt werden, wie er sich dem Benutzer, der sich einem Regal in der Bibliothek nähert, auf den ersten Blick präsentiert. Ein Klick auf das Regal könnte dann zur gewohnten listenartigen Darstellung der einzelnen Titel führen. Der Vorteil liegt darin, daß dem Suchenden der Status der gefundenen Klasse auch graphisch als Klasse der lokalen Bibliothekssystematik deutlich gemacht wird. Nachteile dieser Darstellung können sein, daß virtuelle Klassen (siehe 3.1) zu Verwirrung führen können, weil sie in der Bibliothek ja nicht als Regalbrett existieren. Außerdem könnten Benutzer glauben, daß die schematische Darstellung der Bücher im Regal einen Zusammenhang zum wirklichen Aussehen der gefundenen Bücher hat – was in dieser Version nicht vorgesehen ist, siehe dazu aber den nächsten Absatz.

Eine weitere, derzeit noch utopische Erweiterung der Schnittstelle greift die Metapher des „Stöberns“ im Bestand der Bibliothek auf und versucht sie mit Hilfe einer 3D-Animation umzusetzen. Dabei führt ein Mausklick auf das Symbol für eine Klasse zu einem wenige Sekunden langen Flug, beginnend bei einer möglichst charakteri-

stischen Außenansicht der Bibliothek (z.B. der Blick auf die Gebäudefront mit dem Haupteingang aus der Augenhöhe eines Menschen), gefolgt von einem Zoom durch den Haupteingang in das Treppenhaus und einer kurzen (standardisierten) Bewegung nach oben, die dem Benutzer andeutet, daß er nun ein höheres Stockwerk erreicht (vielleicht wird das Bild überlagert von einer großen Zahl, die das Stockwerk bezeichnet). Im Zielstockwerk muß wieder eine möglichst charakteristische Ausgangsposition eingenommen werden, von der aus eine Kamerafahrt über verschiedene Landmarken (Infoplatz, Schilder etc.) zum Regal beginnt.

Der letzte Kameraschwenk zum Regal hin ist dann eine Überblendung auf ein Digitalphoto oder eine andere realistische Darstellung des Regals, auf der die zur gesuchten Klasse gehörigen Bücher deutlich markiert sind (erleuchtet, umrahmt o.ä.). In einer weiteren Ansicht (deren technische Umsetzung aber vermutlich sehr viel schwieriger ist, weil stark individualisiert) könnten auf einem Regalbrett alle Bücher der in Frage stehenden Klasse nebeneinanderstehen – mit realistischen Proportionen und Farben, um die Wiedererkennung zu gewährleisten.

Bei Überfahren der Buchrücken mit der Maus würde in einer nebenstehend eingeblendeten Liste der bibliographischen Angaben aller Bücher der Klasse der zum jeweils mit der Maus berührten Buch gehörige Datensatz farbig unterlegt und mit Ausleihinformationen versehen am oberen Bildschirmrand platziert werden. Um ein Buch vorzumerken oder aus dem Magazin zu bestellen, wären Ansätze der direkten Manipulation hilfreich, wenn nämlich ein Benutzer ein das Buch repräsentierendes Icon mit der Maus in den „Vormerken“-Korb oder in die „Verlängerungskiste“ ziehen kann.

Nach diesen (zum Teil utopischen) Entwürfen ist es wichtig, noch einmal festzuhalten, daß Interfacegestaltung und Systemfunktionalität nur gemeinsam zu einem besseren, d.h. für den Suchenden erfolgreich zu bedienenden System führen. Das Mißverständnis, allein eine einfache und ansprechende Oberfläche könne schon helfen, lag der Entwicklung der WWW-OPACs (siehe 2.2) zugrunde und zeigt sich auch wieder deutlich in Shneiderman et al. (1998).

Shneiderman et.al. beklagen, daß die Benutzer aktueller Retrievalsysteme überfordert seien. Zur Verbesserung dieser Situation schlagen sie vor, den Suchenden *alle* Möglichkeiten der Systemmanipulation zur Verfügung zu stellen, um ihnen so mehr Einfluß zu geben. Gleichzeitig soll das Interface aber natürlich so verständlich und intuitiv gestaltet sein, wie nur irgend möglich. Shneiderman et.al. ignorieren, daß die geforderte Verfügbarkeit aller Manipulationsmöglichkeiten zu einer deutlichen Belastung der so wieso schon überforderten Benutzer führt. Der einzige Ausweg, der sich anbietet, ist die Entlastung der Suchenden durch stärkere Systemintelligenz: Vor allem eine Verbesserung der Qualität der Rechercheergebnisse würde einen großen Teil der Suchparameter für den Nutzer überflüssig machen. Jedoch wird dieser Aspekt von Shneiderman et al. (1998) übersehen, da sie versuchen, auf beschriebene Art und Weise die Handhabung eines Systemes zu verbessern, das mit herkömmlichen Retrievalmethoden arbeitet, die den (sprachlichen) Daten nicht gerecht werden. Ähnlich Twidale und

Nichols (1997), die statt einer grundlegenden Verbesserung des Gesamtsystems eine Visualisierung von Fehlbedienungen als Hilfe für den Suchenden beschreiben.

Ein anderer Vorschlag für eine verbesserte Systemschnittstelle stammt von Light (1997), der ein System vorschlägt, das auf von Experten kodierten Metadaten operiert und die Interaktion mit dem Benutzer über ein neuentwickeltes graphisches Interface abwickelt. Während die (nicht weiter ausgearbeitete) Idee, Dokumente innerhalb eines domänenabhängigen Rasters von Experten mit Metadaten annotieren zu lassen, aus Kostengründen nicht durchführbar erscheint, ist die Idee einer graphischen Oberfläche, die die verschiedenen Textdimensionen in einem mehrdimensionalen Koordinatensystem repräsentiert, zunächst einleuchtend. Zusammen mit den aktuellen Entwicklungen von Auszeichnungssprachen im Rahmen von XML würde ein graphisch orientierter Zugriff auf Struktur und Inhalt von Dokumenten möglich, der eine Alternative zu den bislang gezeigten Schnittstellen wäre.

Diese Übertragung wäre aber eine grundlegende Erweiterung des Ansatzes von Light, denn dieser betont in seinem Artikel 1997 noch wesentlich die Leistung des Arbeitsplatzrechners, der das an mangelnder Bandbreite leidende Internet entlasten muß. Inwieweit hinter dieser – angesichts von Thin Clients und Net Clients heute überholten – Betonung der lokalen Rechnerressourcen die Produktpolitik von Lights Arbeitgeber Intel Corporation steht, muß hier offenbleiben. In der ursprünglich vorgeschlagenen Form erscheint der Ansatz von Light aufgrund der starken Voraussetzungen jedenfalls nicht realisierbar. Da eine offene Schnittstelle zwischen Oberfläche und Daten fehlt, ist auch nicht abzuschätzen, ob und wie einfach das System skalierbar ist.

Ahlberg und Shneiderman (1994) bemerken, daß die enormen Fähigkeiten des Menschen, visuelle Informationen zu verarbeiten, d.h. große Datenmengen auf einen Blick zu erfassen und interessante, auffällige Objekte zu erkennen, für die Schnittstellengestaltung genutzt werden sollten. Informationen sollten daher visualisiert und dem Benutzer in einem direct manipulation interface angeboten werden – nicht nur als Ersatz, sondern auch als Ergänzung herkömmlicher, textbasierter Schnittstellen.

Als Beispiel, wie dies aussehen könnte, beschreiben Ahlberg und Shneiderman eine Schnittstelle zu einer Filmdatenbank. Die Darstellung des gesamten Datenbestandes erfolgt in einem zweidimensionalen Koordinatensystem, wobei auf der X-Achse die (Entstehungs-)Zeit und auf der Y-Achse die Popularität (gemessen an Zuschauerzahlen) eingetragen sind. Die einzelnen Filme werden als farbige Punkte dargestellt (ohne textuelle Information), wobei die Farben für die Zugehörigkeit zu bestimmten Genres stehen. Eine Suchanfrage wird nicht-sprachlich mit Hilfe von Knöpfen und Schiebereglern realisiert. Die Dimensionen, mit denen gesucht werden kann sind *Title*, *Actor*, *Actress*, *Director*, *Length*, *Rating* und *Genre*. Das Ergebnis einer solchen Anfrage wird wiederum als Menge von Punkten im Koordinatensystem dargestellt. Der Benutzer kann in Bereiche der Zeit-Popularitäts-Matrix zoomen – die einzelnen Filme erscheinen dann größer (dichter herangeholt), und es wird textuelle Information (z.B. der Titel) gegeben. Ein Klick auf einen Film führt zur Vollanzeige.

Der Ansatz von Ahlberg und Shneiderman erscheint interessant in der Darstellung der

Ergebnisse, aber nicht unbedingt überzeugend in der nicht-sprachlichen Realisation der Bedienelemente. Obwohl es einerseits natürlich einen großen Vorteil darstellt, daß jede Veränderung an einem der Regler sofort einen visuellen Effekt hat, ist andererseits aber nicht ersichtlich, wie für komplexere Suchdomänen die Dimensionen der Suche übersichtlich mit Hilfe von Reglern und Knöpfen dargestellt werden können. Hier erscheint eine sprachliche Eingabe vorteilhafter. Auch ist für andere Domänen nicht klar, welche Aspekte auf die Achsen des Koordinatensystems gelegt werden sollen. Zeit mag bei Büchern noch eine gewisse Rolle spielen, aber ob menschliche Experten hiermit gut zu gruppieren sind, darf bezweifelt werden. Ein Wechsel in der Belegung der Koordinatenachsen zwischen Systemen aber führt vermutlich zu Verwirrung und sollte durch eine auffällig andere Gestaltung der Schnittstelle zu kompensieren versucht werden.

Insgesamt erinnern die „starfield displays“ von Ahlberg und Shneiderman stark an Landkarten, eine Möglichkeit der effektiven Informationspräsentation, die auch von Hitch et al. (1986) vorgeschlagen wird. Hitch et al. halten (Land-)Karten für eine lohnende Alternative zu herkömmlichen Systemschnittstellen, da Karten die Vorteile von Schnittstellen mit direkter Manipulationsmöglichkeit weiter steigern würden: Im Vergleich mit Icons werden dem Benutzer nicht nur die zu manipulierenden Objekte, sondern auch deren Verbindungen untereinander veranschaulicht. Hitch et al. schreiben, daß die räumliche Anordnung der Elemente wie bei einer Icon-basierten Oberfläche die Benutzer in ihrem Erinnerungsvermögen bezüglich der Eigenschaften einzelner Objekte unterstütze. Zudem bieten Landkarten den Vorteil, daß sie stets eine Darstellung des gesamten Systems und dessen Struktur sind. Allerdings erscheint es nicht immer einfach möglich, die Suchattribute einer Domäne auf die räumlichen Verhältnisse in einer Landkarte abzubilden.

Hitch et al. versuchen empirisch nachzuweisen, daß Landkarten gegenüber Menüsystemen im Vorteil sind und kommen zu dem Schluß, daß Landkarten „are supportive for naive users and yet permit the use of more efficient, rapid modes of operation for experienced users“ (Hitch et al., 1986, :583). Allerdings überzeugt der Versuch nicht: Die Aufgabe der (englischen) Versuchspersonen lautete, in einer Datenbank nach Vorwahlen zu englischen Städten zu suchen, wobei die Datenbank in drei Hierarchiestufen nach „Towns“, „Counties“ und „Regions“ organisiert war. Es wurden zwei Arten von Landkarten verwendet, solche, die mit den geographischen Verhältnissen Englands korrespondierten, und solche mit anderer, zufälliger Struktur. Daß sich kein Unterschied in der Verwendung der beiden Karten nachweisen ließ, liegt sicherlich an der extremen Abstraktion der Darstellung: Beide Karten waren lediglich in Form einer Windrose aufgebaute Strichdiagramme – bei der geographisch korrekten Karte war die Region X vom geographischen Mittelpunkt Englands aus gesehen korrekt in Richtung NE vermerkt und auf der anderen Karte an einem zufälligen anderen Ort.

Aus technischen Gründen wurden die Karten zudem nicht am Schirm, sondern als Diagramme auf Papier präsentiert, die Interaktion mit dem System erfolgte über die

Tastatur mit Hilfe der Abkürzungen für die Himmelsrichtungen. Man kann Hitch et al. daher nur zustimmen, wenn sie schreiben, daß die Einsatzmöglichkeiten von Karten von der Art der Aufgabe, der Struktur und dem Inhalt der Datenbank abhängen.

Auch in OSIRIS waren für bestimmte Aufgaben Landkarten als Repräsentationsmedium geplant. So könnte mit Hilfe neuronaler Netze eine weitere Feinunterteilung unterhalb der Klassenebene möglich sein, wenn z.B. in der Klasse „Blues und Jazz“ weiter nach „Blues“ differenziert werden soll. Diese Feinunterteilung könnte in Form mausensensitiver 2-dimensionaler Karten dargestellt werden und beim Betrachter Assoziationen wie „die Blues-Insel im Swing-Meer vor der Küste des Jazz-Kontinents“ auslösen. Als mögliche Erweiterung des OSIRIS-Systems ist dies aber vermutlich weniger aufgrund der Darstellung, als durch den Einsatz neuronaler Netze auf klassifikatorisch nicht mehr ausreichend erschlossenen (weil innerhalb *einer* Klasse befindlichen) Daten interessant (vgl. 5.2.3). Die Darstellung in Form einer automatisch generierten Karte erscheint bereits deshalb problematisch, weil die räumlichen Verhältnisse nicht konstant bleiben, sondern durch Änderungen im Klassenbestand durch die Netze im Laufe der Zeit jeweils neu und anders interpretiert werden. Zudem stellt sich die Frage nach der Zahl und der Art der zu berücksichtigen Attribute. Vermutlich wird es den Suchenden eher verwirren, in der Darstellung einer Karte eine Mischung aus inhaltlichen und formalen Attributen zu finden („Blues“ neben „mit Noten“), und es ist offen, welche und wieviele der möglichen Attribute in so eine Darstellung aufgenommen werden sollen. Die Voraussetzung für den sinnvollen Einsatz solcher Techniken scheinen zu sein:

1. Die Darstellungen müssen – obwohl automatisch generiert – in der Konfiguration möglichst konstant, d.h. für den Betrachter einfach wiedererkennbar gehalten werden.
2. Die darstellbaren Attribute müssen zu homogenen Klassen zusammengefaßt werden können. Nur solche Attribute, die einer Klasse angehören, sollten gemeinsam in einer Karte visualisiert werden.
3. Es bedarf genauerer Untersuchungen, wie Benutzer auf eine Ergebnisdarstellung reagieren, die zwar an eine Landkarte erinnert, aber keine geographischen, sondern z.B. bibliographische Informationen zeigt.

Hemmje (1993) bemerkt, daß Informationssuchende sich immer noch häufig mit Anfragesprachen oder Datenbanktechnologie auseinandersetzen müssen, im Gegensatz zu Benutzern von z.B. CAD Systemen, bei denen aufgrund der Aufgabenstellung eine graphische Schnittstelle verfügbar ist. Dies hat mit den eigentlichen Aufgaben der Benutzer nichts zu tun und führt zu einer zusätzlichen kognitiven Belastung bei der Suche. Um diese kognitive Belastung zu reduzieren, schlägt Hemmje (1993) vor, sich auf die natürlichen Fähigkeiten des Menschen zur Orientierung und Navigation in Räumen zu verlassen und diese Fähigkeiten für ein Information Retrieval System mit einer

dreidimensionalen Darstellung des Dokumentenraumes zu nutzen. Werden Dokumente über die in ihnen verwendeten Begriffe als Vektor aufgefaßt, und wird das Suchinteresse der Benutzer ebenfalls als (Begriffs-)Vektor repräsentiert, so kann der Raum der für die Bedürfnisse des Benutzers relevanten Dokumente visualisiert werden.

In der konkreten Umsetzung verwendet Hemmje Kegelbäume zur Visualisierung, da diese durch perspektivische Verzerrung und Schattierung eine räumliche Tiefe andeuten können. Die Frage der hierarchischen Ordnung der Begriffe in einem solchen Baum löst Hemmje, indem die Kegelbäume lediglich chronologisch die Interaktion des Benutzers mit dem System wiedergeben sollen. Die Kegelbäume werden interaktiv entfaltet, beginnend mit dem ersten Begriff, den der Benutzer im Rahmen seiner Suche eingibt, und bestehen immer abwechselnd aus (Such)Begriffen und Dokumentknoten. Kegelbäume können, wie Hemmje es auch in der Einleitung seines Artikels andeutet, als hierarchische Repräsentation ontologischer Abhängigkeiten genutzt werden, in der beschriebenen Anwendung jedoch handelt es sich lediglich um die Visualisierung der Interaktion zwischen Benutzer und System: Ein Benutzer gibt in Schritt 1 einen (möglicherweise komplexen) Suchbegriff A ein, und das System antwortet mit einer Menge von Dokumenten, für die der gesuchte Begriff A inhaltlich charakteristisch ist⁶⁸ und repräsentiert diese Dokumente als Flächen auf dem Kegel der Ebene 2. Der Benutzer hat an dieser Stelle die Möglichkeit, eines dieser Dokumente auszuwählen und entweder dieses im Volltext zu betrachten, oder sich die das ausgewählte Dokument charakterisierenden Begriffe als Kegelbaum der Ebene 3 anzeigen zu lassen. Eine Auswahl eines Begriffes B auf Ebene 3 führt dann wiederum als Antwort des Systems zu einer Menge von Dokumenten, für die die Suchbegriffe A und B (auf dem Pfad von der Wurzel des Baumes bis zur aktuellen Position) charakteristisch sind, und die als Kegelbaum der Ebene 4 dargestellt werden usw.

Obwohl die Darstellung und die Navigationsmöglichkeiten bei Hemmje im Ansatz vielversprechend aussehen, müssen doch zwei Dinge kritisiert werden: Zum einen nützt eine vermutlich für den Benutzer übersichtliche und verständliche Darstellung der Suchergebnisse wenig, wenn die Ergebnisse selbst unzulänglich sind. Dies muß im vorliegenden Fall aber aufgrund der Suche mit Stichwörtern befürchtet werden. Zum anderen scheinen die Möglichkeiten der Ergebnisdarstellung mit Kegelbäumen nicht ausgereizt, da nur die Interaktion des Benutzers mit dem System dargestellt wird und so innerhalb des Baumes sich sprachliche Begriffe und Dokumente mischen. Wünschenswert wäre vielmehr eine homogene und ontologisch sortierte Darstellung. Der Vorschlag zur Ergebnisdarstellung bei Hemmje (1993) erscheint also durchaus ausbaufähig, ist aber noch nicht überzeugend umgesetzt. Zudem steht der Nachweis in Form einer software-ergonomischen Untersuchung noch aus, ob diese Art der Darstellung für den Benutzer wirklich verständlicher ist.

Die Interaktion des Benutzers mit dem System kann weiter verbessert werden, wenn es

⁶⁸Hemmje beschreibt die Anreicherung eines existierenden, probabilistischen Information Retrieval Systems um eine neue Oberfläche – wie das zugrundeliegende System für einen Begriff bestimmt, ob dieser für das Dokument inhaltlich charakteristisch ist bleibt offen.

gelingt, zusätzlich zur Spracheingabe⁶⁹ auch Zeigegesten des Benutzers zu erkennen und Zeigegesten des Systems für den Benutzer zu ermöglichen.⁷⁰ Für die Erkennung von Zeigegesten des Benutzers sind zwei Möglichkeiten denkbar, nämlich einerseits die Erkennung von Zeigeaktionen des Benutzers innerhalb des Systems, z.B. mit Hilfe der Maus auf der Systemoberfläche sowie andererseits die Erkennung von Zeigegesten des Menschen außerhalb des Systems, z.B. durch Videoerkennung o.ä.

Einen sehr frühen Ansatz für die Integration von Zeigegesten und deiktischen Ausdrücken stellen Cohen et al. (1989) vor. In ihrer Schnittstelle kann der Benutzer Demonstrativpronomina verwenden und dazu auf ein Objekt „zeigen“, indem er mit Hilfe der Tastatur aus einer Menge von Objektbezeichnern wie z.B. `rework1` und `rework2` auswählt. Die Reduktion der Ambiguität bei der Anaphernaauflösung, bei der Auflösung elliptischer Folgefragen und auch bei der Auswahl der zu den deiktischen Ausdrücken verfügbaren Objekte ist geschickt durch eine explizite graphische Repräsentation des Kontextes einer Äußerung (mittels eines Rahmens) gelöst. Das System ist damit insgesamt zwar ein interessanter erster Schritt in Richtung auf eine deutliche Erweiterung der Eingabemodalitäten, aber im Ergebnis noch so kompliziert und unnatürlich zu bedienen, daß es lediglich als ein erster Entwurf angesehen werden kann.

Deutlich besser wird die Erkennung von Zeigegesten⁷¹ des Benutzers in EDWARD (Huls und Bos, 1993, 1998) umgesetzt. EDWARD hat zum Ziel, dem Benutzer eine vollständig integrierte, multimodale Schnittstelle zur Verfügung zu stellen, die nicht dem Benutzer die Interaktionsart vorgibt (sei es aufgrund von Antizipation oder Adaptivität), sondern dem Benutzer die Wahl läßt, welche Modi er gerne nutzen und vor allem kombiniert verwenden möchte. EDWARD ist daher in der Lage, natürlich-sprachliche Benutzereingaben wie „Zeige mir den Inhalt *dieses* Ordners.“ zu verarbeiten, wenn die Eingabe von einer Zeigegeste mit der Maus begleitet wird, die einen Ordner bezeichnet. Weitere Eingabemodi, die ebenfalls untereinander und auch mit natürlicher Sprache und Zeigegesten kombiniert werden können, sind Menüs und eine Kommandosprache. Diese Interaktionsmöglichkeiten wiederum können verschiedene Kanäle nutzen: Die Eingabe kann mit Hilfe einer Maus, über die Tastatur oder mittels eines Handschriftenerkennungsgerätes durchgeführt werden, die Ausgabe kann in Form von Bildschirmtext oder einer synthetischen Stimme erfolgen. Die Systemausgabe kann weiterhin durch graphisch dargestellte Zeigegesten unterstützt werden, wenn nämlich das System bei der Textausgabe deiktische, anaphorische oder pronominale Ausdrücke durch einen Pfeil betont und auflöst (vgl. auch Cremers (1998)).

Diese Technik kann auch zur Aufmerksamkeitssteuerung verwendet werden, wenn EDWARD den Benutzer auf eine Änderung der Systemwelt hinweisen möchte. Wird auf mehrere Objekte nacheinander mit einem Pfeil hingewiesen, so bleiben die bereits bezeichneten Objekte graphisch hervorgehoben, um für den Benutzer eine nachvollziehbare Spur der graphischen Ausgabe zu generieren. EDWARD generiert weiterhin

⁶⁹Unabhängig davon, ob es sich um gesprochene Sprache oder Eingaben über die Tastatur handelt.

⁷⁰„The desire to point while speaking is often irresistible.“(Cohen et al., 1989, :231)

⁷¹Zeigegesten innerhalb des Systems.

linguistische Beschreibungen der Interaktion mit dem Benutzer: Wird ein Icon auf der Oberfläche vom Benutzer selektiert, gibt das System automatisch zusätzliche Informationen dazu („Dies ist der Jahresbericht der DFG 1999“). Wenn der Benutzer ein Objekt löscht, z.B. indem er dessen Icon durch direkte Manipulation mit der Maus in den Papierkorb zieht, beschreibt das System den Vorgang natürlichsprachlich mit „Sie sind dabei, die Dissertation von Meyer zu löschen“, um dem Benutzer im Fall eines Versehens (z.B. weil er sich in der Referenz des Dateisymbols getäuscht hatte) zu warnen. In einer Benutzeruntersuchung (Huls und Bos, 1998) kommen die Autoren zu dem Schluß, daß die linguistischen Beschreibungen der Benutzerinteraktion durch das System den Nutzern helfen, da diese im Vergleich weniger Fehler machen und weniger Zeit für die Aufgaben benötigen, indem sie die Aufgaben mit weniger Aktionen lösen können.

Auch das DENK-System (Bunt et al., 1998) verfolgt das Ziel, dem Benutzer eine vollständig integrierte Schnittstelle zu bieten, an der natürlichsprachliche Eingaben über die Tastatur mit Zeigegesten auf der Oberfläche des Systems kombiniert werden können. Dazu ist eine systeminterne Repräsentation der sichtbaren Oberfläche sowie eine ständige Aktualisierung dieser Repräsentation nötig, um Benutzerfragen nach Objekten dieser Domäne beantworten und auf Benutzeraktionen innerhalb dieser Domäne reagieren zu können (zu DENK siehe auch Beun (1993)).

Ähnlich wie EDWARD und das DENK-System versuchen auch Stock und Strapparava (1993) mit ihrem ALFRESCO System einen integrierten, multimodalen Dialog zu ermöglichen (s. a. Stock et al. (1997)). Durch die Anwendungsdomäne (Fragen zu italienischen Fresken) ist die Komplexität des Gegenstandes, über den der Benutzer mit dem System verhandelt, ungleich höher. Der Benutzer kann dem System z.B. Fragen stellen wie „In welchen Fresken hat Giotto diesen Heiligen abgebildet?“, wobei die Verwendung des Demonstrativartikels einhergeht mit einer Berührung des touch screens an der Stelle, an der der entsprechende Heilige gezeigt wird. Dazu ist natürlich eine genaue Repräsentation der Inhalte und des Aufbaus der dem System bekannten Fresken erforderlich. Es ist daher verständlich, daß die Autoren sich darauf beschränken, lediglich die Figuren im Bildvordergrund zu repräsentieren, den Bildhintergrund aber, der bei Fresken aus beliebigen, oft sehr detaillierten Alltagsszenen bestehen kann, unberücksichtigt zu lassen. Obwohl diese Entscheidung den Aufwand der Systemerstellung reduziert, bleibt offen, ob die Benutzer diese Einschränkung erkennen und nachvollziehen können: Das Bedürfnis, zu einer Figur des reichhaltig gestalteten Bildhintergrundes eine Frage zu stellen, ist nachvollziehbar, und diese Frage durch eine Berührung des entsprechenden Bildschirmbereiches zu unterstützen, erscheint durchaus sinnvoll.

Im Unterschied zu den beiden bisher genannten Systemem liefert ALFRESCO auch eine graphische Repräsentation der Analyse der natürlichsprachlichen Benutzereingabe und versucht dem Benutzer so die Möglichkeit eines Metadialoges zu eröffnen. Durch die Darstellung der Frage-Antwort-Paare des Diskurses und die als Antezedenten verfügbaren Objekte soll der Benutzer eine bessere Kontrolle der Systeminteraktion er-

halten und auch Eingaben wie „Kommen wir auf meine vorherige Frage zurück ...“ erfolgreich tätigen können. Ob diese an sich sinnvolle Funktionalität aber dem Benutzer vermittelt werden kann durch eine einfache graphische Darstellung von etwas, das an thematische Rollen erinnert, darf bezweifelt werden. Die gesamte Darstellung der systeminternen Analyseergebnisse ist so unübersichtlich und aufwendig, daß der Nutzen für den Suchenden in keinem Verhältnis zum (vermuteten) kognitiven Aufwand steht. Zudem wird soviel Raum auf dem Bildschirm zu diesem Zweck verbraucht, daß unklar bleibt, wo die eigentlichen Fresken abgebildet werden, oder ob es sich um einen zusätzlichen Darstellungsmodus handelt.

ALFRESCO kann bei der Antwortgenerierung zwischen verschiedenen Präsentationsmodi (Verbalisierung, tabellarische Zusammenfassung, Photo, Animation etc.) wählen. Komplexe Informationen („Wieviele Forscher gehören den einzelnen Gruppen des Institutes an?“) können durch Balkendiagramme oder andere Infographiken ausgedrückt werden und Fragen nach der Lokalisierung von Räumen oder Gebäuden durch die Anzeige von Karten oder Raumplänen. Über die Wahl einer angemessenen Darstellungsart hinaus soll ALFRESCO auch die Inhalte der Systemantwort strukturieren: Für weiterführende, im ersten Schritt nicht präsentierte Informationen sollen automatisch Hinweise an den inhaltlich relevanten Stellen z.B. in Form von Knöpfen generiert werden, um einen individuell vom Benutzer zu explorierenden Hypertext zu generieren. Was im Prinzip einleuchtend erscheint, ist in der Praxis aber nur dann sinnvoll, wenn eine Komponente vorhanden ist, die die einzelnen Präsentationsmodi hinsichtlich ihrer Nützlichkeit für die jeweils darzustellende Information bewertet und daraus resultierend die eigentliche Präsentation plant. Über diese für den erfolgreichen Einsatz adaptiver Präsentationsmodi entscheidende Komponente schweigen sich Stock und Strapparava (1993) aber aus.

In die gleiche Richtung argumentieren auch André und Rist (1993), wenn sie versuchen, in ihrem System WIP einen systematischen Zusammenhang zwischen Bild und Text bei der (online) Generierung von Systemreaktionen zu integrieren. Ziel ist es natürlich, die jeweiligen Vorteile der einzelnen Medien zu nutzen, um die Informationen möglichst verständlich und der Situation angemessen zu präsentieren. André und Rist behaupten, daß sich viele der in Texten gefundenen Relationen zwischen Textteilen wie z.B. „Elaboration“, „Sequenz“ oder „Interpretation“ auch zwischen Bildern, Bildteilen oder auch Text und Bildern oder Bildteilen nachweisen lassen. Unter dieser Annahme ist die Ausweitung der Organisationsprinzipien textueller Systemreaktionen auf bildliche Teile der Ausgabe nur folgerichtig. André und Rist versuchen, die rhetorische Struktur einer multimodalen Systemausgabe zu bestimmen, für deren Teile dann relativ zur Informationsart (räumlich, zeitlich etc.) und zur kommunikativen Aufgabe (vergleichen, zusammenfassen, Aufmerksamkeit generieren etc.) einzelne Präsentationsmedien ausgewählt werden.

Als Unterschied z.B. zu EDWARD führen André und Rist (1993) an, daß in WIP nicht nur die Darstellung auf Objektebene intern repräsentiert wird, um die Planung einer Systemausgabe auch mit nicht-textuellen Elementen zu ermöglichen, sondern auch

auf Ebene der Merkmale. So werde es möglich, über die Farbe (z.B. „rot“) eines Teils einer Graphik diesem Teil eine Eigenschaft (z.B. „defekt“ zu sein) zuzuschreiben.

„Tellim“, ein System zur multimedialen Produktpräsentation im WWW (Jörding et al. (1998) und Jörding und Meissner (1998)) versucht ebenfalls die Präsentation auf die (erkannten) Vorlieben und die Situation des Benutzers abzustimmen. Dazu wird einerseits mit Hilfe eines inkrementellen Lernalgorithmus versucht, aufgrund der Nutzerinteraktion Aussagen über die Interessenlage des Benutzers zu tätigen. Andererseits werden die Mittel der Präsentation (Graphiken, Texte, Bilder etc.) nach ihren Ansprüchen an die Übertragungsbandbreite gewichtet und auf die jeweils herrschende Netzlast hin abgestimmt eingesetzt, um ein Optimum an Darstellungsqualität und Darstellungsgeschwindigkeit sowie (aufbauend auf der Interpretation der Interaktion) an Individualisierung der Systemreaktion zu ermöglichen. Für diesen aus wirtschaftlichen Erwägungen heraus derzeit überaus florierenden Bereich stellt „Tellim“ nur ein Beispiel von vielen dar. Wie weit die Überlegungen gehen, eine individualisierte Sicht, abgestimmt auf die Interessen und die aktuelle Situation des Benutzers zu generieren, beweist Jameson (1998), wenn er auch Situationsconstraints wie Zeitdauer und Ablenkung des Benutzers in Betracht zieht.

Die Erkennung von Zeigegesten des Menschen außerhalb des Systems ist technisch und konzeptionell sehr viel aufwendiger. Jung et al. (2000) stellen ein System zum virtuellen Konstruieren vor, das gesprochene natürliche Sprache und natürliche koverbale Gesten vor einer interaktiven Wand verarbeiten kann. Bei den Gesten werden drei Typen unterschieden. Die *deiktischen Gesten* sind das systemexterne Äquivalent zu den Zeigegesten mit der Maus, die für die vorangegangenen Systeme beschrieben wurden. Eine deiktische Geste spezifiziert zusammen mit einer sprachlichen Äußerung ein Objekt oder einen Ort. *Mimetische Gesten* qualifizieren eine sprachlich geäußerte Handlung, indem z.B. gestisch die Richtung einer sprachlich eingeführten Bewegung angegeben wird. *Ikonische Gesten* schließlich beschreiben die Eigenschaften von Objekten, wenn z.B. die Form eines sprachlich genannten Objektes gestisch beschrieben wird. Die letzteren beiden Gestenarten haben kein Pendant in den bislang betrachteten Systemen.

Technisch erfolgt die Erkennung der Gesten mittels elektromagnetischer Positionssensoren am Körper und Datenhandschuhen an den Händen des Benutzers (die zugehörige Spracherkennung ist sprecherunabhängig). Es wird z.B. für eine Zeigegeste mit Hilfe des ausgestreckten Zeigefingers keine Bilderkennung betrieben, sondern es wird allein anhand der Daten zur Position der Hand im Raum, der Lage der Finger relativ zum Handkörper und relativ zueinander auf eine Konstellation geschlossen, die eine Zeigegeste sein könnte und deren Richtung im Raum ermittelt werden kann. Welche Merkmale für eine Zeigegeste konstitutiv sind, kann mit Hilfe eines für die Gestensprache entwickelten Notationssystems festgelegt werden. Schwierig ist die Erkennung bedeutungstragender Gesten in den stets vorliegenden Bewegungen der Person – hier spielen starke Abweichungen von zuvor erkannten Ruhepositionen, starke Beschleunigung oder auch symmetrische Bewegungen eine Rolle. Welche Bedeutung letztlich

einer Geste über ihren Grundtypus („Zeigegeste“) hinaus zugewiesen wird, kann nur im Kontext der sprachlichen Interaktion geklärt werden. Durch diese stets koverbale Interpretation von Gesten bestehen aber sehr hohe Ansprüche an die zeitliche Auflösung und die Verarbeitungsgeschwindigkeit, um den Zusammenhang zwischen sprachlicher und gestischer Benutzeraktion herstellen zu können. Durch die enge Domäne (virtuelles Konstruieren) und die vollständige Repräsentation der Objekte der Domäne und deren Eigenschaften (ein endlicher Satz an Bauteilen und deren Kombinationsmöglichkeiten) wird die Interpretation oftmals stark ambiger Gesten und Äußerungen vereinfacht.

Obwohl Interaktionsmöglichkeiten, wie sie Jung et al. (2000) beschreiben, sich zur Zeit noch im Forschungsstadium befinden und für erweiterte Kontexte durchaus nicht abzusehen ist, wie der dann dramatisch wachsende Anteil der Pragmatik in den sprachlichen wie gestischen Benutzeraktionen handhabbar sein wird, scheint dies langfristig dennoch die richtige Richtung zu sein, mit dem Benutzer einer Schnittstelle umfassend zu interagieren. Nicht alle Anwendungen werden solche Schnittstellen benötigen, aber ein unbestreitbarer Vorteil wird die zunehmende Freiheit des Menschen an solchen Schnittstellen sein, wenn Datenhandschuhe und Positionssensoren entweder standardmäßig in Kleidungsstücken integriert oder Wege gefunden werden, die benötigten Daten ohne Berührung des Akteurs (z.B. durch Echtzeit-3D-Scannen) zu erheben. Bereits jetzt aber wird über Anwendungsszenarien nachgedacht, die durch die von Jung et al. vorgestellten Möglichkeiten erst umgesetzt werden können – Grasso et al. (1998) beschreiben bspw. neue Interaktionsformen in kulturell vernetzten Städten, für die neue Formen der Mensch-Maschine-Schnittstelle im Sinne von Jung et al. gefunden werden müssen.

Wie bereits zu Beginn dieses Kapitels erwähnt, sind generelle Überlegungen, mit welchen Konzepten Mensch-Maschine-Schnittstellen verbessert werden können, wichtig. Nur so können neue, vielleicht erst in einigen Jahren tragfähige Lösungen angeschoben oder einsatzreife Systeme in die richtige Richtung weiterentwickelt werden. Lieberman (1996) beispielsweise, dessen System „Letizia“ in 5.2.3 angesprochen wird, propagiert die Medienkonvergenz und versucht eine Schnittstelle zu entwerfen, die die „zapping“ genannte Herangehensweise des modernen Fernsehkonsumenten auf das WWW überträgt. Dazu versucht ein in den Browser-Client integrierter Agent (ein „guide“ in der Terminologie von Dryer (1997)) die Vorlieben und Interessen des Betrachters zu erkennen und aus den in der aktuell betrachteten WWW-Seite weiterführenden Links diejenigen auszuwählen, die diesen Vorlieben entsprechen. So werden nicht nur Ladezeiten reduziert, sondern der Benutzer wird auch aktiv vom System aufgefordert, andere und als interessant eingestufte Angebote des WWW zu nutzen.

Von stetig wachsender wirtschaftlicher Bedeutung sind die Überlegungen, die Buxton et al. (1989) bereits 1989 diskutierten: Software für den Endbenutzer, auch Seiten im WWW, sollten in eine Geschichte eingebettet sein, um das Verständnis des Benutzers durch einen stärkeren Zusammenhalt zwischen den Informationsangeboten zu

befördern und den Reiz der längerfristigen Nutzung des Angebotes zu erhöhen.⁷² Die wirtschaftliche Bedeutung dieses gar nicht so neuen Ansatzes liegt für aktuell im Internet operierende E-Commerce-Unternehmen in der stärkeren Kundenbindung, die durch eine unterhaltsame und insbesondere für die stetig wachsende Zahl von Nutzern ohne Vorkenntnisse verständlichere Interaktion an das Unternehmen gebunden werden. Dabei ist die Einbettung des Online-Angebotes in eine Geschichte durchaus im Sinne des Theaters gemeint, wenn Auftritte im Internet inszeniert und dramatisch arrangiert werden sollen: Die darstellenden Künste haben seit sehr langer Zeit Erfahrungen darin, Betrachtern mit einfachen Mitteln der Abstraktion, der Metaphorik und der Überhöhung in einem zeitlich und räumlich engen Rahmen Gefühle und Eindrücke zu entlocken, die nicht selten mit einer tiefen inneren Anteilnahme einhergehen, bis zu einem gewissen Grade gezielt hervorgerufen und eingesetzt werden können und oft einen bleibenden Eindruck hinterlassen. Diese Effekte auch nur ansatzweise für Mensch-Maschine-Schnittstellen umzusetzen, könnte enorme Auswirkungen auf die Zufriedenheit der Benutzer haben.

5.2.3 Unterstützung und Verbesserungen der Suchverfahren

Nach dem Überblick über die Erweiterungsmöglichkeiten des Datenbestandes und den Verbesserungsmöglichkeiten an der Schnittstelle soll es nun um Verfahren gehen, die den Suchprozeß selbst verbessern oder erweitern können. Möglichkeiten sind hier z.B. die Anreicherung des Systems mit ontologischen Informationen und der Versuch, Dokumente automatisch zu klassifizieren. Manchmal wird in diesem Zusammenhang auch aus dem Scheitern oder den Schwierigkeiten bislang eingesetzter Verfahren der Schluß gezogen, daß eine grundlegende Neubestimmung der Werkzeuge und Verfahren nötig sei. Winiwarter (1996) konstatiert z.B. in seinem Überblicksartikel zu den Werkzeugen, mit denen man der Informationsflut Herr werden könne, eine Abkehr von den wissensbasierten Ansätzen hin zu statistischen und konnektionistischen Ansätzen bei der Sprachverarbeitung. Indem Winiwarter die Schwächen älterer Ansätze aus der KI kritisiert, deren „Toy Systems“ (siehe auch 4.1) auf die Hoffnung bauten, die für Beispieldaten gezeigte Lösung später auf größere Datenbestände ausweiten zu können, meint er die Ablösung des Paradigmas des Textverstehens durch die Informationsextraktion begründen zu können. Wie die von Winiwarter vorgestellten statistischen und konnektionistischen Werkzeuge allerdings mit pragmatischen und komplexen syntaktischen Phänomenen umgehen sollen, bleibt offen. Der bloße Verweis auf den Erfolg dieser einfachen Verfahren im Alltagsbetrieb z.B. von Spracherkennern ist unzureichend – das Vertrauen, durch Erweiterung und Verfeinerung z.B. statistischer Verfahren auch elaboriertere Phänomene behandeln zu können, steht der Hoffnung in die Erweiterbarkeit von „Toy Systems“ in nichts nach.

⁷²In gewisser Hinsicht gehört in diesen Bereich auch die Idee einer Kamerafahrt für die Ergebnispräsentation in OSIRIS (s. S. 186), die dem Benutzer den Zusammenhang zwischen seinem Suchergebnis und der Situation in der Bibliothek möglichst plastisch und eindrücklich vor Augen führen soll. Zur Integration von Geschichten in Retrievalsysteme vrgl. auch Mark Pejtersen (1998).

Ganz offenbar sind noch so viele Probleme zu lösen und noch so viele Wege zur Verbesserung des status quo offen, daß die Beschäftigung mit den unterschiedlichsten Verfahren und Werkzeugen immer noch sinnvoll erscheint.

Integration ontologischer Information Eine zentrale Stärke von OSIRIS war es, daß die bislang für das Retrieval im Katalog kaum genutzte Bibliothekssystematik als wichtiges Element des Suchverfahrens eingesetzt werden konnte (vgl. 3.1). Der allgemeine Nutzen des Einsatzes einer Ontologie beim Information Retrieval wird auch von anderen erkannt, wobei insbesondere die Bereiche von Interesse sind, in denen formale Merkmale, die eine Einordnung in eine Ontologie erleichtern, bereits in den Daten vorhanden sind. Weinstein (1998) beschreibt z.B. für bibliographische Daten aus der Medizin, wie (vorhandene) Metadaten⁷³ auf eine Ontologie abgebildet und über deren hierarchische Beziehungen die Suchmöglichkeiten verbessert werden können. Embley et al. (1998) beschreiben ebenfalls den Einsatz einer Ontologie für das Information Retrieval, im Gegensatz zu Weinstein (1998) allerdings für unstrukturierte Dokumente, für die noch keinerlei Metadaten vorliegen. Voraussetzung dazu ist eine detaillierte Ontologie einer bestimmten Domäne sowie zu klassifizierende Dokumente, von denen mit Sicherheit gesagt werden kann, daß sie (z.B. aufgrund ihrer Quelle) aus der Domäne der Ontologie stammen. Unter diesen (in der Praxis nicht trivialen) Voraussetzungen versuchen Embley et al. mit Hilfe manuell erstellter und parametrisierter syntaktischer Regeln im Volltext eines Dokumentes gemäß der Ontologie einschlägige Konzepte und deren Beziehungen untereinander zu erkennen. Das Ergebnis dieses Vorgehens ist laut Embley et al. nicht allein die Klassifikation eines Dokumentes, sondern die Extraktion der (im Sinne der Ontologie relevanten) Information, d.h. die Verfügbarkeit der Konzepte und deren Relationen zueinander in einer Datenbank.

Das Bedürfnis der Benutzer von Faktendatenbanken (z.B. einer Datenbank mit Meßwerten zu den Eigenschaften verschiedener Werkstoffe), auch in den Suchanfragen unscharfe Ausdrücke zu verwenden und entsprechend unscharfes Wissen der Datenbank zu erfragen, wird nach Roppel (1993) durch Anfragen wie die folgenden belegt, die einem Simulationstest entnommen wurden:

- „Werkstoff auf Nickel-Basis mit *möglichst niedriger* Wärmeausdehnung.“
- „Nickelbasislegierung mit *minimalem* Längenausdehnungskoeffizienten.“
- „Gibt es Daten zur chemischen Analyse eines *ähnlichen* Werkstoffs?“
- „Welche *charakteristischen* Festigkeitswerte hat IN713?“

Neben der (problematischen) Interpretation der linguistischen Ausdrücke⁷⁴ ist insbesondere auch deren Kontextabhängigkeit ein Problem: So ist eine Temperatur von 800

⁷³Im verbreiteten **Machine Readable Cataloging** (MARC) Format.

⁷⁴Vgl. 4.2.15 für die ähnlich schwierige Interpretation von non-standard Quantoren.

Grad Celsius für eine Titan-Legierung extrem hoch, für eine Nickel-Basis-Legierung jedoch nicht. Solche Abhängigkeiten können aber nur mit Hilfe einer Ontologie der Domäne (in diesem Fall der Werkstoffe) gelöst werden.

Allerdings kann die Integration einer Ontologie und die Abbildung der Daten auf diese Ontologie immer nur eine von mehreren Maßnahmen für die Verbesserung eines Retrievalsystems sein. Die Möglichkeit, mit Hilfe einer Ontologie bessere Suchergebnisse erzielen zu können, muß dem Benutzer entweder systemseitig durch eine Modifikation der Suchverfahren oder aber explizit durch eine überarbeitete Schnittstelle ermöglicht werden.

Eine fachlich hinreichend spezielle Ontologie kann, wenn die Abbildung der vorhandenen Daten auf sie mit der nötigen Genauigkeit gelingt, auch für die Klassifikation neuer Objekte der Domäne verwendet werden (vgl. de Lima et al. (1998) sowie zu OSIRIS den folgenden Abschnitt).

(Semi)Automatische Klassifikation Die wissenschaftliche Bibliothek, in der gut ausgebildete Bibliothekare die neuerworbenen Bücher manuell klassifizieren, stellt eine Ausnahmesituation dar, verglichen mit anderen Datenbeständen wie z.B. Preprints im Internet. An diesen Stellen kommen entweder keine oder nur von den Teilnehmern selbst organisierte Möglichkeiten der Klassifikation in Betracht. Hier wären Verfahren hilfreich, die die Klassifikation von Dokumenten entweder automatisch durchführen oder den Benutzer durch qualitativ hochwertige Klassifikationsvorschläge signifikant entlasten.

OSIRIS unternimmt mit dem *Computer Aided Indexing (CAI)* einen ersten Versuch, aufbauend auf dem in der OSIRIS-Wissensbasis gesammelten Wissen neue Dokumente automatisch zu klassifizieren. Dazu werden die zu einem Dokument vorliegenden Daten wie *Titel*, aber auch die zum Teil vorliegenden Informationen *Fremdnotation* und *Schlagwort* usw. (vgl. 3.1) als Anfrage benutzt: Gegeben ein Werk mit dem Titel „Algebraic Geometry“ und der Angabe der Klasse der Library of Congress Classification, der das Werk beim Erscheinen zugeordnet wurde, so kann OSIRIS Vorschläge generieren, in welche Klasse der lokalen Systematik dieses Buch am ehesten gehört. Diese Vorschläge werden am hochwertigsten aufgrund von Fremdnotationen und Schlagwörtern ausfallen, aber auch eine Titelanalyse kann in vielen wissenschaftlichen Disziplinen die Klassifikation ermöglichen. Dennoch kann es sich bei diesem Verfahren nur um ein semi-automatisches Vorgehen handeln: Der Bibliothekar muß stets die Vorschläge des Systems auf Plausibilität prüfen und ggfs. korrigierend eingreifen. Allerdings ist zu dieser Arbeit mit stark kontrollierendem Charakter nun nicht mehr ein ganz so umfassend ausgebildeter (und daher teurerer) Angestellter nötig. Dieses Verfahren wird seit längerer Zeit in der Universitätsbibliothek Osnabrück für die klassifikatorische Erschließung erprobt – für die Information über die Neuerwerbungen der Bibliothek wird sogar erfolgreich eine vollautomatische Klassifikation der Titel nach Fachgebieten durchgeführt.

Die Klassifikation eines Dokumentes wird umso besser sein, je mehr sprachliches

Material (neben klassifikatorischen Merkmalen) zu einem Dokument vorliegt. Eine Möglichkeit, geeignetes sprachliches Material zu einem Dokument zu erlangen ist das Scannen und Verarbeiten des sogenannten Klappentextes eines Werkes bzw. des Abstracts eines Artikels. Hier ist mit einer überblicksartigen Beschreibung der Inhalte zu rechnen, deren Analyse als Klassifikationshilfsmittel (bei der Abbildung auf ein Klassifikationssystem) denkbar ist. Zur eigentlichen Analyse-Aufgabe hinzu kommen dabei aber die mit OCR derzeit immer noch verbundenen Schwierigkeiten, auf die in 3.4.4 im Zusammenhang des KASCADE-Projektes eingegangen wurde (vgl. Junger (1999)).

Ähnlich geht GERHARD⁷⁵ vor, das in Zusammenarbeit mit dem Institut für Semantische Informationsverarbeitung der Universität Osnabrück entwickelt wurde, und daher inhaltlich einige Parallelen zu OSIRIS aufweist (Wätjen, 1998; Möller et al., 1999). GERHARD besteht aus einer Suchmaschine, deren Suchraum manuell beschränkt ist (auf den akademischen Teil des WWW) und einer über den Ergebnissen der Suchmaschine arbeitenden automatischen Klassifikation. Für die Klassifikation wird der Volltext der WWW-Dokumente ähnlich wie in OSIRIS analysiert, und anschließend versucht, auf das sprachliche Material der UDK⁷⁶ abzubilden. Durch die Auswertung der erfolgreichen Abbildungen sprachlicher Elemente auf Klassen der UDK soll so eine Klassifikation des gesamten Dokumentes möglich werden.

Eine andere Möglichkeit, einen Artikel automatisch in einen fachlichen Zusammenhang zu stellen und damit eine grobe Klassifikation vorzunehmen, stellt die Auswertung der Literaturliste eines Artikels als Kontext des Artikels dar. Diese Technik wird bereits zur Verbesserung von Suchergebnissen in WWW-Suchmaschinen eingesetzt und könnte durch die Betrachtung der umgekehrten Verweis-Richtung (von welchen Dokumenten aus wird auf das aktuelle Dokument verwiesen?) weiter verbessert werden.

Isahara et al. (1998) versucht, für die Domäne der Internet News Groups die Struktur eines vom Benutzer bestimmten Fokus-Artikels zu erkennen und zu den einzelnen Strukturteilen verwandte Artikel zu finden, auf die z.B. inhaltlich (direkt oder indirekt) Bezug genommen wird. Dies geschieht mit dem Ziel, den „thread“ zu erkennen, in den der vom Benutzer genannte Artikel einzuordnen ist. Diese Technik erscheint aber durchaus auch für die Klassifikation des Artikels einsetzbar zu sein.

Stairmand (1997) stellt ein Verfahren zur automatischen Klassifikation von Texten mittels WordNet vor. Zunächst wird der Volltext auf semantisch ähnliche Wörter mit Hilfe der WordNet Synonym Sets (synsets) untersucht. Die Wörter, die zu einem Synset gehören und deren Abstand ein vorgegebenes Abstandsmaß nicht überschreiten, werden als lexikalische Kette betrachtet, mehrere lexikalische Ketten zu einem Synset werden

⁷⁵German Harvest Automated Retrieval and Directory

⁷⁶Universal Dezimal Klassifikation – dreisprachiges, bibliothekarisches Klassifikationssystem. Die UDK verfolgt einen universellen Ansatz, d.h. sie modelliert die ontologischen Gegebenheiten der Welt, und nicht die in einer Bibliothek vorherrschenden lokalen Aufstellungskriterien wie z.B. das Fächerangebot der Hochschule.

als lexikalischer Cluster innerhalb eines Dokumentes angenommen. Dabei wird die Stärke eines solchen Clusters durch die Zahl der beteiligten Ketten und der sie konstituierenden Wörter bestimmt. Die Relevanz eines Dokumentes für eine Frage wird dann durch den Vergleich der in der Anfrage gefundenen Synsets mit denen des Dokumentvolltextes bestimmt.

Ähnliche Objekte Eine Möglichkeit, die Suchverfahren in OSIRIS zu erweitern und auch zu verbessern, könnte eine neue Funktionalität der Suche nach Ähnlichkeit sein. Für verschiedene textuelle Objekte wie Artikel, Projektbeschreibungen, technische Anleitungen sowie medizinische oder juristische Fallbeschreibungen könnte dem Benutzer die Möglichkeit angeboten werden, mit „Ich suche einen Artikel wie ...“ nach ähnlichen Dokumenten zu suchen. Als Ausgangspunkt der Suche eignet sich eine möglichst umfangreiche natürlichsprachliche Beschreibung also z.B. ein bereits als interessant erkanntes Objekt oder eine Klasse in der Systematik. Die Frage nach ähnlichen Objekten kann auch gut als Ausgangspunkt einer Suche im WWW genutzt werden: Wenn der Benutzer zunächst auf einer klassifikatorisch erschlossenen kleinen Domäne wie dem Bibliothekskatalog sucht und dann zu einem als inhaltlich relevant befundenen Buch oder Artikel nach ähnlichen WWW-Seiten fragt, kann der gesamte Volltext (z.B. der Klappentext oder das Abstract des Artikels) für die Suche im WWW verwendet werden. Mit Rechercheverfahren, wie sie z.B. unter dem Namen *IC-find* von der Firma ICDM entwickelt wurden, werden die Suchergebnisse im WWW dann dramatisch besser sein und zwar umso besser, je größer das natürlichsprachliche Ausgangsobjekt für die Ähnlichkeitssuche ist.

Nach Roppel (1993) ist es aber auch für das Einsatzgebiet „Faktenretrieval“ (im Gegensatz zur bislang besprochenen Suche in Texten oder sprachlich annotierten Daten) nötig, „das zugrundeliegende Datenbankmanagementsystem mit seiner Retrievalsprache durch Komponenten, die den Benutzer bei der Formulierung einer adäquaten Anfrage und bei der Interpretation des Ergebnisses unterstützen, zu ergänzen“ (Roppel, 1993, :155). Dabei geht es Roppel neben der für OSIRIS bereits angesprochenen Unterstützung des Suchenden durch z.B. query expansion auch darum, das Problem unscharfen Wissens in der Datenbank und den daraus resultierenden Wunsch der Benutzer nach einer Suche nach ähnlichen Objekten zu lösen. Unscharfes Wissen befindet sich laut Roppel in verschiedenen Varianten in der Datenbank (s. a. S. 198). „Unscharfes Wissen“ liegt vor, wenn zu einem Suchaspekt für ein Objekt keine Messergebnisse vorliegen, aber zu einem ähnlichen Suchaspekt. In einem solchen Fall ist es laut Roppel der einhellige Wunsch der Benutzer, auf „ähnliche“ Suchaspekte auszuweichen. „Unvollständiges Wissen“ findet sich regelmäßig in Faktendatenbanken, da nicht für alle Objekte in allen Stadien ihrer Entwicklung vollständige Werte vorliegen. Bspw. kann ein Objekt in einem Entwicklungsstadium getestet worden sein, aber nicht in der aktuell verwendeten Version seiner Weiterentwicklung. „Unsicheres Wissen“ entsteht, wenn Messwerte in der Datenbank aus unterschiedlichen Quellen stammen, deren Zuverlässigkeit unterschiedlich bewertet wird. „Widersprüchliches Wissen“ schließlich

kann in Sammlungen von Meßwerten auftreten, wenn Messungen eines Kennwertes an einem Objekt aus unterschiedlichen Meßreihen voneinander abweichen.

Adaptivität Um die Retrievalergebnisse besser auf die Bedürfnisse des Suchenden abstimmen zu können, ist es wünschenswert, daß sich das System individuell an den Benutzer anpaßt, sich also adaptiv verhält.

Eine Möglichkeit für ein System, sich adaptiv zu verhalten, besteht darin, explizit die Pläne und Ziele des Benutzers (soweit diese bekannt sind) zu modellieren, und anhand dieses systeminternen Benutzermodells die Interaktion mit dem Suchenden zu steuern. Um die dazu benötigten Informationen zu erlangen, kann das System den Benutzer explizit zu seinen Zielen und Wünschen befragen – dies wird allerdings von den meisten Benutzern abgelehnt, teils, weil sie dies als zu aufwendig und störend empfinden, teils, weil sie sich gegenüber dem System nicht festlegen wollen. Und selbst wenn Benutzer gezwungen sind, ein Programm zu benutzen und diesem Angaben zu ihrem Interessenprofil zu machen, ist das Ergebnis nicht unbedingt hilfreich: Suchende sind sich oft über ihre Bedürfnisse nicht vollständig im Klaren, ihre Bedürfnisse ändern sich im Laufe der Zeit, so daß es nicht reicht, sie ein einziges Mal zu befragen, und Benutzer werden in einer Situation, in der sie gezwungen sind, Angaben zu ihren Interessen zu machen (ohne daß sie begreifen, wie dies zu ihrem Vorteil gereicht), vermutlich qualitativ unzulängliche Angaben machen.

Eine andere, wenig erfolgversprechende Möglichkeit wird von Davenport und Weir (1986) vorgestellt, die eine Reihe möglicher Pläne der Benutzer antizipieren und durch Stichwörter der Benutzereingabe und antizipierte Folgen von Interaktionen den jeweils wahrscheinlichsten Plan bestimmen. Davenport und Weir belassen es aber nicht dabei, einen Plan des Benutzers zu erkennen, sondern beschränken die Interaktionsmöglichkeiten des Benutzers explizit auf die vom aktuellen Plan vorgesehenen Möglichkeiten. Entspricht der vom System erkannte Plan nicht dem vom Benutzer intendierten Plan oder ändert der Benutzer kurzfristig sein Vorhaben, dann zwingt das System ihn zum expliziten Widerruf des aktuellen Planes und zur Auswahl eines neuen Planes aus der Menge der antizipierten Pläne. Dadurch wird die ursprüngliche Absicht, Pläne zur Unterstützung der Benutzerinteraktion und zur Steigerung der Nutzerzufriedenheit einzusetzen, konterkariert, da dem Benutzer bei Davenport und Weir (1986) eine ganze Reihe neuer Zwänge aufgebürdet werden, die ihn in seiner Arbeit mit dem System behindern.

Aus diesen Gründen sind Computerprogramme auf die Beobachtung des Benutzers und den Schluß auf dessen Interessen aufgrund externer Informationen angewiesen. Ein wichtiger Schritt hin zu einer solchen Beobachtung ist die eindeutige Identifikation des Suchenden.

Eine Möglichkeit für ein System wie OSIRIS besteht z.B. in der Koppelung von Systemfunktionalität an die Registrierung des Benutzers, ohne daß dieser gezwungen wäre, mehr als seinen Namen und seine Email-Adresse zu nennen. Der Nutzen für den

Suchenden könnte in der Benachrichtigung über das bevorstehende Ende einer Leihfrist oder der Ergebnisübermittlung bei asynchroner Suche liegen. Der Nutzen für das System liegt dabei in der einfachen Wiedererkennung des Suchenden (ein ansonsten unlösbares Problem) und der Suche nach zusätzlichen Informationen über den Nutzer. Handelt es sich beispielsweise um einen Angehörigen der Universität, so kann über Namen und Email-Adresse dessen Zuordnung zu einem Fachbereich, einem Studiengang, verschiedenen Veranstaltungen usw. eruiert werden, was einen ersten Anhaltspunkt für die Interessen des Suchenden ergibt.

Allerdings ist nicht in allen denkbaren und sinnvollen Einsatzszenarien für ein System wie OSIRIS (und erst recht nicht für andere Systeme) mit der Möglichkeit zu rechnen, daß eine Identifikation des Benutzers zu erreichen ist. Es ist deshalb wichtig, über Möglichkeiten der Benutzermodellierung nachzudenken, die ohne explizite Befragung des Benutzers möglich wird. Dies könnte z.B. durch die Beobachtung der Interaktion des Benutzers mit dem System auf der Grundlage von Standardannahmen über Benutzerverhalten geschehen.

Lieberman (1995) beschreibt Letizia, einen Softwareagenten, der in einen herkömmlichen Browser integriert, das Verhalten des Benutzers bei der Suche im WWW beobachtet und versucht, Schlüsse über das Interessenprofil des Benutzers zu ziehen. Ausgenutzt wird dieses Wissen, um im Hintergrund Links zu identifizieren, die für den Benutzer von Interesse sein könnten. Solche Links werden dem Benutzer in einem separaten Fenster angeboten. Um das Interesse des Benutzers zu erkennen, wird sein Verhalten in Bezug auf einzelne Dokumente und Links innerhalb dieser Dokumente bewertet. Dabei werden das Anlegen eines Bookmarks oder das lokale Sichern eines Dokumentes als Zeichen des Interesses an deren Inhalt gewertet. Ein weiteres Zeichen von Interesse des Benutzers gegenüber einem Dokument ist für Letizia die Tatsache, daß der Benutzer einem Verweis innerhalb des Dokumentes folgt. Auch die Rückkehr zu einem bereits zuvor besuchten Dokument wird als Interesse gewertet. Desinteresse erkennt Letizia an der Auswertung der linearen Abfolge von Verweisen in einem Dokument: Unter der Annahme, daß Verweise top-down und left-right wahrgenommen werden, ist die Anwahl des zweiten Verweises von links in der ersten (sichtbaren) Reihe neben der Interessenbekundung für eben diesen Verweis auch eine Desinteressenbekundung bezüglich des ersten, übergangenen Verweises.

Für eine tiefergehende Analyse des natürlichsprachlichen Inhalts dieser Objekte verfügt Letizia leider nicht über die notwendigen Werkzeuge und muß sich deshalb auf aus dem Dokument extrahierte Stichwörter verlassen. Dazu verwendet Letizia einen Ansatz zum automatischen Indexieren von Dokumenten, indem die für das in Frage stehende Dokument relevanten Stichwörter aufgrund einer einfachen Heuristik (TFIDF)⁷⁷ bestimmt werden (Lieberman, 1996). TFIDF betrachtet die Auftretenshäufigkeit von Termen innerhalb des Dokumentes in Relation zur Auftretenswahrscheinlichkeit im gesamten Korpus. Wörter, die relativ häufig im Dokument, aber relativ selten insgesamt auftreten, werden dabei als besonders für die Beschreibung des Dokumentes ge-

⁷⁷Term Frequency Times Inverse Document Frequency (TFIDF)

eignet identifiziert. Lieberman selbst gibt allerdings zu, daß dieses Vorgehen nicht sehr zuverlässig ist und durch Parsingtechniken deutlich verbessert werden könnte.⁷⁸

Die Anforderungen, die an die Qualität der Systemvorschläge gestellt werden, sind nicht übermäßig hoch: Letizia soll nur auf Anfrage hin Vorschläge unterbreiten, was vom aktuellen Standpunkt des Benutzers im WWW aus gesehen interessant erscheint. Lieberman wörtlich: „Its guesses only need be better than no guess at all“ (Lieberman, 1995, :925). Ob das dem Suchenden für eine langfristige Nutzung des Programms ausreicht, darf bezweifelt werden. Mit einer aufwendigeren Inhaltserschließung erscheint es aber durchaus realistisch, daß Letizia dem Suchenden deutlich mehr bieten kann, als nur einen Vorschlag, der besser ist als „no guess at all“.

Wærn (1997) widmet sich in diesem Zusammenhang der Frage, wie zu erkennen sein könnte, ob ein Benutzer zufrieden ist mit dem, was er als Antwort auf seine Suchanfrage geliefert bekommen hat. Dabei stellt sie fest, daß in *direct manipulation interfaces* die Benutzer das System als Werkzeug behandeln, um ihre Ziele zu erreichen. Im Gegensatz dazu versuchen Benutzer z.B. an NL-Schnittstellen, gemeinsam mit dem System eine Aufgabe zu lösen.⁷⁹ Diese Unterschiede haben Auswirkungen auf die Erkennung von Benutzerplänen, da in der nicht kooperativen Situation (*key hole plan recognition*) an Schnittstellen mit direkter Manipulation nur auf sehr niedriger Ebene Informationen über Benutzeraktionen vorliegen und zudem die Benutzer häufig ohne Vorwarnung ihre Suchziele ändern, ohne das System darüber zu unterrichten (da sie diesem ja nur den Status eines Werkzeuges zubilligen). Wærn (1997) bemerkt, daß der plötzliche Wechsel der Suchziele insbesondere an Information Retrieval Systemen anzutreffen sei, da hier die vom System gefundenen Ergebnisse unmittelbar die Intentionen des Suchenden beeinflussen und zu neuen, geänderten Anfragen führen würden. Wærn übersieht dabei allerdings, daß die Änderungen der Suchintentionen nicht willkürlich sind, sondern von den Systemmeldungen ausgelöst werden. Da diese aber dem System bekannt sind, ist die Lage für das Information Retrieval vielleicht nicht ganz so negativ, wie von Wærn dargestellt.

In einer empirischen Untersuchung versuchte Wærn herauszufinden, ob *key hole plan recognition* beim Lesen von Nachrichten ausreicht, um zu entscheiden, welche Artikel für den Betrachter besonders interessant oder gerade uninteressant waren. Dabei stellte Wærn fest, daß es zwar relativ einfach sei, eine Reihe abstrakter Lesestrategien zu unterscheiden, daß der Erfolg dieser Strategien (ob der Suchende interessante Information gefunden hat) aber nicht oder nur sehr schwer zu erkennen gewesen sei. Insbesondere habe sich herausgestellt, daß die Benutzer regelmäßig ihre Suchstrategien geändert hätten. Aus dieser Untersuchung schließt Wærn, daß die Daten, die aufgrund von *key hole plan recognition* gewonnen werden, von zu geringer Qualität sind, um als Grundlage für einen Filtermechanismus für das Lesen von Nachrichten dienen zu

⁷⁸Vrgl. dazu in 3.4.4 auch die vernichtende Kritik von Junger (1999) an den Ergebnissen des KASCADE-Teilprojektes SELIX-JB, das dieselbe Heuristik einsetzt.

⁷⁹Wærn (1997) geht nicht darauf ein, warum dies für NL-Schnittstellen so ist – vielleicht ist sie der Ansicht, daß NL-Schnittstellen aus Benutzerperspektive keinen reinen Werkzeugcharakter haben.

können.

An Schnittstellen, an denen der Benutzer kooperativ mit dem System eine Aufgabe bearbeitet wird für die Planerkennung auch von *intended plan recognition* gesprochen. Hier kann mit Hilfe von Parsing-Ansätzen versucht werden (vgl. z.B. VERBMOBIL, S. 109), eine Sequenz von Benutzeraktionen mit einem der in der Plangrammatik verzeichneten Benutzerpläne zu identifizieren. Dabei kann mit Hilfe von Meta-Plänen (wenn auch sehr aufwendig) der Strategiewechsel von Benutzern erkannt werden – jedoch ist, so Wærn, im Fall der *key hole plan recognition* die Datenlage zu schlecht, als daß sinnvoll Meta-Pläne eingesetzt werden könnten. Weiterhin berichtet Wærn, daß es eine auffällige Diskrepanz gegeben habe zwischen der hohen Komplexität des Planerkennungsmechanismus einerseits und den tatsächlich erkannten Strategien andererseits. Letztere seien allesamt sehr rudimentär gewesen und nur aufgrund von einer oder maximal zwei charakteristischen Benutzeraktionen zu erkennen gewesen.

Um den genannten Schwierigkeiten zu begegnen, schlägt Wærn vor, die Planerkennung auf lokale Pläne zu beschränken – dies sei auf Grundlage der Daten möglich und verlange keine Metapläne zum Strategiewechsel, sondern lediglich einen Anhaltspunkt, wie häufig ein solcher Wechsel vorkomme. Die Erkennung lokaler Pläne wurde in zwei Beispielanwendungen (News Filtering und adaptiver Hypertext) mit unterschiedlichen Konzepten für die Beschränkung der Pläne erfolgreich erprobt. Zum einen wurde dafür der Fokus der Planerkennung auf eine feste Anzahl an Benutzeraktionen beschränkt. Zum anderen wurden den einzelnen Benutzeraktionen mit steigendem Abstand geringere Gewichte zugewiesen.

Dabei erwies sich der zweite Ansatz als deutlich flexibler und erfolgversprechender für die weitere Entwicklung, als die mit festen, vorgegebenen Grenzen arbeitende erste Beschränkung.

6 Die Wirkung der OSIRIS-Schnittstelle auf den Benutzer

In den vorangegangenen Kapiteln wurden einige Aspekte herausgearbeitet, die für den „Erfolg“ eines Dialogsystems hinsichtlich der Zufriedenheit der Benutzer als wichtig angesehen wurden:

- Oberflächengestaltung / ergonomische Adäquatheit
- Eingabemöglichkeiten / Interaktionsformen
- Für Domäne und Medium optimierte intelligente Suchverfahren
- Ergebnispräsentation / Möglichkeiten der Weiterverwendung der Ausgabe
- Präsuppositionen bei Bezeichnungen und Bedienung (wieviel Fach- oder Domänenwissen wird benötigt?)
- Flexibilität / Robustheit gegenüber Eingabefehlern
- Bedienbarkeit ohne Training und ohne Hilfetexte (wenn überhaupt möglich)

Kriterien dieser Art helfen den Entwicklern bei der Konzeption und Erstellung eines Dialogsystems, lassen die konkrete Umsetzung aber offen. Um herauszufinden, wie die von den Entwicklern gefundenen Lösungen sich auf die Nutzerzufriedenheit auswirken, müssen die fertigen Systeme evaluiert werden. Aufbauend auf der Interpretation der Ergebnisse einer Evaluation können dann konkrete Verbesserungen des Systems vorgenommen werden.

In 6.1 wird die Evaluation von OSIRIS im Vergleich mit dem konventionellen OPAC beschrieben. Eines der interpretationsbedürftigsten Ergebnisse ist hier die Zufriedenheit der Nutzer mit der Qualität der Rechercheergebnisse: Es zeigt sich, daß die in 3.3.1 dargestellte deutliche Überlegenheit von OSIRIS (im Schnitt 10 mal höherer Recall bei konstant hoher Precision) keine Entsprechung in der Zufriedenheit der Benutzer findet, da OSIRIS und der OPAC hinsichtlich der subjektiven Beurteilung der Qualität der Rechercheergebnisse gleich abschneiden. Um dieses Faktum erklären zu können, werden in 6.2 die Erkenntnisse zur Akzeptanz von NL-Schnittstellen aus der Literatur zusammengetragen.

6.1 Psychologische Untersuchungsergebnisse zu OSIRIS

Die Evaluation von OSIRIS wurde unter der Leitung von Dr. Kai-Christoph Hamborg aus der Arbeitsgruppe „Arbeits- und Organisationspsychologie“ des Fachbereiches Psychologie der Universität Osnabrück durchgeführt. Sein Interesse lag zunächst im Vergleich verschiedener Evaluationsmethoden, deren Eigenschaften er am Untersuchungsgegenstand OSIRIS erprobte (Hamborg, 1997; Hamborg et al., 1998; Hamborg,

1998). Darüber hinaus regte er in einem von ihm geleiteten Studienprojekt zum Thema „Evaluation von Dialogsystemen“ die teilnehmenden Studenten an, auch OSIRIS und den OPAC als Untersuchungsgegenstände mit einzubeziehen (Moranz, 1998). Und schließlich betreute er die Arbeit einer Teilnehmerin des Studienprojektes, die mit einer Benutzbarkeitsstudie zum Vergleich von OSIRIS und dem OPAC ihr Diplom in Psychologie erwarb (Moranz, 2000).

Die Evaluation von OSIRIS begann 1997 mit der Fertigstellung der ersten (hochschul)öffentlichen Version 1.0. Da klar war, daß OSIRIS noch zahlreiche Veränderungen und Erweiterungen erleben würde, wurden zur Bewertung der Gestaltung von OSIRIS *formative* Verfahren eingesetzt, also solche, die während des Entwicklungsprozesses einsetzbar sind und deren Rückmeldungen den Entwicklungsprozeß begleiten und einen Zyklus aus Verbesserung und erneuter Bewertung anstoßen sollen. Zum Zweck des Vergleichs unterschiedlicher Evaluationsmethoden wurde sowohl eine expertenzentrierte, heuristische Evaluation (Hamborg, 1997) als auch eine nutzerorientierte Evaluation mit Hilfe des in Osnabrück entwickelten IsoMetrics-Verfahrens (Willumeit et al., 1995, 1996; Gediga und Hamborg, 1997) durchgeführt.

Die heuristische Evaluation wird im frühen Designstadium eingesetzt – eine kleine Gruppe von Experten soll dabei auf der Grundlage einer geringen Anzahl relativ allgemeiner Richtlinien mögliche Benutzbarkeitsprobleme aufdecken. Für OSIRIS wurden fünf Experten aus dem Bereich Software-Ergonomie und Softwaredesign im Ergonomie-Labor des Fachbereiches Psychologie gebeten, anhand der nachfolgend wiedergegebenen Leitfragen (Heuristiken) OSIRIS zu bewerten (Hamborg, 1997):

1. Einfacher und natürlicher Dialog.
Anforderung: Informationen werden in natürlicher und logischer Anordnung dargeboten. Es werden alle relevanten Informationen angeboten, das sind die Informationen, die der Benutzer aktuell benötigt.
2. Spreche die Sprache des Benutzers.
Anforderung: Jede Information wird in solchen Worten oder Konzepten dargestellt, die der Benutzer verstehen kann. Es wird kein „Fachchinesisch“ gesprochen.
3. Minimiere die Gedächtnisbelastung des Benutzers.
Anforderung: Relevante Information steht im Dialog immer zur Verfügung. Benutzerhilfen stehen schnell und einfach zur Verfügung.
4. Sei konsistent.
Anforderung: Die gleichen Aktionen, Wörter oder Situationen sollen als gleiche erkennbar sein.
5. Gib Rückmeldung.
Anforderung: Der Benutzer wird über die aktuellen Systemzustände und erfolgte Systemreaktionen ausreichend informiert. Die Rückmeldung paßt zeitlich.

6. Sorge für schnelle Abbruchpunkte.
Anforderung: Der Benutzer ist in der Lage, jede Dialogeinheit schnell wieder abzubrechen.
7. Unterstütze Abkürzungen.
Anforderung: Erfahrene Benutzer sollen die Möglichkeit haben, den Dialog durch Kurzkommandos oder Makros abzukürzen.
8. Gute, präzise Fehlermeldungen.
Anforderung: Gute Fehlermeldungen sind defensiv, präzise und konstruktiv. Der Benutzer wird nicht kritisiert. Es wird genau über den Grund des Problems informiert, und es wird Hilfe für die Fehlerbehebung gegeben.
9. Verhüte Fehler.
Anforderung: Sorgfältiges Design der Software minimiert die potentiellen Fehlersituationen.
10. Graphische Gestaltung.
11. Sonstiges.

Insgesamt wurden von den Experten über 70 Anmerkungen zu den einzelnen Leitfragen mit Bezug auf OSIRIS gemacht. Solche Anmerkungen betrafen z.B. die nicht immer eindeutige Verwendung von Icons und Bildschirmbezeichnungen, die manchmal unübersichtliche Präsentation der Suchergebnisse sowie Mißverständnisse aufgrund des neuen, weil natürlichsprachlichen Anfragemodus. Diese Punkte wurden intensiv im OSIRIS-Projekt diskutiert und unterschiedlich schnell und unterschiedlich weitgehend geändert.

Während die Änderung von textuellen Bezeichnungen und graphischen Elementen sehr schnell umgesetzt werden konnte, dauerte es hingegen sehr lange, bis OSIRIS auch um einen Anfragemodus erweitert wurde, in dem mit Booleschen Operatoren gearbeitet werden kann. Diese Verzögerung erklärt sich durch die in 2.2 zur Genüge angeführten Belege, daß Benutzer mit Booleschen Operatoren eigentlich nicht umgehen können. Zusammen mit dem deutlichen Eindruck, daß die wiederholten Forderungen von Benutzern nach einem solchen Interaktionsmodus eigentlich auf einem Mißverständnis über die natürlichsprachlichen Anfragemöglichkeiten von OSIRIS beruhen, entstand die erst spät aufgegebene Überzeugung, daß es nur eine Frage der Zeit und der Erfahrung sei, bis Nutzer von OSIRIS die Vorteile der natürlichsprachlichen Anfrage erkennen würden. Inzwischen hat auch das OSIRIS-Projekt zur Kenntnis genommen, daß es letztlich der Wunsch des Benutzers ist, der zählt, und der umgesetzt werden sollte. Wir glauben aber ebenfalls, daß es immer wieder neue Benutzer sind, die die Benutzung Boolescher Operatoren aufgrund von Vorerfahrungen mit anderen Programmen vermissen, und daß diejenigen, die länger mit OSIRIS gearbeitet haben,

Boolesche Operatoren nur noch vereinzelt und für ganz spezielle Aufgaben einsetzen, z.B. wenn sie Bibliotheksbeauftragte ihrer Organisationseinheit sind und in dieser Rolle spezifisch bibliothekarische Aufgaben erfüllen müssen.

Das zweite eingesetzte formative Evaluationsverfahren war IsoMetrics^L.⁸⁰ IsoMetrics^L orientiert sich an dem speziell auf die Gestaltung und Bewertung von Dialogsystemen ausgerichteten Teil 10 der ISO-Norm 9241, in dem sieben Gestaltungsgrundsätze formuliert sind (Willumeit et al., 1996):

1. Aufgabenangemessenheit
2. Selbstbeschreibungsfähigkeit
3. Steuerbarkeit
4. Erwartungskonformität
5. Fehlerrobustheit
6. Individualisierbarkeit
7. Erlernbarkeit

Für jeden dieser Gestaltungsgrundsätze existiert in IsoMetrics^L eine Subskala mit mehreren Einträgen, so daß insgesamt 90 Eigenschaften des Systems abgefragt werden. Zu jeder Eigenschaft, d.h. zu jedem Eintrag auf einer Subskala, werden in IsoMetrics^L drei unterschiedliche Daten erhoben:

1. Für die zu bewertende Software werden auf einer Skala von „stimme überwiegend nicht zu“ über „teils/teils“ bis hin zu „stimme überwiegend zu“ die einzelnen Eigenschaften abgefragt. Beispiel: „Wenn Menü-Optionen in bestimmten Bearbeitungsschritten nicht zur Verfügung stehen, wird mir die Sperrung dieser Option z.B. durch Farbigkeit sichtbar gemacht.“
2. Für jede Eigenschaft wird auf einer Skala von „nicht wichtig“ über „teils/teils“ bis hin zu „wichtig“ deren Bedeutung für den Gesamteindruck der zu bewertenden Software abgefragt. Beispiel: „Wie wichtig ist dieser Aspekt für den Gesamteindruck der Software?“.
3. Für jede Eigenschaft werden (wenn möglich) ein oder mehrere Beispiele in Freitext erbeten, bei denen die in Frage stehende Eigenschaft nicht zutrifft. Beispiel: „Können Sie konkrete Beispiele nennen, bei denen Sie dieser Aussage nicht zustimmen können?“.

⁸⁰Das „L“ in IsoMetrics^L steht für *long* und unterscheidet die formative Variante des Evaluationsverfahrens von der summativen Variante IsoMetrics^S, wobei in IsoMetrics^S das „S“ für *short* steht.

Mit Hilfe von IsoMetrics^L bewerteten fünfzehn Testpersonen OSIRIS im Ergonomie-Labor des Fachbereiches Psychologie. Nach einer etwa 30minütigen Explorationsphase mußten die Versuchspersonen vier Testaufgaben bearbeiten und im Anschluß das System mit dem IsoMetrics^L-Fragebogen bewerten.

Auf diese Weise entstanden 168 Anmerkungen zu den Eigenschaften des Systems. Neben einleuchtenden und hilfreichen Anmerkungen zu fehlenden oder unzureichenden Hilfeseiten sowie der inkonsistenten Nutzung von Symbolen oder Bezeichnungen wurden auch vielfach unerfüllbare oder nicht in den Verantwortungsbereich von OSIRIS gehörige Wünsche geäußert sowie allgemeine Merkmale der Browser- und Betriebssystembedienung OSIRIS angekreidet. Zu den im Rahmen von OSIRIS unerfüllbaren Wünschen gehört z.B. eine ausführliche Inhaltsangabe zu jedem Titel (siehe aber 5.2.3 zur Idee, Klappentexte und Abstracts zu scannen und zu verwerten). Ebenfalls „unerfüllbar“ waren zahlreiche Verbesserungsvorschläge, die die Funktionalität des Betriebssystems (Verständlichkeit von Windows-Fehlermeldungen) oder die Funktionalität des Browsers („Zurück“-Taste) betrafen: Viele Nutzer konnten ganz offensichtlich nicht zwischen OSIRIS und der Programmumgebung wie dem Browser oder dem Betriebssystem trennen. Auch zeigte sich, daß elementare Kenntnisse im Umgang mit Internet-Programmen nicht vorhanden waren.

Die Ergebnisse der softwareergonomischen Untersuchungen der Version 1.0 und 1.1 von OSIRIS wurden im Rahmen eines Review in den Entwicklungsprozeß aufgenommen und zusammen mit anderen, systeminternen Veränderungen in der Version 2.0 umgesetzt. Um den Effekt der Weiterentwicklung zu testen und auch, um einen Vergleich zwischen der natürlichsprachlichen Eingabeschnittstelle von OSIRIS und dem herkömmlichen OPAC zu erreichen, wurden in einer Feldstudie 120 Personen zu ihrer Arbeit mit den Programmen befragt und ihre Bedienungszeiten erfaßt (Hamborg et al., 1998). Dazu wurden die Personen an der Tür zur Bibliothek angesprochen – in einem vorbereiteten Raum wurden mit Hilfe eines Fragebogens Angaben zur Person sowie zum Rechercheziel und zur erwarteten Dauer der Recherche erfragt. Zum größten Teil wollten die Befragten Themensuchen vornehmen, an zweiter Stelle stand die Suche nach einem bestimmten Buch, Titel oder Autoren, gefolgt von komplexeren Suchabsichten, die sowohl Themensuche, als auch die konkrete Suche nach einem bestimmten Buch, Titel oder Autor umfaßten. Nur sehr selten bestand die Absicht, nach Zeitschriften zu suchen.

Im Anschluß an die Befragung konnten die Testpersonen das ihnen eventuell noch nicht bekannte Rechercheprogramm ausprobieren, bevor sie gebeten wurden, die Recherche, derentwegen sie in die Bibliothek gekommen waren, durchzuführen. Es wurden vier Systeme eingesetzt, OSIRIS in den Versionen 1.1 und 2.0, der konventionelle OPAC und die HTML-Version des OPAC. Mit jeder Programmversion arbeiteten 30 Personen, ohne daß ihnen ein Zeitlimit für ihre Recherche gesetzt wurde. Dabei gab es keine unterschiedliche Verteilung der Rechercheziele über die untersuchten Systeme. Im Anschluß an die Recherche wurden die Personen gebeten, die Rechercheergebnisse zu bewerten, Angaben zu ihrer Vorerfahrung mit Computern zu machen, das System

nach Akzeptanzgesichtspunkten zu bewerten und Kommentare zum Programm abzugeben.

Für den Punkt *Bearbeitungszeit* ergab sich ein bedeutsamer Unterschied zwischen OSIRIS 2.0 und OSIRIS 1.1: Die Bearbeitungszeiten in der Version 2.0 sind kürzer als die in der Version 1.1. Kein bedeutsamer Effekt hingegen zeigte der Vergleich zwischen dem konventionellen OPAC und OSIRIS 1.1.

Zur Bewertung der *Rechercheergebnisse* wurden den Testpersonen fünf Fragen gestellt, denen auf einer fünfstufigen Skala zugestimmt oder widersprochen werden sollte. Es wurde gefragt, ob alle Informationen, nach denen gesucht wurde, auch gefunden wurden, ob die gefundenen Informationen inhaltlich den Erwartungen entsprachen und ob die Informationen genau die Informationen enthielten, nach denen gesucht wurde. Außerdem wurde die Menge der gefundenen Informationen und die benötigte Recherchezeit bewertet.

Im Ergebnis wird OSIRIS 1.1 durchgängig schlechter als der herkömmliche OPAC bewertet, während sich für die anderen Systeme keine statistisch bedeutsamen Abweichungen zeigen.

Für die Bewertung der Akzeptanz der untersuchten Programme wurde der „Questionnaire for User Interface Satisfaction“ (QUIS) verwendet. QUIS umfaßt fünf neunstufige Skalen und enthält die Möglichkeit, freie Kommentare vorzunehmen. Die fünf Skalen sind

1. Gesamtbewertung
2. Bildschirm
3. Fachwörter und Systeminformationen
4. Erlernen
5. Fähigkeiten des Systems

OSIRIS 1.1 und OSIRIS 2.0 werden auf der Skala „Bildschirm“ hinsichtlich Gestaltung und Aufbau des Bildschirms, Lesbarkeit von Zeichen, der Qualität von Hervorhebungen und zur Abfolge von Bildschirmseiten bedeutsam besser als der herkömmliche OPAC bewertet. Ebenfalls bedeutsam besser als der herkömmliche OPAC wird OSIRIS 1.1 in Bezug auf die Erlernbarkeit bewertet. OSIRIS 2.0 wird tendenziell besser als der herkömmliche OPAC in Bezug auf die Erlernbarkeit beurteilt.

In den Untersuchungen, die im Rahmen des Studienprojektes „Evaluation von Dialogsystemen“ (Moranz, 1998) am OPAC und an OSIRIS durchgeführt wurden, stand die Frage im Vordergrund, ob die Interaktion mit natürlicher Sprache wie in OSIRIS oder einer strukturierten Abfragesprache wie im OPAC die Handlungsstrategien der Nutzer beeinflusst oder ihren Suchabsichten entgegenkommt. Zu untersuchen war daher,

ob die Nutzer in Abhängigkeit vom verwendeten System und den von ihnen wahrgenommenen Merkmalen des Systems andere Suchverfahren wählen. Dazu wurde ein Rechercheziel als Aufgabe formuliert und für 20 Versuchspersonen untersucht, ob es in Hinblick auf die Strategiebildung der Nutzer einen Unterschied zwischen der natürlichsprachlichen Schnittstelle von OSIRIS⁸¹ und der Schnittstelle des OPAC gibt.

Die Suchanfragen wurden u.a. mit dem Ziel formuliert, im Gegensatz zu früheren Untersuchungen von Bibliothekssystemen realitätsnah zu sein. So sollte sichergestellt werden, daß der Nutzen für den Alltag des Bibliotheksbenutzers und nicht für die spezifischen Belange⁸² der Bibliothekare überprüft wird. Es wurde dazu eine Vorbefragung von 24 Bibliotheksbenutzern durchgeführt, in deren Rahmen mit einem zu diesem Zweck entworfenen Fragebogen nach dem Rechercheziel und den Erwartungen der Suchenden an das Ergebnis gefragt wurde. 70,8 %⁸³ wollten nach Literatur zu einem bestimmten Thema suchen, 41,7 % der Befragten suchten nach einem bestimmten Buch und 16,7 % eine Zeitschrift. Daraus ergab sich, daß drei Aufgaben, eine thematische Recherche, eine Suche nach einem bestimmten Werk („Autor-Titel-Suche“) und eine nach einer bestimmten Zeitschrift, formuliert und am OPAC sowie an OSIRIS getestet wurden. Trotzdem stellte es sich später heraus, daß die Suchaufgabe in Bezug auf die Zeitschrift wegen mißverständlicher Formulierung⁸⁴ teilweise nicht in die Auswertung übernommen werden konnte.

Um den gesamten Prozeß der Suche betrachten zu können, wurden die 20 Versuchspersonen nach ihren Zielen, Strategien und Problemen im Umgang mit Computern befragt. Ausgewertet wurden die Strategien bei der Ersteingabe, die Nutzungshäufigkeit der Strategien im weiteren Suchverlauf, die Anzahl der Strategiewechsel, Art und Anzahl der Reformulierungen sowie je Aufgabe die benötigte Zeit und der Sucherfolg. Um Aufschluß über die Ziele und Strategien der Versuchspersonen zu bekommen, wurde nach Wegen gesucht, durch Mitteilungen der Versuchspersonen über ihre Wahrnehmungen, Gedanken und Gefühle in der Interaktion mit dem Computer Zugang zu den an der Suche beteiligten kognitiven Prozessen zu erhalten. Dazu wurde eine Kombination⁸⁵ aus Verhaltensbeobachtung durch Videoaufzeichnung und retrospektivem lauten Denken gewählt. Konkret wurden je zehn Versuchspersonen gebeten, mit dem OPAC bzw. mit OSIRIS die vorbereiteten Aufgaben zu bearbeiten. Nach einer kurzen Erklärung des Experimentes und dem Hinweis, daß es um den Vergleich zweier Bibliothekssysteme in Hinblick auf deren „Benutzbarkeit“ gehe, wurden ggfs. die Besonderheiten von OSIRIS kurz erläutert (natürlichsprachliche Eingabeschnittstelle, Möglichkeit der Themensuche), da in Bezug auf den OPAC alle Versuchspersonen ausreichende Vorkenntnisse hatten.

⁸¹Untersucht wurde OSIRIS in der Version 1.1 und 2.0

⁸²Z.T. auch: die theoretischen Interessen der Bibliothekare.

⁸³Mehrfachnennungen waren erlaubt.

⁸⁴Die Probanden suchten einen Zeitschriftenartikel statt einer Zeitschrift, Artikel sind jedoch gar nicht im Datenbestand enthalten.

⁸⁵Heterarchische Aufgabenanalyse (HAA), vgl. Hamborg und Greif (1998).

Anschließend wurde mit einem Fragebogen die Computervorerfahrung der Versuchspersonen überprüft und eine 15minütige Explorationsphase mit dem entsprechenden System angesetzt. Danach wurden von den Probanden drei Suchaufgaben ohne Zeitvorgabe bearbeitet und im Anschluß nach einer fünfminütigen Pause ein Interview mit Videokonfrontation durchgeführt.

Auf dem Video waren mit Hilfe von drei Kameras die Tastatur mit den Händen der Versuchsperson, ihr Gesicht sowie die Suchmaske des Systems, wie sie auf dem Bildschirm der Versuchsperson gezeigt wurde, aufgezeichnet. In zwei gemeinsamen Durchläufen wurden zunächst die Aktionen der Versuchsperson protokolliert und durch die Versuchsperson subjektiv in größere Einheiten („Schritte“) unterteilt. Im zweiten Durchgang wurde die Versuchsperson in kurzen Abständen immer wieder nach den erhofften Ergebnissen der gerade aktuell zu sehenden Handlung, eventuellen Schwierigkeiten beim Erreichen des Ergebnisses sowie einer eventuellen Einbettung der aktuellen Handlung in einen Plan befragt, um Ziele, Strategien und Probleme der gesamten Suche zu ermitteln. Auf Grundlage dieser Daten aber auch anderer, unangeforderter Hinweise der Probanden wurden dann später Art der Strategie und die Strategiewechsel ermittelt.

In der beschriebenen Untersuchung ergab sich hinsichtlich der verwendeten Strategien ein signifikanter Unterschied für die „Autor-Titel-Suche“ bei der Ersteingabe. Alle Benutzer des OPAC wenden die Strategie „Grober Filter“ an, d.h. sie übernehmen nicht die gesamte Information der Aufgabe in die Suchanfrage, sondern selektieren bereits für die erste Anfrage an das System. Dies ist nicht unbedingt verwunderlich, da die direkte Übernahme natürlichsprachlich formulierter Anfragen in den OPAC praktisch unmöglich ist, was den Probanden aufgrund ihrer Vorerfahrung vermutlich bekannt war. Die Übernahme größerer Teile der Suchformulierung in den OPAC scheiterte sicher aber auch daran, daß nachweislich weniger als 10 % der Benutzer klar ist, daß in der Anfrage eine Verknüpfung der Suchaspekte „Autor“ und „Titel“ möglich ist (vgl. 2.2).

Demgegenüber übernahm die Hälfte der Versuchspersonen, die OSIRIS benutzte, die Suchformulierung direkt aus der Aufgabe. Dies hängt sicherlich auch damit zusammen, daß in der entsprechenden Maske ein Beispiel für eine solche Anfrage explizit gegeben wird.

Für die Themensuche ergab sich kein signifikanter Unterschied zwischen OPAC und OSIRIS bei der Ersteingabe. Die Themensuche in OSIRIS ist sehr frei gestaltet, d.h. es gibt eine Beispielfrase, aber es fehlen natürlich für die Benutzer möglicherweise hilfreiche Leitlinien zur Eingabe (Anzahl der Wörter usw.), so daß die Probanden eventuell auf ihre Erfahrungen am OPAC zurückgegriffen haben. Und da alle Probanden über Vorerfahrung mit dem OPAC verfügten, waren sie es auch gewohnt, ihr Suchinteresse nur selektiv und bereits interpretiert dem System mitzuteilen.

Für den weiteren Suchverlauf wurden keine signifikanten Unterschiede zwischen dem OPAC und OSIRIS hinsichtlich der verwendeten Strategien festgestellt. Zwar wurden von den mit OSIRIS arbeitenden Probanden signifikant häufiger Zusatzinforma-

tionen über einen Titel eingeholt und auch signifikant häufiger das Suchergebnis durch scrollen des Bildschirms betrachtet, insgesamt ergab sich aber über alle betrachteten Strategien kein signifikanter globaler Unterschied zum OPAC.

Im Gegensatz zur „Autor-Titel-Suche“ wurde für die Themensuche aber ein signifikant höherer Wechsel der Strategien durch die OSIRIS-Nutzer festgestellt. D.h. daß die Versuchspersonen, die OSIRIS verwendeten, eine größere Anzahl unterschiedlicher Strategien während der Suche einsetzten und auch interindividuell sich stärker voneinander im Suchverhalten unterschieden, als dies bei den OPAC-Nutzern der Fall war, die eine einzige (vermutlich durch Vorerfahrung bekannte) Erfolgsstrategie verwendeten.

Hinsichtlich des Sucherfolges ergaben sich keine signifikanten Unterschiede zwischen dem OPAC und OSIRIS. Auch der zeitliche Aufwand der Probanden unterschied sich nicht für die beiden Systeme: Sowohl die Zeit zwischen dem Beginn der Suche und dem ersten relevanten Treffer als auch die insgesamt für die Suche benötigte Zeit zeigte für die beiden Systeme keinen signifikanten Unterschied.

Zusammenfassend scheint OSIRIS den Benutzern offenbar mehr Raum zu geben, unterschiedliche und persönlich bevorzugte Herangehensweisen umzusetzen (vgl. auch Moranz et al. (1999)). Einer negativen Interpretation der häufigeren Strategiewechsel im Sinne einer Konfusion der Benutzer steht die Tatsache entgegen, daß trotz des reichhaltigeren Spektrums an Strategien und der größeren interindividuellen Unterschiede der Sucherfolg und die dafür benötigte Zeit verglichen mit dem OPAC nicht anstiegen. Als wesentliche Schwäche der Untersuchung im Rahmen des Studienprojektes (Moranz, 1998) wurde von allen Beteiligten die durchgehende Vorerfahrung der Versuchspersonen in Hinblick auf den OPAC betrachtet. Aus diesem Grunde wurden im Anschluß an das Studienprojekt im Rahmen einer Diplomarbeit (Moranz, 2000) die beschriebenen Untersuchungen mit 18 neuen Versuchspersonen wiederholt, die in Hinblick auf beide Systeme als unerfahren einzustufen sind.

Sowohl für die Ersteingabe wie für den weiteren Suchverlauf wurde bei der Suche nach einem bestimmten Werk („Autor-Titel-Suche“) die Strategie „Grober Filter“ von den OPAC-Nutzern häufiger als von den OSIRIS-Nutzern angewandt – allerdings waren die Unterschiede nicht signifikant.

Vor dem Hintergrund der Ergebnisse der ersten Untersuchung und unter Zuhilfenahme der Kommentare der Nutzer zu den von ihnen durchgeführten Aktionen kommt Moranz (2000) zu dem Schluß, daß an dieser Stelle offenbar Merkmale der verwendeten Systeme ausschlaggebend sind für das beobachtete Verhalten.

Für die thematische Recherche ergab sich jedoch sowohl bei der Ersteingabe wie auch im weiteren Suchverlauf ein signifikanter Unterschied zum OPAC: Bei der Ersteingabe übernahmen die mit OSIRIS arbeitenden Probanden signifikant häufiger die gesamte Suchformulierung als Systemeingabe. Und im weiteren Suchverlauf wurde die Möglichkeit, mittels der durch OSIRIS gefundenen Klassen das Fachgebiet und die angrenzenden Bereiche durch Browsing zu überblicken, so häufig in Anspruch genommen, daß sich ein globaler Unterschied zwischen OSIRIS und dem OPAC ergibt.

Dieses Ergebnis steht im Widerspruch zu dem Ergebnis der ersten Untersuchung, weshalb Moranz folgert, daß es sich nicht um ein systemspezifisches Merkmal, sondern um eine Folge der Vorerfahrung der Probanden der ersten Untersuchung mit dem OPAC handelt. Es sei „anzunehmen, daß Probanden, die durchschnittlich vorerfahren im Umgang mit OPAC sind, die vermeintlichen Restriktionen dieses Systems auch auf OSIRIS übertragen, während gänzlich unerfahrene Nutzer da möglicherweise eher unvoreingenommen agieren“ (Moranz, 2000, :85).

Die Anzahl der Strategiewechsel unterschied sich weder für die „Autor-Titel-Suche“ noch für die thematische Recherche signifikant zwischen dem OPAC und OSIRIS. Insbesondere zeigt die zweite Untersuchung – im Gegensatz zur ersten Untersuchung, wo dieses Phänomen noch charakteristisch für die OSIRIS-Nutzer war – für alle Probanden eine große Anzahl unterschiedlicher Strategien sowie interindividuell große Unterschiede.

Moranz schließt daraus, daß es sich weniger um ein systemspezifisches Phänomen, als vielmehr wiederum um eine Folge der Vorerfahrung der Probanden der ersten Untersuchung mit dem OPAC handelt.

Allerdings zeigten sich signifikante Unterschiede für die thematische Recherche beim Sucherfolg und der dafür benötigten Zeit. Während die „Autor-Titel-Suche“ von allen Versuchspersonen, unabhängig vom verwendeten System und ohne signifikanten Unterschied in der benötigten Zeit, erfolgreich durchgeführt wurde, fanden die OSIRIS-Benutzer bei der thematischen Recherche im Schnitt 8,5 Bücher, während es bei den OPAC-Nutzern nur 6,6 Bücher waren. Zudem brauchten die OSIRIS-Benutzer weniger als die Hälfte der von den OPAC-Nutzern benötigten Zeit, um den ersten relevanten Treffer zu erzielen. Die zeitliche Überlegenheit von OSIRIS spiegelt sich auch in der Zeitschriftensuche wieder: OSIRIS-Nutzer fanden die gesuchte Zeitschrift in lediglich 25 % der Zeit, die die OPAC-Nutzer benötigten.

Nach Ansicht von Moranz zeigen diese Ergebnisse im Gegensatz zu zahlreichen, von Moranz (2000) selbst zitierten Evaluationsstudien (vgl. auch 6.2), daß ein mit einer natürlichsprachlichen Schnittstelle ausgerüstetes System hinsichtlich der Effektivität konventionellen Systemen durchaus überlegen sein kann. Allerdings schränkt Moranz diese Aussage zu Recht ein, wenn sie auf den ungeklärten Beitrag der qualitativ sehr viel besseren Ergebnispräsentation von OSIRIS zur Überlegenheit hinsichtlich der Effektivität verweist.

6.1.1 Fazit

Die Ergebnisse der software-ergonomischen Evaluation (heuristische Evaluation und IsoMetrics^L-Evaluation) wurden in den Entwicklungsprozeß integriert und stellten ein enorm hilfreiches Korrektiv für das OSIRIS-Projekt dar.

Die Ergebnisse der Feldstudie sorgten dagegen für Erstaunen und teilweise auch Enttäuschung. Die einzig unter dem Kriterium der Nutzerzufriedenheit durchgeführte Untersuchung belegte, daß hinsichtlich der Qualität der Rechercheergebnisse und der Zu-

friedenheit der Nutzer mit dem System OSIRIS genauso gut abschnitt wie der herkömmliche OPAC.

Dieses Ergebnis ist aus bibliothekspolitischer Perspektive durchaus brisant und zunächst auch ein Erfolg für das OSIRIS-Projekt. Der in Osnabrück eingesetzte OPAC (vgl. 2.2) wird von einer europaweit agierenden Firma vertrieben und ist das Standardbibliothekssystem im ganzen Norddeutschen Bibliotheksverbund. Die für den Betrieb des OPAC allein am lokalen Standort Osnabrück nötige Hardware kostet etwa 20 mal mehr als die für OSIRIS 1.0 benötigte Maschine und zusätzlich zu den lokal zur Betreuung eingesetzten Kräften ist in Göttingen ein eigenes „Bibliotheksrechenzentrum“ mit über 100 Mitarbeitern damit beschäftigt, allein im Norddeutschen Bibliotheksverbund den OPAC zu pflegen und zu betreuen. Die genauen Lizenzgebühren sind aufgrund der Sonderbedingungen des verbundweiten Einsatzes nicht bekannt, sie liegen je lokaler Bibliothek aber in der Größenordnung mehrerer zehntausend Mark jährlich. Daß OSIRIS nach nur zwei Jahren Entwicklungsarbeit zu einem praxisreifen System wurde, das für einen Bruchteil der Kosten zukunftsweisende Konzepte für die Suche auf dem gesamten Datenbestand der Universitätsbibliothek Osnabrück umsetzt und in den Augen der Benutzer ein dem OPAC vollständig ebenbürtigen System darstellt, ist ein Politikum, das die zentralistischen Tendenzen in der deutschen Bibliothekslandschaft in Frage stellt. Das große Interesse, das OSIRIS in der Öffentlichkeit erfahren hat, ist auch vor diesem Hintergrund zu sehen.

Unter Berücksichtigung der (verglichen mit dem OPAC) deutlich *größeren* Leistungsfähigkeit (siehe 3.3) und den enormen Anstrengungen hinsichtlich der natürlichsprachlichen Eingabe erscheint das Ergebnis der Feldstudie aber als eine herbe Enttäuschung. Aus Sicht der Entwickler schafft es OSIRIS, das neue Interaktionsmöglichkeiten bietet und im Schnitt zehnmal mehr relevante Treffer liefert nicht, seine Überlegenheit gegenüber einem altertümlichen Großrechnerprogramm dem Benutzer zu vermitteln.

Auch die im Anschluß durchgeführten Evaluationsstudien (Moranz, 1998; Moranz et al., 1999; Moranz, 2000), deren Ergebnisse bereits referiert wurden, bestätigten OSIRIS wohl den Rang eines professionellen und alltagstauglichen Bibliotheksrecherche-systems, konnten aber hinsichtlich der Nutzerzufriedenheit und dem Nutzerverhalten keine eklatante Überlegenheit von OSIRIS gegenüber dem OPAC feststellen.

Es bleiben so am Ende dieses Abschnittes eine Reihe grundsätzlicher Fragen:

- Waren die Untersuchungen der Psychologen unzureichend ? Angesichts des zeitlichen Umfangs der Arbeiten, der gewissenhaften Vorbereitung der Versuche und der vollständigen Transparenz gegenüber dem OSIRIS-Projekt bei der Durchführung der Versuche besteht zu dieser Annahme keine Veranlassung. Die durchgeführten Untersuchungen haben durchaus Grenzen, und einige Fragen sind noch ungeklärt. Wie Moranz (2000) selbst schreibt, treffen die bisher durchgeführten Untersuchungen keine Aussage darüber, ob *erfahrene* OSIRIS-Nutzer ähnlich wie die OPAC-Nutzer sich auf einige wenige Suchstrategien festlegen, oder ob sie weiterhin das volle Spektrum der möglichen Strategien ausschöpfen.

Ebenfalls ungeklärt bleibt bislang der Einfluß der verbesserten Ergebnispräsentation von OSIRIS auf die Gesamtwirkung des Systems. Nichts deutet jedoch darauf hin, daß die Untersuchungen verfälschte Ergebnisse geliefert hätten.

- Waren die Erwartungen des OSIRIS-Projektes bezüglich der Wirkung des Systems auf den Nutzer übertrieben ? Angesichts der in 2.2 geschilderten Misere der Benutzer am herkömmlichen Katalog und den in 3 (insbesondere 3.3) dargestellten Möglichkeiten von OSIRIS ist die ebenbürtige Wahrnehmung der beiden Systeme durch die Nutzer tatsächlich merkwürdig.
- Sind die Benutzer überhaupt „reif“ für eine fortschrittliche Software ? Diese Frage ist natürlich provokant überspitzt und betont dabei eine völlig falsche Einstellung, daß nämlich der Benutzer sich der Software anzupassen habe. Der durchaus bedenkenswerte Punkt hinter dieser Frage ist aber die Erfahrung aus den psychologischen Untersuchungen, daß Benutzer oftmals nicht zwischen Systemfehlern und Programmfehlern unterscheiden können. Vielfach fielen die Unzulänglichkeiten des Betriebssystems, die Ideosynkrasien der Browser oder schlicht die Unkenntnis der Benutzer auf OSIRIS zurück. Es drängt sich die Frage auf, wieviel Voraussetzungen an ein allgemeines Verständnis hinsichtlich der Funktionsweise eines PCs eine Software, die von großen Benutzergruppen einfach und sicher bedient werden soll, stellen darf. Vielleicht müssen für Anwendungen wie OSIRIS aus der Sicht der Benutzer vollkommen isolierte d.h. frei von jedem Betriebssystemkontext agierende Klienten gebaut werden ?
- Haben die Benutzer in OSIRIS etwas ganz anderes gesehen oder vermutet als das, was von den Entwicklern umgesetzt und in den Untersuchungen der Psychologen überprüft wurde ? Diese Frage liegt nahe, da die Ansprüche, die die Benutzer an OSIRIS stellen, ganz andere sind als die Maßstäbe, die an den OPAC angelegt werden. Dies wird insbesondere in den Anmerkungen der Versuchspersonen in den freien Kommentarfeldern deutlich, wenn z.B. nach Inhalten verlangt wird. Überspitzt formuliert: Statt sich über die bequeme und gute Recherchemöglichkeit zu freuen, sind die Benutzer enttäuscht, daß sie wie im OPAC nur Buchtitel, nicht aber die Inhalte der Bücher finden. Auch erwarten Benutzer offensichtlich ein tieferes Verständnis der Maschine für ihre Eingaben, wenn sie z.B. Warum-Fragen stellen.

Zwei Konsequenzen sind in dieser Situation wünschenswert:

1. Spezifisch für OSIRIS sollte eine weitere Folgeuntersuchung mit sorgfältig ausgewählten Probanden und überarbeiteten Suchaufgaben weitere Klarheit bringen über den Anteil der natürlichsprachlichen Schnittstelle auf die Gesamtwirkung des Systems. In diesem Zusammenhang wäre auch die Berücksichtigung des inzwischen im Einsatz befindlichen HTML-OPAC wünschenswert. Weiterhin sollten bei dieser Untersuchung Aufschlüsse über das Suchverhalten erfahrener OSIRIS-Nutzer gewonnen werden.

2. Andere Projekte haben nicht unähnliche Erfahrungen mit der Wirkung ihrer Software auf die Benutzer gemacht. Die einschlägigen Berichte über die Wirkung sogenannter intelligenter, z.T. auch natürlichsprachlicher Schnittstellen auf den Benutzer sollten diesbezüglich ausgewertet und auf ein plausibel auf „den Fall OSIRIS“ übertragbares Muster hin untersucht werden.

Konsequenz Nummer 1 kann nur von den Psychologen geleistet werden – wir hoffen, daß sich solche Folgeuntersuchungen auch nach Abschluß des offiziellen OSIRIS-Projektes aufgrund ihrer generellen, das Projekt übersteigenden Aussagekraft noch umsetzen lassen.

Konsequenz Nummer 2 soll im verbleibenden Teil dieser Arbeit angegangen werden. Dazu wird in 6.2 die Literatur zur Akzeptanz von NL-Schnittstellen referiert und auf die Forschung zur Wirkung (anthropomorpher) Schnittstellen eingegangen.

6.2 Natürlichsprachliche Schnittstellen in der Literatur

In diesem Kapitel soll ein Überblick über die aus der Literatur bekannten Untersuchungen zum Verhalten von Menschen an natürlichsprachlichen Schnittstellen gegeben werden. Dabei werden die gängigen Vorstellungen über diese Schnittstellen – z.B. daß Benutzer selbiger kein Training benötigen – anhand der Literatur hinterfragt.

6.2.1 Gibt es generelle Alternativen zu natürlicher Sprache ?

Der Einsatz von Methoden zur Analyse natürlicher Sprache in Retrievalsystemen ist keine Modeerscheinung, sondern macht bereits für die Suche auf textuellen Daten Sinn. Insbesondere kann der Einsatz solcher linguistischer Methoden völlig getrennt von der Entscheidung gesehen werden, ob die Schnittstelle natürlichsprachlich oder anders bedient werden kann. In diesem eingeschränkten Sinne gibt es – auch angesichts der stetig zunehmenden Verknüpfung von lokaler Suchfunktionalität mit internetweiter Suche – keine Alternative zum Einsatz von Analysemethoden für natürliche Sprache.

In einer groß angelegten Untersuchung weist Glöckner-Rist (Glöckner-Rist, 1993a,b) die Notwendigkeit nach, den Endnutzer, der nicht über Information Retrieval Kenntnisse verfügt, mit natürlichsprachlichen Verarbeitungsmethoden zu unterstützen. Dazu untersucht sie drei Benutzergruppen:

- Kooperative Vermittlersuchen: Drei erfahrene Vermittler der Universitätsbibliothek führten je neun Datenbankrecherchen gemeinsam mit informationssuchenden Studenten ohne Retrievalerfahrung durch.
- Nicht-kooperative Vermittlersuchen: Drei erfahrene Vermittler führten je neun Datenbankrecherchen in Abwesenheit der informationssuchenden Studenten ohne Retrievalerfahrung durch.
- Endnutzersuchen: Die neun Klienten der kooperativen Vermittlersuchen führten je drei Recherchen alleine durch.

Den Rechercheaufgaben lag eine schriftliche Problembeschreibung zugrunde, die ein bis zwei Tage vorher durch einen Rechercheiteilnehmer aus einem Artikel ohne Titel, Abstract, Quellenangabe oder Literaturverzeichnis extrahiert wurde. Ziel der Aufgabe war es, den Artikel aufgrund der erstellten Beschreibung in der Datenbank zu identifizieren. Den Teilnehmern war bekannt, daß sie lediglich auf der Grundlage der Beschreibung des Artikels recherchieren würden. Für die kooperative Vermittlersuche und die Endnutzersuche wurde die Beschreibung von den Studenten erstellt, und für die nicht-kooperative Vermittlersuche von den Vermittlern selbst. Die Vermittler bei der kooperativen Suche bekamen den Artikel nicht zu sehen, sondern wurden von den Studenten, als informationssuchenden Klienten, erst anhand der von diesen angefertigten Problembeschreibung informiert. In allen drei Benutzergruppen wurden stets

dieselben neun Artikel verwendet, die alle eine experimentalpsychologische Untersuchung beschrieben.

Als Ergebnis der Datenbankrecherchen⁸⁶ stellt Glöckner-Rist fest, daß die nicht-kooperativen Vermittler deutlich effektivere⁸⁷ Problembeschreibungen als die Endnutzer erstellten. Dabei entsprechen die von den professionellen, nicht-kooperativen Vermittlern gewählten Suchwörter offenbar solchen, von denen Retrievalspezialisten erwarten, daß sie in den Nachweisen der Artikel auftreten würden. Hingegen leiteten die Endnutzer häufiger als die nicht-kooperativen Vermittler erfolglose Suchwörter ab: „Endnutzer formulieren also auch dann relativ ineffektive Problembeschreibungen, wenn sie genau wissen, wonach sie suchen“ (Glöckner-Rist, 1993a, :196).

Der Grund dafür ist nach Glöckner-Rist, daß den Endnutzern Kenntnisse über den datenbankspezifischen Sprachgebrauch fehlen: Obwohl sie die Artikelinhalte ähnlich wie die Vermittler (im nicht-kooperativen Fall) konzeptualisieren, d.h. im Prinzip dieselben Suchwörter aus dem Artikel auswählen, scheitern sie an der Wahl der Wortformen, die nicht mit dem in der Datenbank verwendeten Vokabular übereinstimmen. Dieses Problem beeinträchtigt nicht nur die Formulierung der Suchanfragen (im Fall der Endnutzerrecherche), sondern bereits die Formulierung der Problemstellung, wie sich an den Schwierigkeiten bei den kooperativen Datenbankrecherchen zeigte: „Die kooperativen Vermittler, die nur vermittelte Kenntnisse über die Suchprobleme hatten, konnten die Erfolgswahrscheinlichkeit der von den Endnutzern vorgegebenen Problemwörter offensichtlich weniger sicher beurteilen als die Rechercheure der beiden anderen Aufgabenkontexte, die selbst ausgewählte Problemwörter zu beurteilen hatten“ (Glöckner-Rist, 1993a, :196). Dies führt im Ergebnis dazu, daß die kooperativen Vermittler weniger effektive Suchanfragen formulieren als die nicht-kooperativen Vermittler.

Glöckner-Rist (1993a) und Glöckner-Rist (1993b) untersuchen nicht, ob der Benutzer mit einem Information Retrieval System in natürlicher Sprache kommunizieren sollte. Überdeutlich wird aber, daß zumindest bei der Abbildung der Suchanfrage auf die Daten linguistisches Wissen nötig ist. Ohne dieses Wissen erscheinen Retrievalsysteme, die auf textuellen Daten suchen, nicht für Nutzer ohne Spezialkenntnisse erfolgreich bedienbar zu sein.

Ob die – für die Suche auf textuellen Daten bereits vorhandenen – Analysekomponenten auch für die Interaktion mit dem Benutzer verwendet werden, muß unter Berücksichtigung der Bedienfreundlichkeit der Schnittstelle entschieden werden.

Der Artikel von Shneiderman et al. (1998) macht in diesem Zusammenhang eine Reihe von Mißverständnissen deutlich, die in der Diskussion um die vermeintlich bessere Schnittstelle vorherrschen. Shneiderman et al. stellen einen Entwicklungsprozeß in vier Stufen vor, der zu besseren Schnittstellen führen soll. Dabei wird viel Wert gelegt auf eine verständliche, übersichtliche und aus der Sicht des Benutzers beherrschbare Schnittstelle. An den Stellen, an denen die Autoren sich aber über die Funkionali-

⁸⁶DIMDI und PsycInfo, beide konventionell mit Booleschen Ausdrücken zu bedienen.

⁸⁷Gemessen an der Ergebnismenge der Suchanfragen.

tät des entstehenden Systems auslassen⁸⁸ wird deutlich, daß ihr Ansatz unterhalb der Oberfläche stark traditionellen Systemen verpflichtet ist, deren Schwächen die Vorteile der Oberfläche deutlich überwiegen.

So werden an der Oberfläche Möglichkeiten zur Suche nach zusätzlichen Dokumentenmerkmalen vorgesehen, obwohl aus den Benutzeruntersuchungen bekannt ist, daß die Nutzer von Online-Katalogen bereits jetzt mit den verfügbaren Suchaspekten überfordert sind (vgl. 2.2 sowie Recker et al. (1996)). Bei der Formulierung der Suchanfragen wird auf Wiedererkennbarkeit Wert gelegt und daher die Syntax der WWW-Suchmaschinen („city guide + Boston“) imitiert – allerdings ohne zu hinterfragen, ob die Benutzer mit diesem Instrumentarium zurechtkommen,⁸⁹ und ohne die Angemessenheit der Suchanfrage für den Datenbestand zu überprüfen. Dies wird insbesondere dann deutlich, wenn die Verwendung von Stopwörtern vorgeschlagen wird (siehe 2.2). Gleichzeitig wird der Begriff „natural language“ überstrapaziert, wenn jede Möglichkeit, Wörter ohne Operatoren – wie „+“ – einzugeben, als natürlichsprachliche Eingabe gewertet wird. In diesem Sinne sind auch syntaktisch zusammenhanglose Reihungen von Wörtern „natürlichsprachliche Eingaben“ und insbesondere tritt die Schwierigkeit auf, daß die Verwendung von „und“, „oder“ usw. ambig wird, da es sich sowohl um Stopwörter, Boolesche Operatoren oder eben natürlichsprachliche Ausdrücke handeln kann. Dies ist aber keine Schwäche einer natürlichsprachlichen Schnittstelle, sondern beruht auf der unzureichenden Definition natürlicher Sprache durch Shneiderman et al. Unter diesen Voraussetzungen nimmt es aber nicht Wunder, daß die Autoren „natürlichsprachliche“ Schnittstellen kritisch betrachten und die oben gezeigten herkömmlichen Eingabemethoden, wie sie an Suchmaschinen verwendet werden, zusammen mit umfangreichem Systemfeedback präferieren.

Es geht im Folgenden in dieser Arbeit also nicht mehr darum, zu entscheiden, ob natürlichsprachliche Analysemethoden in Dialogsystemen eingesetzt werden sollten, sondern nur noch um die Kriterien für die Entscheidung über den genauen Zweck und Umfang ihres Einsatzes.

6.2.2 Natürliche Sprache im Vergleich zu formalen Sprachen

Was ist nun über die konkreten Auswirkungen des Einsatzes natürlicher Sprache als Anfragesprache an einer Schnittstelle aus der Literatur bekannt ?

Zunächst muß einschränkend vorweggeschickt werden, daß die meisten Untersuchungen zu diesem Thema angesichts der schnellen Entwicklung von Software, Hardware, sowie in Hinblick auf Art und Umfang der Suchdomäne überholt sind. Daher erscheinen sie zum Teil nur schwer untereinander und mit den aktuellen Untersuchungen zu OSIRIS überwiegend gar nicht vergleichbar. Für den Test einer existierenden natürlichsprachlichen Schnittstelle gilt auch immer die, von Jarke et al. erfreulich klar artikulierte, Beschränkung auf den verwendeten Stand der Technik:

⁸⁸Hier greifen Shneiderman et al. stets auf Beispiele zur Suche in Bibliothekskatalogen zurück.

⁸⁹Was zu bezweifeln ist, siehe 2.2.

„Thus, what is really being investigated is the extent to which restrictions (characteristics) of a particular [natural language interface] influence how subjects use that system rather than how they use 'pure' natural language [...]. Consequently, the degree to which any evaluation study of a particular system can be generalized is open to question.“(Jarke et al., 1985, :97)

Eine Ausnahme bilden sorgfältig angelegte Wizard-of-Oz Experimente (s. S. 224), die – weil es sich um Simulationen handelt – bis zu einem gewissen Grade von den aktuellen Möglichkeiten der Technik unabhängig sind und zu generellen Aussagen über das Verhalten von Menschen gelangen können.

Eine weitere Schwierigkeit vieler Untersuchungen liegt in der Vermischung verschiedener Untersuchungsbedingungen (vgl. auch Moranz (1998)). Zoepritz (1983) bemängelt z.B., daß Versuchspersonen mit einer formalen Anfragesprache gleichzeitig auch etwas über die zugrundeliegende Datenstruktur lernen. Insofern sind sie den Benutzern des natürlichsprachlichen Systems überlegen, weil von denen erwartet wird, daß sie ohne jede Vorkenntnis das System bedienen können. Diesen Effekt macht Zoepritz z.B. bei Shneiderman (1980) aus, dessen Versuchspersonen mit natürlicher Sprache zunächst deutlich schlechter abschnitten, als mit SEQUEL.⁹⁰ Hatten die Versuchspersonen jedoch zunächst SEQUEL gelernt und damit am Datenbestand gearbeitet, verschwand der Unterschied.

Um zu beschreiben, wie einfach, natürlich und effizient eine Anfragesprache ist, wird der Begriff der „Habitabilität“⁹¹ verwendet. Eine Anfragesprache gilt als habitable, wenn der Benutzer mit Hilfe der Sprache alles ausdrücken kann, wovon er glaubt, daß es das System verstehen müsse, und daß es für die Lösung einer Aufgabe nötig sei. Ogden und Bernick (1997) unterscheiden vier Domänen der Habitabilität:

- Die konzeptuelle Domäne einer Sprache beschreibt die Menge der Objekte und Aktionen, auf die der Benutzer mit Hilfe der natürlichsprachlichen Schnittstelle referieren kann. Die konzeptuelle Domäne einer Sprache ist zu unterscheiden von der konzeptuellen Domäne des darunterliegenden Systems: Im Fall einer Bibliotheksschnittstelle ist es durchaus möglich, daß die konzeptuelle Domäne der Sprache Objekte wie z.B. „Filme“ enthält, obwohl in der darunterliegenden Datenbank keine solchen Objekte enthalten sind. Dies ist sinnvoll, um auf Benutzeranfragen adäquat reagieren zu können.
- Die funktionale Domäne beschreibt die Einschränkungen hinsichtlich Komplexität und Kombinierbarkeit in Anfragen. Z.B. können Anfragen, die insgesamt zu komplex sind, weil nach mehr als einer Information auf einmal gefragt wird, in mehrere, einzeln bearbeitbare Anfragen aufgeteilt werden. Ogden und Bernick (1997) versäumen es an dieser Stelle, zwischen der funktionalen Domäne

⁹⁰Ein Vorläufer von SQL.

⁹¹„Habitability“

der Sprache und der des darunterliegenden Systems zu unterscheiden: Es ist ein Unterschied, ob eine Anfrage innerhalb der Anfragesprache nicht formulierbar und analysierbar ist, oder ob sie nicht von der darunterliegenden Datenbank beantwortet werden kann.

- Die syntaktische Domäne umfaßt die innerhalb der Anfragesprache möglichen syntaktischen Konstruktionen. Z.B. ist die syntaktische Domäne der Anfragesprache von OSIRIS nicht vollständig,⁹² da u.a. keine eingebetteten Nebensätze eingegeben werden können.
- Die lexikalische Domäne beschreibt den durch die Systemlexika abgedeckten Wortschatz des Systems.

Eine geeignete Anfragesprache sollte im Idealfall in allen vier Domänen habitable sein. Dabei handelt es sich immer um Ziele, denen man sich annähern möchte. Wie Ogden und Bernick (1997) schreiben, wäre dasjenige System, das in allen vier Domänen vollständig den Erwartungen der Nutzer entspricht, ein guter Kandidat für den Turing Test.

Alternativ zur unbedingten Annäherung an diese Ziele ist es aber auch denkbar, daß die Einschränkungen im Systemdesign so offenbar sind, daß der Nutzer keine Schwierigkeiten hat, sie zu erkennen und zu verstehen, und damit die Sprache immer noch habitable ist.

Um die Eigenschaften einer Schnittstelle zu überprüfen, sind neben den richtigen Versuchspersonen auch die passenden Aufgaben nötig. Aufgaben, die von den Experimentatoren selbst entworfen werden, haben den Vorteil, daß sie auch für Bereiche und Anwendungen verfügbar sind, zu denen es noch keine Benutzer gibt.

Ein Nachteil dieser Aufgaben ist, daß sie in der Regel alle lösbar sind, weil Experimentatoren normalerweise nur lösbare Aufgaben generieren. Im Gegensatz dazu testen Benutzer die konzeptuelle und funktionale Domäne eines Systems voll aus, da sie in der Formulierung ihrer Probleme frei sind. Entsprechend werden von Nutzern auch unlösbare Aufgaben erstellt.

Ein Nachteil von durch Nutzer generierten Aufgaben liegt in der mangelnden Systematisierung: Im Gegensatz zu den Experimentatoren erforschen die Nutzer die Möglichkeiten der Schnittstelle nicht systematisch. Da sie andererseits aber ihre eigenen Suchbedürfnisse zugrundelegen, entsprechen die so generierten Aufgaben eher der Praxis der Nutzer und testen die Schnittstelle in Bezug auf ihre Alltagstauglichkeit. Als vielversprechende Alternative führen Ogden und Bernick (1997) an, den Versuchspersonen Tabellen oder Graphen zu zeigen, in denen an einigen Stellen Information fehlt, die dann mit Hilfe des Systems ergänzt werden soll. In diesem Szenario hat der Experimentator maximalen Einfluß auf die Aufgabengestaltung, ohne im konkreten Fall die Eingaben der Versuchspersonen vorwegzunehmen.

⁹²Die syntaktische Domäne ist aus den in 3.1 beschriebenen Erwägungen bewußt unvollständig gehalten.

In vielen Fällen sind die Systeme, deren Eigenschaften getestet werden sollen, noch nicht in einem Zustand, in dem sie von unbedarften Personen ernsthaft benutzt werden könnten. Auch können die Systembeschränkungen so stark sein, daß ein natürlicher Umgang dem Benutzer nicht möglich erscheint. In diesen Fällen, sowie als Vorbereitung für eine noch nicht begonnene Systementwicklung, oder auch als Beitrag zur Benutzerforschung, werden sogenannte „Wizard-of-Oz Simulationen“ durchgeführt. In Anlehnung an den Roboter mit dem menschlichen Herz aus dem gleichnamigen Märchen simuliert dabei ein menschlicher Experimentator die natürlichsprachliche Schnittstelle und kommuniziert über eine Rechnerverbindung mit der Versuchsperson. Der Vorteil einer solchen Simulation liegt in dem geringen technischen Aufwand, verglichen mit der Entwicklung eines vollständigen Systems. Außerdem können auf diese Weise – bereits vor Beginn der eigentlichen Entwicklungen – wichtige Erkenntnisse über das Entwicklungsziel gewonnen werden. Von Nachteil ist die unter Umständen mangelnde Konsistenz der Reaktionen des Experimentators: So beschreibt Diaper (1986) ein Wizard-of-Oz Experiment, bei dem die Versuchspersonen aufgrund der gelegentlichen Tippfehler des Experimentators die Überzeugung, mit einer Maschine zu kommunizieren, verloren (zu Diaper s.a. S. 244).

Alternative Vorgehensweisen, eine natürlichsprachliche Schnittstelle zu testen, sind Testsuiten, schriftliche Aufgaben für Testpersonen und Feldstudien (vgl. Whittaker und Stenton (1989) und Saracevic (1995)).

Testsuiten bestehen aus einem Korpus natürlichsprachlicher Anfragen, die in syntaktischer, semantischer, lexikalischer und pragmatischer Hinsicht das Spektrum der vom System erwarteten Analyseleistung testen sollen (vgl. Callan und Croft (1993)). Die Überprüfung, ob die Testsuite tatsächlich komplett ist, erscheint im allgemeinen Fall aber nur sehr schwer möglich zu sein, wenn ernsthaft auch Diskursphänomene getestet werden sollen. Diese sind sehr stark aufgabenabhängig, d.h. eine Testsuite, die auch Diskursphänomene testen soll, wird nicht nur vollständig in Hinblick auf die Phänomene der natürlichsprachlichen Eingabe sein müssen, sondern auch dem Aufgabenbereich der Schnittstelle Rechnung tragen müssen. Damit wird die Testsuite aber nicht mehr einfach auf andere Schnittstellen übertragbar sein. In Kombination mit anderen Verfahren erscheint der Einsatz einer Testsuite (so man denn eine hat) aber sinnvoll.

Schriftliche Aufgaben umgehen alle Probleme, die ein prototypisches System hinsichtlich Stabilität oder Domänenabdeckung haben kann. Die Versuchspersonen bekommen eine Aufgabe gestellt und sollen für diese eine Systemanfrage konstruieren. Diese Anfrage wird aber nicht an ein System gestellt, sondern (auch) von menschlichen Experten evaluiert. Dieser Ansatz eignet sich überhaupt nicht zum Test von Diskursphänomenen, da keine Interaktion zustandekommt. Entscheidend ist auch hier die Auswahl der Aufgaben, deren Beschaffenheit die Repräsentativität der Untersuchung bestimmt.

Für Feldstudien wiederum gelten die möglicherweise aus dem Entwicklungsstand des Systems resultierenden Einschränkungen hinsichtlich Robustheit oder Domänenabdeckung. Dafür haben Feldstudien den Vorteil, daß die Nutzer mit dem echten System arbeiten. Sie sind in ihrer Interaktion nicht eingeschränkt, sondern lösen die Aufga-

ben, die ihren persönlichen Bedürfnissen entsprechen. Ob die Eingaben, die sie dabei tätigen, allerdings durch die Grenzen des Systems beeinflußt werden, d.h. ob es die gleichen Eingaben sind, die sie auch an einem robusteren oder vollständigeren System tätigen würden, ist unklar. Verschiedene Untersuchungen zeigen, daß Benutzer geneigt sind, sich den von ihnen wahrgenommenen Möglichkeiten des Systems anzupassen (s. 6.2.5).

Im Folgenden werden nun die Ergebnisse einiger Untersuchungen zu den Eigenschaften formaler Sprachen im Vergleich zu natürlicher Sprache als Anfragesprache kritisch widergegeben.

Bei Turtle (1994) handelt es sich um eine der wenigen neueren Untersuchungen, die zudem auch einen klaren Vorteil der natürlichsprachlichen Schnittstellen nachweist. Turtle (1994) beschreibt einen Vergleichstest zwischen mehreren Information Retrieval Systemen, die natürlichsprachliche Anfragen akzeptieren und einem System, das mit Booleschen Ausdrücken befragt wird. Das Retrieval fand auf sehr großen textuellen Daten statt. Die Aufgaben für diesen Versuch wurden natürlichsprachlich von menschlichen Experten der Domäne (Jura) entworfen – die Booleschen Ausdrücke wurden von Retrieval Experten erstellt. Dabei schnitten die natürlichsprachlichen Anfragen hinsichtlich Precision und Recall signifikant (24.2 %) besser ab.

Trotz dieser klaren Vorteile für natürliche Sprache als Retrievalmöglichkeit in großen textuellen Datenmengen geht Turtle davon aus, daß in Zukunft weiterhin Boolesche Ausdrücke verwendet werden, da es durchaus Retrievaldomänen gäbe, in denen sie im Vorteil seien. Welche Art von Domänen das sind und wodurch sie sich auszeichnen, erwähnt Turtle leider nicht.

Jarke et al. (1985) überprüfen aus einer vorangegangenen Laboruntersuchung (Turner et al., 1984) abgeleitete Hypothesen zum Vergleich von natürlichsprachlichen und strukturiertsprachlichen Systemen in einem Feldversuch. Dabei stellen die Autoren fest, daß weder hinsichtlich der Performanz der erfolgreichen Aufgabenbearbeitung, noch hinsichtlich des Erfolges eine Überlegenheit des beschränkt natürlichsprachlichen Systems über die formale Anfragesprache (SQL) nachgewiesen werden kann.

Die von den Versuchspersonen zu bearbeitenden Aufgaben wurden von den verantwortlichen Verwaltungskräften einer US-amerikanischen Universität für die Stipendiatendatenbank dieser Universität und deren Partneruniversitäten formuliert, z.B.: „Es wird nach einer Liste der Stipendiaten im Bundesstaat Kalifornien gefragt, deren Nachname mit 'S' beginnt. Erstelle eine solche Liste mit Angabe von Vorname und Nachname.“ Diese Aufgabe sollte dann von den eigentlichen Versuchspersonen bearbeitet werden, indem diese eine oder mehrere natürlichsprachliche Anfragen („What are the last names and first names of all California Alumni whose last name is like S%?“) bzw. SQL-Anfragen⁹³ formulieren und an das jeweilige System richten.

⁹³

```
Select  lastname,  firstname
From    donors
Where   srccode = 'al' and state = 'ca' and lastname like 's%';
```

Hinsichtlich der Bearbeitbarkeit der Aufgaben stellen Jarke et al. fest, daß 75 % der Aufgaben in beiden Anfragesprachen lösbar und alle Aufgaben zumindest in Teilen lösbar waren. Da die Aufgaben nicht von den Experimentatoren, sondern von Versuchspersonen generiert wurden, war dieser Punkt nicht selbstverständlich. 15.6 % der Aufgaben waren nicht in SQL lösbar, verglichen mit 26.2 % für natürliche Sprache. Damit erscheint SQL mächtiger in der Ausdruckskraft zu sein.

Auch hinsichtlich des Anfrageerfolges schneidet SQL besser ab, und zwar sowohl auf der Ebene der Aufgaben, als auch auf der Ebene der einzelnen Systemanfragen. Auf der Ebene der Aufgaben waren die SQL-Benutzer in etwa doppelt so vielen Fällen erfolgreich,⁹⁴ verglichen mit den NL-Benutzern – unabhängig davon, ob nur vollständig lösbare (52.4 % vs. 23.6 %) oder auch teilweise lösbare Aufgaben betrachtet werden (44.2 % vs. 17.1 %). Auf der Ebene der Systemanfragen beträgt das Verhältnis sogar 3:1 zugunsten von SQL. Jarke et al. betonen, daß mit einer Erfolgswahrscheinlichkeit von bestens 70 % beide Anfragesprachen enttäuschend abschnitten und ein professioneller Einsatz unter diesen Umständen für keine der beiden Anfragesprachen vorstellbar sei.

Für die SQL-Nutzer zeigten sich interindividuell und über die Zeit keine signifikanten Unterschiede, während die NL-Benutzer mit einer Erfolgsquote von nur 4.8 % begannen und am Ende bei 30 % lagen, ohne daß sich die Zahl der formulierten Anfragen dabei verändert hätte. Zudem gab es große individuelle Unterschiede. Das bedeutet andersherum, daß die SQL-Benutzer im ersten Anlauf sehr viel besser als die NL-Nutzer abschnitten, daß sich im weiteren Verlauf dieser Unterschied allein durch eine Verbesserung der NL-Nutzer aber stark verringerte, wobei für einzelne Aufgaben die NL-Benutzer sogar die SQL-Nutzer überholten. Jarke et al. vermuten, daß es die leichte Zerlegbarkeit in mehrere einfache, natürlichsprachliche Anfragen ist, die diese Aufgaben auszeichnet.

Gründe für das Scheitern der Benutzer waren, laut Jarke et al., häufig technische Probleme mit der Oberfläche (21.1 % für SQL bzw. 25 % für NL). Dies ist ein weiterer Beweis für die bereits in 5.2.2 vertretene These, daß nur eine Kombination aus verlässlicher und einfach zu bedienender Oberfläche und intelligenten Verarbeitungsmechanismen erfolgreich sein kann. Zudem unterschieden sich die Benutzerfehler deutlich: SQL-Nutzer scheiterten zum überwiegenden Teil an Tippfehlern, Syntaxfehlern und semantisch nicht adäquaten Anfragen, während das Hauptproblem der NL-Nutzer mangelnder Sprachumfang und fehlende Funktionalität waren. Wichtig ist in diesem Zusammenhang die Feststellung von Jarke et al., daß in einem natürlichsprachlichen Anfragesystem die Verantwortung für Eingabefehler vom Benutzer zum System verlagert wird. Während die Fehler der SQL-Nutzer ihre Ursache in der mangelnden Beherrschung von SQL hatten, waren bei den Benutzern der natürlichsprachlichen Schnittstelle Beschränkungen des Systems die Fehlerursache. Die subjektive Beschreibung des Scheiterns durch die Versuchspersonen lautete denn auch für die SQL-Nutzer

⁹⁴Erfolg ist das Mittel über die individuellen Erfolgsraten, d.i. das Verhältnis zwischen bearbeiteten und erfolgreich gelösten Aufgaben je Person.

überwiegend „Zeitmangel“, und für die NL-Nutzer „Konnte nicht herausfinden, wie die Aufgabe zu lösen ist“.

Die Benutzer der NL-Schnittstelle stellten je Aufgabe 50 % mehr Anfragen an das System als die SQL-Nutzer, brauchten dennoch aber nur etwa 15 % mehr Zeit als die SQL-Nutzer, um eine Aufgabe zu bearbeiten. Dem steht gegenüber, daß die von den SQL-Nutzern formulierten Systemanfragen dreimal so umfangreich⁹⁵ wie die natürlichsprachlichen Anfragen waren. Die Formulierung einer solchen Anfrage brauchte 40 % mehr Zeit als die Formulierung einer natürlichsprachlichen Anfrage. Trotz der deutlich höheren Anzahl natürlichsprachlicher Systemanfragen, erweist sich die natürliche Sprache insgesamt dennoch als kompakter, verglichen mit SQL. Natürlichsprachliche Anfragen sind also schneller und einfacher zu formulieren als SQL-Anfragen. Für Jarke et al. ist dies eine vielversprechende Aussicht für zukünftige, verbesserte Systeme.

Walker und Whittaker (1990) untersuchen, ob Personen, die Aufgaben aus ihrem unmittelbaren Berufsalltag lösen sollen, dies mit einer Menü-Schnittstelle oder mit einer natürlichsprachlichen Schnittstelle erledigen. Untersucht wurde, welche Merkmale das natürlichsprachliche System gegenüber dem Menü-System aufweist und welche dieser Merkmale die Aufgabenbearbeitung besser unterstützen. Dazu wurden Zahl und Art der Fehler sowie der Umgang mit diesen Fehlern betrachtet. Walker und Whittaker stellen fest, daß die natürlichsprachliche Schnittstelle häufiger und trotz einer relativ hohen Fehlerrate konstant weiter genutzt wurde. Walker und Whittaker führen dies auf die erweiterten Eingabemöglichkeiten der natürlichsprachlichen Schnittstelle im Vergleich mit der Menü-Schnittstelle zurück.

In einem Wizard-of-Oz Experiment von Small und Weldon (1983) sollten die Versuchspersonen Aufgaben nacheinander, sowohl an einem strukturiertsprachlichen (SEQUEL), als auch an einem (simulierten) natürlichsprachlichen Eingabesystem erledigen. Es wurden zwei Gruppen untersucht, von denen eine zunächst mit SEQUEL arbeitete, und danach die natürlichsprachliche Schnittstelle benutzte. Die andere, zweite Gruppe ging in umgekehrter Reihenfolge vor.

Es waren vier Aufgaben zu bearbeiten:

1. Eine „Simple Mapping (SM)“ Aufgabe, bei der in einer Tabellenspalte nach einem einzelnen Kriterium gesucht werden sollte.
2. Eine „Composing from more than one table (COMP)“ Aufgabe, bei der Informationen aus zwei unterschiedlichen Tabellen verknüpft werden mußten.
3. Eine Aufgabe mit Booleschem „und“ sowie
4. eine Aufgabe mit Booleschem „oder“.

⁹⁵Anzahl der verwendeten Wörter je Anfrage.

Insgesamt zeigte sich, bis auf einen Geschwindigkeitsvorteil für SEQUEL, bei Aufgabentyp eins und zwei kein Unterschied zwischen SEQUEL und dem natürlichsprachlichen System. Insbesondere ließ sich auch für die Reihenfolge, in der die beiden verschiedenen Schnittstellen benutzt wurden, kein Effekt nachweisen. Daß Aufgaben vom Typ eins und zwei (im Gegensatz zu Typ drei und vier) schneller mit SEQUEL bearbeitet werden konnten, führen Small und Weldon auf die strukturelle Ähnlichkeit zwischen der Aufgabe und der Anfragesprache SEQUEL zurück, die explizit in Termini von Tabellen und Spalten arbeitet. Im Gegensatz dazu, so Small und Weldon, erfordere die Bearbeitung der Aufgaben vom Typ drei und vier die Auseinandersetzung mit der Aufgabe, um eine adäquate Anfrage zu generieren. Dabei sei keine der beiden Anfragesprachen im Vorteil gewesen und entsprechend habe sich auch kein Unterschied ergeben.

Als Fazit stellen Small und Weldon fest, daß bis auf die Situation, in der die Struktur einer Aufgabe eine Übereinstimmung mit den Merkmalen der Anfragesprache zeigt, keine Unterschiede zwischen strukturierter und natürlichsprachlicher Anfrage festgestellt werden konnte – insbesondere auch nicht hinsichtlich der Reihenfolge, in der die beiden Anfragesprachen gelernt wurden. Damit stehen die Ergebnisse von Small und Weldon (1983) im Widerspruch zur häufig genannten These (siehe Zoeppritz (1983), aber auch Small und Weldon (1983) selbst), daß Benutzer strukturierter Anfragesprachen in der Trainingsphase durch eine Exploration der Struktur der Daten einen Vorteil gegenüber den Benutzern der natürlichsprachlichen Anfrage genießen.

Borenstein (1986) verglich in einem Wizard-of-Oz Experiment ein besser aufbereitetes, und mit weiteren Hilfetexten angereichertes UNIX „MAN-Page“-System, mit einem prototypischen System, mit graphischer Oberfläche sowie einer simulierten natürlichsprachlichen Schnittstelle. Alle drei Systeme enthielten dieselbe Information, und Borenstein (1986) konnte keine Unterschiede in der Verwendung dieser Systeme ausmachen. Er schließt daher, daß die Qualität der verfügbaren Information das entscheidende Kriterium ist, und nicht die Art der Schnittstelle.

Bell und Rowe (1992) kommen zu einem konträren Ergebnis bei der Untersuchung von drei unterschiedlichen (kommerziellen) Schnittstellen, denen dieselbe Datenbasis mit Informationen über Lehrer, Schüler und Unterrichtsfächer zugrundelag. Verglichen wurden eine graphische Schnittstelle für SQL, eine Schnittstelle mit einer formalen Anfragesprache⁹⁶ und eine natürlichsprachliche Schnittstelle.

Untersucht wurden 55 Versuchspersonen, die ihren unterschiedlichen Vorkenntnissen entsprechend in vier Gruppen (Anfänger, mit normaler Computervorerfahrung, mit Datenbankkenntnissen, mit Programmierkenntnissen) eingeteilt wurden. Keine der Versuchspersonen konnte für sich beanspruchen, ein Experte zu sein. Alle Versuchspersonen durchliefen vor dem eigentlichen Versuch eine Lernphase, in der ihnen anhand verschiedener – von Hilfestellung begleiteten – Aufgaben der Umgang mit der Datenbank erläutert wurde. Beide Phasen unterschieden sich für die Teilnehmer nur

⁹⁶Eine Variante von SQL.

durch die verwendeten Schnittstellen voneinander, so daß alle Teilnehmer die gleichen Vorkenntnisse über den Datenbestand hatten. Im eigentlichen Versuch wurden die in der Versuchsphase trainierten Aufgabenarten eingesetzt – ergänzt um zwei weitere, untrainierte Aufgaben.

Für die trainierten Aufgaben stellten Bell und Rowe (1992) in der Versuchsphase einen deutlichen, positiven Einfluß der Vorerfahrung der Nutzer auf die Erfolgswahrscheinlichkeit fest. D.h. Versuchspersonen mit Programmier- oder Datenbankerfahrung schnitten deutlich besser ab, als Anfänger und Nutzer mit normaler Computererfahrung. Zudem ist die Erfolgsrate, die die Versuchspersonen an einer Schnittstelle erzielen, aufgabenabhängig: Es gibt Aufgaben, die in SQL und mit der graphischen Schnittstelle zu SQL sehr komplex sind (und von keiner Versuchsperson an diesen Schnittstellen gelöst werden konnten), die aber natürlichsprachlich ohne größeren Aufwand formulierbar und lösbar sind.

Als nicht signifikante Tendenz stellen Bell und Rowe weiter fest, daß für SQL und die graphische Schnittstelle zu SQL Anfänger genauso gut abschneiden wie solche Personen, die normale Computererfahrung haben. An der natürlichsprachlichen Schnittstelle hingegen schnitten die Anfänger schlechter ab, verglichen mit den Versuchspersonen, die über normale Computervorerfahrung verfügten.

Die erste, nicht trainierte Aufgabe (Typ Nicht-Existenz), wurde von den Nutzern der natürlichsprachlichen Schnittstelle hervorragend⁹⁷ gelöst. Im Gegensatz dazu gelang es keiner Versuchsperson, die mit der SQL-Schnittstelle arbeitete, diese Aufgabe zu lösen.⁹⁸ Die zweite, untrainierte Aufgabe (selbstreferentieller join), wurde insgesamt nur von einer einzigen Versuchsperson⁹⁹ gelöst, was darauf hindeutet, daß hier die Art der Aufgabe und nicht die Eigenschaften der Schnittstelle eine Rolle spielten.

Als Ergebnis halten Bell und Rowe (1992) fest:

- Kein System war so gut, daß es allen anderen Systemen in allen Belangen überlegen war. Jedes System war aber jeweils in bestimmten Situationen den anderen überlegen. Und jedes System war verbesserungsbedürftig.
- Die Interaktion mit der natürlichsprachlichen Schnittstelle war anders beschaffen als die mit SQL und der graphischen Oberfläche zu SQL. Das Abschneiden der Nutzer an der natürlichsprachlichen Schnittstelle unterschied sich deutlich von dem an den anderen Schnittstellen. Auch waren an der natürlichsprachlichen Schnittstelle ganz andere Aufgaben schwierig zu bearbeiten als an den anderen beiden sonst konformen Schnittstellen.
- Die Vorerfahrung eines Benutzers beeinflusst an jeder der drei Schnittstellen direkt das Abschneiden eines Benutzers. Überraschend daran war, daß gerade

⁹⁷Drei von fünf der Anfänger und der normal erfahrenen Nutzer, sowie alle Nutzer mit Datenbank- und Programmiererfahrung.

⁹⁸Der Aufgabentyp war unlösbar mit der graphischen Oberfläche.

⁹⁹Ein Nutzer der graphischen Oberfläche mit Datenbankerfahrung.

die natürlichsprachliche Schnittstelle sowohl von der Computer-, als auch der Datenbank-Vorerfahrung der Benutzer profitierte. Auch der besondere Vorteil, den die Programmiererfahrung der Benutzer für die Bedienung der graphischen Oberfläche darstellte, kam unerwartet. Eine graphische Oberfläche zu einer formalen Datenbankanfrage sollte eigentlich so gestaltet sein, daß Erfahrung im Umgang mit Programmiersprachen nicht ins Gewicht fällt.

Entscheidend ist im Zusammenhang mit dieser Arbeit die neue Qualität, die Bell und Rowe (1992) eindeutig für die Benutzung der natürlichsprachlichen Schnittstellen feststellen. Bell und Rowe werten diese Erkenntnis als Erfolg, da es ein elementares Ziel der Entwicklung natürlichsprachlicher Schnittstellen sei, die Mensch-Maschine-Interaktion grundlegend zu verändern.

Obwohl dies zweifellos richtig ist, fällt es doch schwer von einem echten Erfolg zu sprechen, solange wesentliche Ziele, wie die leichtere Bedienbarkeit für Benutzer ohne Vorkenntnisse, nicht erreicht werden. Dennoch bleibt festzuhalten, daß natürlichsprachliche Schnittstellen offenbar „anders“ sind, wie auch immer man das bewerten will.

Zu einem ähnlichen Ergebnis kommen Capindale und Crawford (1990), die erfreut von der sehr positiven Reaktion der Benutzer auf die (kommerzielle) natürlichsprachliche Schnittstelle zu einer Datenbank berichten:

„In the study it was found that natural language is an effective method of interaction for casual users with a good knowledge of the database, who perform question-answering tasks, in a restricted domain.“ (Capindale und Crawford, 1990, :360)

Auffällig ist allerdings, daß die Kenntnis des Datenbestandes und eine eingeschränkte Domäne, als Voraussetzung für eine erfolgreiche Interaktion genannt werden.

Zusammenfassend läßt sich feststellen, daß natürlichsprachliche Schnittstellen im Vergleich zu SQL- oder anderen Schnittstellen offenbar nicht so erfolgreich sind, daß ihre bloße Existenz einen positiven Effekt auf die Nutzer haben würde. Andererseits muß auch gesagt werden, daß die zitierten Untersuchungen – soweit sie konkrete Systeme beschreiben – aufgrund ihres z.T. beträchtlichen Alters und der daraus resultierenden Unterschiede im Entwicklungsstand schwer oder gar nicht zu vergleichen sind. Sicher erscheint nur, daß zum jetzigen Zeitpunkt natürlichsprachliche Schnittstellen im besten Fall die Akzeptanz beim Nutzer erreichen, die auch eine herkömmliche Schnittstelle genießt. Das liegt zum Teil an den vorhandenen technischen Möglichkeiten, aber auch an den spezifisch anderen Problemen, die natürlichsprachliche Schnittstellen aufwerfen (z.B. der Verlagerung der Verantwortung im Fehlerfall vom Benutzer zum System). Ob natürliche Sprache oder eine andere Möglichkeit die bessere Anfrageform ist, muß nach dem Überblick über die Literatur weiter offen bleiben.

6.2.3 Linguistische Phänomene an NL-Schnittstellen

Zu den an natürlichsprachlichen Schnittstellen zu erwartenden und zu beobachtenden linguistischen Phänomenen liegen unterschiedliche Aussagen vor. Tennant (1979) behauptet, daß die Interaktion zwischen Benutzer und Datenbank oft wenig mit dem allgemein angenommenen, geordneten Wechselspiel aus Frage und Antwort zu tun habe. Er zählt eine Reihe unterschiedlicher Arten von Äußerungen auf, die Benutzer nach seinen Erfahrungen an Datenbanken richten würden, die aber nicht im herkömmlichen Sinne als Anfragen an die Datenbank zu interpretieren seien:

- Fragen über den Datenbestand: „Ist das alles, was über 'Math 195' bekannt ist?“
- Fragen zum Vokabular: „Was ist ein 'buser'?“
- Explizites Setzen des Kontextes: „Mich interessiert 'Math 195'.“
- Referenz auf Diskursobjekte: „Jetzt addiere die beiden.“
- Verifikation oder Zusammenfassung von Ergebnissen: „Also brauche ich im Grundstudium nur die Scheine aus 'Math 195' und 'KI 1-6'?“
- Kombination mehrerer Fragen in einer Äußerung: „Welche Studenten haben bestanden, und wo lag der Notendurchschnitt?“
- Aufteilung einer (komplexen) Frage auf mehrere Äußerungen: „Wieviele Studenten haben 1997 ihre Zwischenprüfung gemacht? Und 1998? Und 1999? Liste die Studenten nach ihren Matrikelnummern ab 1997.“

Auch Burton und Steward (1993) widmen sich der Frage nach den satzübergreifenden Referenzen an natürlichsprachlichen Schnittstellen. Dazu untersuchen sie mit einer experimentellen, natürlichsprachlichen Schnittstelle die möglichen Verbesserungen, die eine solche Schnittstelle durch die Behandlung solcher satzübergreifender Beziehungen, wie Ellipsen und Anaphern, erfahren kann. Als Vergleich wird eine natürlichsprachliche Schnittstelle herangezogen, die durch außersprachliche Mittel – eine bequeme Editierfunktion der bereits getätigten Systemeingaben – die Erkennung und Behandlung von Ellipsen und Anaphern ersetzt.

Burton und Steward (1993) können, im Vergleich der beiden Systeme, für den Nutzer keinen Vorteil im ersteren System ausmachen. Im Gegenteil, es erscheint Burton und Steward, bis auf wenige spezielle Situationen, sogar insgesamt komplizierter zu bedienen zu sein. Hatten die Nutzer beide Möglichkeiten zur Verfügung, wurde doppelt so häufig die Editierfunktion gewählt als die Möglichkeit genutzt, natürlichsprachlich eine Ellipse zu verwenden.

Beaulieu (1997) stellt an einer (nicht natürlichsprachlichen) Schnittstelle ebenfalls fest, daß Benutzer überwiegend die ursprüngliche Anfrage modifizieren und erneut

stellen, statt interaktiv Folgefragen zu stellen (s. a. 3.4.5). Beaulieu führt dies darauf zurück, daß das System insgesamt zu kompliziert zu bedienen sei.

Das Untersuchungsergebnis erscheint kurios, denn gerade durch die Behandlung der in der zwischenmenschlichen Kommunikation so häufigen Kontextphänomene soll eine natürlichsprachliche Schnittstelle ja an Natürlichkeit in der Wirkung und Einfachheit in der Bedienung gewinnen. Vor dem Hintergrund, daß Menschen sich einer Maschine gegenüber sprachlich anders verhalten als einem Menschen gegenüber (vgl. 6.2.5), machen die Ergebnisse von Burton und Steward (1993) allerdings Sinn. Für Burton und Steward (1993) ist daher der weitere Ausbau der linguistischen Möglichkeiten einer natürlichsprachlichen Schnittstelle in Hinblick auf Diskursphänomene ein Fehler – sie schlagen die Verwendung einer Subsprache als besser zu bedienende Alternative vor.

Der lexikalischen Variabilität widmen sich Landauer et al. (1983). Sie untersuchen das intuitive Ziel vieler Entwickler, Programmbezeichnungen besonders verständlich zu halten, indem natürlichsprachliche Ausdrücke verwendet werden. Landauer et al. stellen fest, daß je nach Ausbildungsstand der Nutzer sowohl die Benennung der einzelnen Programmfunktionen, als auch die zugrundeliegende Konzeptualisierung dieser Funktionen, sich sehr stark unterscheiden. Auch Nutzer, die sich auf demselben Ausbildungsniveau befinden und in etwa die gleichen Konzepte für Programmfunktionen verwenden, variieren sehr stark hinsichtlich des Vokabulars, mit dem sie die Funktionen bezeichnen.

Hingegen stellt Miller (1981) in einem Papiertest, ohne Beschränkung der natürlichsprachlichen Ausdrucksmöglichkeiten der Versuchspersonen, einen beschränkten Wortschatz in den Anfragen fest, der zudem durch eine große Zahl an Synonymen gekennzeichnet war.

Unbestritten scheinen allerdings die Schwierigkeiten zu sein, die Menschen mit der logischen Interpretation des natürlichsprachlichen „und“ haben (in 4.2.14 ausführlich dargestellt, siehe auch Das-Gupta (1987) und Ogden und Kaplan (1986)). Capindale und Crawford (1990) bestätigen diese Ergebnisse. Damit die Maschine darauf angemessen reagieren kann, schlagen Ogden und Kaplan (1986) eine Interpretationsstrategie vor (s. S. 148), um die vom Benutzer aktuell intendierte Bedeutung der Verwendung von „und“ zu erschließen. Siehe in diesem Zusammenhang auch Willie und Bruza (1995) mit einem Versuch, die Schwierigkeiten von Menschen mit Booleschen Ausdrücken kognitionswissenschaftlich durch zwei unterschiedliche, zugrundeliegende mentale Modelle zu erklären.

Dekleva (1994) untersucht eine kommerzielle natürlichsprachliche Schnittstelle, und versucht in einem Papierexperiment herauszufinden, wie sinnvoll (in Hinblick auf ihren Sprachumfang) diese Schnittstelle eingesetzt werden kann. Dabei stützt er sich auf die (in Sethi (1987) überarbeitete) Klassifikation natürlichsprachlicher Schnittstellen als Ergebnis einer Expertenkonferenz (Hendrix, 1982). Dabei werden zwölf mögliche Eigenschaften der Eingabemöglichkeiten natürlichsprachlicher Schnittstellen betrachtet: Grammatikprüfung, komplexe Nomen, Pronomenverwendung, andere

Referenzen, Ellipsis, Quantifikation, Möglichkeit von Inferenzen, Identifikation von Benutzerzielen, Alltagswissen, Beantwortung von Metafragen, Hilfestellung im Fehlerfall und Hilfe bei der Anfrageformulierung. Anhand dieser zwölf Eigenschaften der Eingabemöglichkeiten (sowie weiterer Charakteristika) werden die Schnittstellen in einer vierstufigen Klassifikation eingeordnet, beginnend bei der niedrigsten Stufe L1:

L1: Schnittstellen dieser Stufe zeichnen sich durch eine extrem eingeschränkte Modellierung der Anwendungsdomäne aus und übersetzen Nutzeranfragen ohne interne Repräsentation und Verarbeitung direkt in Datenbankabfragen. Von den zwölf genannten Eigenschaften kommen nur sechs auf dieser Stufe zum Tragen: Typischerweise ist eine Grammatikprüfung nur rudimentär vorhanden, komplexe Nomen, Quantifikation und Ellipsis in der Anfrage werden zurückgewiesen, die Beantwortung von Metafragen ist nur begrenzt möglich, und die Hilfestellung im Fehlerfall ist antizipiert.

eL1: Schnittstellen der erweiterten ersten Stufe („extended L1“) zeichnen sich aus durch begrenzte Inferenzmöglichkeiten auf dem Datenbestand, die Verwendung von Pronomen und die bessere Übertragbarkeit auf andere Domänen.

L2: Schnittstellen dieser Stufe verfügen über eine explizite Theorie der Anwendungsdomäne.

L3: Schnittstellen der höchsten Stufe zeichnen sich durch explizite Theorien über die Benutzer, ihre Ziele und Pläne, sowie die Integration von Weltwissen aus.

Dekleva (1994) referiert die einschlägige Literatur (Jarke et al. (1985), Small und Weldon (1983), Hauptmann und Green (1983), Napier und Guadango (1989) u.a.) und kommt zu dem Schluß, daß Vergleichsstudien zwischen natürlichsprachlichen und konventionellen Schnittstellen überwiegend negativ für die natürlichsprachlichen Varianten ausfallen.

Dekleva selbst versucht mit seiner Untersuchung einer Schnittstelle – die ein typischer Vertreter der niedrigsten Kategorie L1 der Hendrix/Sethi-Skala ist – nachzuweisen, daß natürlichsprachliche Schnittstellen im Praxiseinsatz durchaus bestehen können. Dekleva untersucht dazu zwei Personengruppen, von denen die erste nicht darüber aufgeklärt war, daß die Anfragen, die sie auf Papier konstruierten, für ein Computerprogramm bestimmt waren. Die zweite Gruppe hingegen war darüber informiert, daß ihre Anfragen für eine natürlichsprachliche Schnittstelle gedacht waren – diese Gruppe wurde darüber hinaus explizit für den Umgang mit Datenbanken – konventionellen wie solchen mit natürlichsprachlichem Interface – trainiert.

Dekleva (1994) kommt auf extrem hohe Werte für die korrekte Interpretation. Er schließt daraus, daß die weitere Entwicklung natürlichsprachlicher Schnittstellen überaus vielversprechend sei, da bereits mit einem Vertreter der niedrigsten Kategorie L1 so gute Ergebnisse erzielt werden konnten. Allerdings erscheinen die Zahlen, zu denen

Dekleva kommt, nicht sehr zuverlässig zu sein, da er zahlreiche der ursprünglich 800 Anfragen aus verschiedenen Gründen als „nicht angemessen“ ausschließt.

Sicher scheint aber zu sein, daß Anfragen derjenigen Versuchspersonen, die darüber informiert waren, Eingaben für ein Computerprogramm zu generieren, und zudem mindestens ein minimales Training bekommen hatten, sehr viel bessere Ergebnisse lieferten als die der Vergleichsgruppe. Dekleva kommt hier auf 83.8 % korrekt interpretierte Anfragen insgesamt, sowie 90 % für die von ihm als „angemessen“ bezeichneten Anfragen.

Interessant ist ein integriertes Training für das System: Im Anschluß an die Auswertung der Anfrageverarbeitung wurde das Lexikon um bislang unbekannte Wörter aus den Anfragen erweitert, um es explizit auf die von den Versuchspersonen generierten Anfragen zu optimieren. Auf diese Weise konnte für die untrainierten, nicht aufgeklärten Versuchspersonen der Anteil der korrekt interpretierten Anfragen verdoppelt werden (von 23.1 % auf 53.8 % der Anfragen insgesamt). Für die Vergleichsgruppe ergab sich immerhin noch eine Steigerung von 83.8 % auf 92.2 % (Anfragen insgesamt).

Interessant ist auch die Verteilung der Verarbeitungsprobleme auf die Anfragen insgesamt: Neben den als „nicht angemessen“ bezeichneten Fragen, die den größten Anteil ausmachen, gibt es Probleme in Bezug auf die Interpretation von Ellipsen, hinsichtlich der lexikalischen Ambiguität, sowie der Ambiguität von Abkürzungen und bei der Auflösung von Pronomen. Weitere Probleme sind unbekannte Wörter und die Kombination zweier Fragen zu einer Systemanfrage. Wie nicht anders zu erwarten war, stellen die unbekanntes Wörter vor der Optimierung des Systems die größten Fehlerquellen dar, sind nach der Optimierung aber verschwunden. Daß die (untrainierte, nicht aufgeklärte) erste Gruppe von Versuchspersonen mit 19.8 % (von absolut 459 Anfragen) einen sehr viel höheren Anteil unbekannter Wörter (vor der lexikalischen Optimierung) hatte als die zweite Gruppe, ist nicht weiter verwunderlich, da die (trainierte und aufgeklärte) zweite Gruppe mit dem – bereits für die Anfragen der ersten Gruppe optimierten – System zu arbeiten begann.

Sehr hoch ist bei den untrainierten, nicht aufgeklärten Versuchspersonen der Anteil der strukturellen Probleme. Dabei handelt es sich aber eigentlich nicht um ein Problem der Schnittstelle, sondern um eines der darunterliegenden Datenbank, die z.B. die Verknüpfung zweier Fragen zu einer Systemanfrage nicht bewältigen kann.¹⁰⁰ Diese Probleme werden durch die Systemoptimierung um 80 % reduziert. Interessant ist auch der Anteil der Probleme mit Ellipsen: Mit 5 % vor der Optimierung, und immer noch 1.1 % nach der Optimierung, ist diese Fehlerquelle vergleichsweise gering.

Abschließend ist festzustellen, daß Dekleva (1994) – trotz einiger seltsamer Details in seiner Untersuchung – zu überzeugenden Ergebnissen kommt, die vielversprechend für die weitere Entwicklung von natürlichsprachlichen Schnittstellen sind. Wichtig

¹⁰⁰Allerdings wäre es wünschenswert, daß die Schnittstelle genau diesen Umstand erkennen und den Benutzer daraufhinweisen würde.

ist der enorme Einfluß der inkrementellen, domänenabhängigen Verbesserung des Lexikons. Auf die Bedeutung einer solchen Adaptionsphase weisen auch Ogden und Bernick (1997) hin. Beunruhigend erscheint in der Untersuchung von Dekleva (1994) der starke Einfluß von Training auf die Nutzer. Daß zudem die Aufklärung der Versuchspersonen darüber, Anfragen für eine natürlichsprachliche Schnittstelle zu generieren, großen Einfluß auf die Qualität der Anfragen hatte, spricht für die weitverbreitete Annahme, das Bewußtsein, mit einer Maschine zu kommunizieren, verändere den menschlichen Kommunikationsstil (s. 6.2.5).

Zusammenfassend läßt sich mit einer gewissen Vorsicht, angesichts der uneinheitlichen Literaturlage, festhalten, daß Benutzer von natürlichsprachlichen Schnittstellen hinsichtlich des Sprachumfanges und der Diskursphänomene offenbar eingeschränkte Eingaben tätigen. Dies ist überraschend, da die Entwickler natürlichsprachlicher Schnittstellen ja gerade davon ausgehen, daß Benutzer mit ihren Systemen einen natürlichen, kommunikativen Umgang pflegen sollen.

Zusammen mit den in 6.2.5 wiedergegebenen Erkenntnissen zur Steuerung des Eingabeverhaltens von Benutzern natürlichsprachlicher Schnittstellen durch Systemfeedback steht zu vermuten, daß natürlichsprachliche Schnittstellen von Benutzern deutlich anders wahrgenommen werden als menschliche Gesprächspartner. Zur Zeit ist aber unklar, ob dies nur an der noch nicht ausreichenden sprachlichen Kompetenz der Systeme oder an anderen Aspekten (z.B. ihrer Körperlosigkeit) liegt. Unter Umständen könnte sich herausstellen, daß der Ansatz vieler Forscher, eine natürlichsprachliche Schnittstelle einem menschlichen Gesprächspartner kommunikativ so ähnlich wie möglich zu machen, falsch ist, weil Menschen eventuell auch solche Programme kommunikativ immer noch anders behandeln als echte Menschen (vgl. Churcher et al. (1997) sowie Shneiderman und Maes (1997)). Von Bedeutung sind in diesem Zusammenhang die in 6.2.6 wiedergegebenen Erkenntnisse zu den Erwartungen von Benutzern an natürlichsprachliche Schnittstellen.

6.2.4 Wie einfach sind NL-Schnittstellen zu bedienen ?

Ein wesentlicher Anspruch der Entwickler natürlichsprachlicher Schnittstellen, ist die Bedienbarkeit dieser Schnittstellen für Nutzer ohne Vorkenntnisse und ohne ein Training. Dieser Anspruch läßt sich aber offenbar nicht halten, wenn man die einschlägige Literatur betrachtet.

Ogden und Bernick kommen in ihrem Übersichtsartikel zu natürlichsprachlichen Schnittstellen im „Handbook of Human-Computer Interaction“ nach Auswertung verschiedenster Laborexperimente, Evaluationen und Feldstudien zu dem Schluß, daß für natürlichsprachliche Schnittstellen, trotz aller anderslautender Ansprüche der Entwickler, dennoch Training notwendig sei. Capindale und Crawford (1990) schreiben, daß alle Teilnehmer ihrer Untersuchung trainiert werden mußten und daß dennoch diejenigen am besten abschnitten, die mit den Strukturen der zugrundeliegenden Datenbank vertraut waren.

Auch Tennant (1979) und Dekleva (1994) kommen zu dem Schluß, daß die Benutzer natürlichsprachlicher Schnittstellen trainiert werden müssen, um erfolgreich mit dem System umgehen zu können. Dekleva (1994) (s. S. 232) stellt in einem Papierexperiment fest, daß Versuchspersonen, die trainiert und darüber informiert worden waren, Anfragen für ein Computerprogramm zu generieren, über 50 % mehr korrekt interpretierbare Anfragen stellen als eine untrainierte Vergleichsgruppe ohne Aufklärung über den Einsatz der Anfragen für den Computer. Laut Tennant (1979) liegt dies vor allem daran, daß die Benutzer sich mit der Struktur der Daten vertraut machen müssen. Insbesondere haben Benutzer Schwierigkeiten, innerhalb der konzeptuellen Grenzen der natürlichsprachlichen Schnittstelle zu bleiben.

Schwierigkeiten entstehen auch, wenn Personen, die keine Erfahrung im Umgang mit dem System haben, Aufgaben für andere Benutzer generieren. Die Tatsache, daß diese Aufgaben oft nicht lösbar sind, zeigt laut Ogden und Bernick (1997), daß das untersuchte System konzeptuell unvollständig sei, d.h. nicht die gesamte Menge der aus Sicht des Menschen für die Schnittstelle nötigen Konzepte umfaßt (vgl. S. 222). Dieses Argument erscheint schwach, denn auch Menschen wissen in zwischenmenschlicher Kommunikation nicht immer alles. Das Ziel einer natürlichsprachlichen Schnittstelle sollte es sein, sprachlich die aufgabenbezogenen Anforderungen des Benutzers gut bewältigen zu können, aber nicht jeder Idee eines Menschen gewachsen zu sein.

Biermann et al. (1983) untersuchen (ähnlich wie Miller (1981), s. S. 238) die Möglichkeiten, Englisch als Programmiersprache einzusetzen. Ogden und Bernick (1997) interpretieren die Ergebnisse von Biermann et al. (1983) so, daß Benutzer Schwierigkeiten hätten, innerhalb der funktionalen, syntaktischen und lexikalischen Domäne der NL-Schnittstelle zu bleiben. Lediglich die konzeptuelle Domäne werde eingehalten, da innerhalb des von Biermann et al. (1983) untersuchten Systems Referenzen nur auf die am Bildschirm dargestellten Objekte möglich seien. Zudem werden nur von den Systemdesignern entworfene Aufgaben verwendet.

Die Interpretation der Untersuchungen von Tennant (1979) und Biermann et al. (1983) durch Ogden und Bernick befremdet, weil das Alter der Untersuchung und die damit verbundenen Systemgrenzen nicht in Betracht gezogen werden. Bei einem Lexikon mit nur 350 Einträgen (Biermann et al., 1983) erscheint es aus heutiger Sicht erstaunlich, wie gut das System abschneidet. Biermann et al. (1983) ziehen, im Gegensatz zu der eher negativen Aussage von Ogden und Bernick, am Ende auch ein neutrales Fazit, wenn sie schreiben, daß keines der Standardprobleme natürlicher Sprache wie Vagheit, Ambiguität usw. ein Problem dargestellt habe.¹⁰¹

Wie viele andere Systeme auch, ist das von Biermann et al. (1983) untersuchte allein durch die Entwicklung der graphischen Oberflächen obsolet geworden. Um im Stile einer Tabellenkalkulation Zahlen in Spalten einzutragen oder Feldern arithmetische Funktionen zuzuweisen, würde heute erfolgreich die Maus an einer graphischen Oberfläche eingesetzt.

¹⁰¹(Biermann et al., 1983, :71).

Jarke et al. (1985) kommen in ihrer Untersuchung einer beschränkt natürlichsprachlichen Schnittstelle zu dem Schluß, daß Benutzer durchaus Training brauchen. Allerdings stellen sie auch fest, daß die Lernrate der Versuchspersonen, die SQL als Anfragesprache verwendeten, ohne individuelle Unterschiede konstant war, während die Lernrate derjenigen, die eingeschränkt natürlichsprachliche Anfragen stellten, bei großen individuellen Unterschieden sehr stark war.¹⁰² Jarke et al. weisen hinsichtlich der individuellen Unterschiede darauf hin, daß die Nutzer einer natürlichsprachlichen Schnittstelle in zwei Extreme zerfallen und entweder verstehen, wie eine natürlichsprachliche Schnittstelle funktioniert, oder aber größte Schwierigkeiten im Umgang mit einer solchen Schnittstelle haben. Unter Umständen müsse daher, je nach Nutzer, mit einer sehr langen Adaptionzeit für natürlichsprachliche Systeme gerechnet werden. Jarke et al. haben keine Erkenntnisse zu den möglichen Ursachen dieser starken individuellen Unterschiede.

Ogden und Sorknes (1987) stellen in einem Laborexperiment fest, daß ein (nicht namentlich genanntes) kommerzielles, natürlichsprachliches Datenbankinterface für Nutzer ohne Kenntnis der Struktur der zugrundeliegenden Datenbank nur sehr schwer zu benutzen ist. Nur 28 % der Anfragen waren im ersten Versuch erfolgreich, und es waren im Schnitt 3.6 Versuche nötig, um zu einem Ergebnis zu gelangen. 16 % der dabei aufgetretenen Fehler blieben unbemerkt. Insgesamt wurden überhaupt nur 72 % der Aufgaben korrekt gelöst.

Allerdings zeigen die Beispiele, die Ogden und Sorknes für die Schwierigkeiten der Nutzer geben, daß die Schwierigkeiten in der Bedienung nicht grundsätzlich für natürlichsprachliche Schnittstellen gültig sein müssen, sondern zu einem Teil auch in der unzureichenden Leistungsfähigkeit des untersuchten Systems begründet sein können. So bemängeln Ogden und Sorknes, daß das System unterschiedliche Antworten auf die Frage „How many credits does David Lee have?“ und „What are the total credits for David Lee?“ gibt. Im ersten Fall wird die Anzahl der Kurse ausgegeben, die David Lee besucht hat und im zweiten Fall die Summe der insgesamt erworbenen Punkte. Wie in 4.2.13 bereits angesprochen, ist Ambiguität eine der wesentlichen Eigenschaften natürlicher Sprache.¹⁰³ Natürlichsprachliche Schnittstellen müssen daher ganz besonders auf ambige Eingaben eingehen, u.a. indem sie dem Benutzer explizite Rückmeldung geben, welche Lesart der Interpretation der Anfrage zugrunde liegt. Daß dies im allgemeinen Fall schwer zu lösen ist, wurde in 4.2.13 bereits diskutiert. Eine Ausgabe wie die des von Ogden und Sorknes getesteten Systems (`count: 2` vs. `total credits: 7`) ist aber sicher unzureichend und kann nicht als Schwäche der natürlichsprachlichen Eingabe an sich gewertet werden.

Der Schluß, den Ogden und Sorknes aus ihrer Untersuchung ziehen, daß nämlich Benutzer natürlichsprachlicher Schnittstellen Training brauchen, überzeugt daher nicht

¹⁰²Von unter 5 % korrekten Anfragen zu Beginn der Interaktion auf über 30 % im weiteren Verlauf.

¹⁰³Und die Tatsache, daß dies menschlichen Sprechern (im Gegensatz zur Maschine) zumeist überhaupt keine Schwierigkeiten bereitet, ist eine der wesentlichen Eigenschaften menschlichen Sprachverstehens.

völlig. Das gleiche gilt für die Feststellung von Ogden und Sorknes, daß Wissen über die Struktur des Datenbestandes Voraussetzung für die erfolgreiche Bedienung einer natürlichsprachlichen Schnittstelle sei: Wie schon die Aussage zur Notwendigkeit eines Trainings für Benutzer natürlichsprachlicher Schnittstellen, scheint auch diese Feststellung nicht zwingend aus den Ergebnissen des Experimentes von Ogden und Sorknes zu folgen.

Angesichts der hochgradig interpretationsbedürftigen Systemantworten (s.o.) kann der von Ogden und Sorknes registrierte häufige Blick der Versuchspersonen auf die Struktur der Datenbank auch als Versuch einer Interpretation der Systemantwort gewertet werden. Ogden und Sorknes schlagen als Resultat ihrer Untersuchung vor, natürlichsprachliche Schnittstellen dahingehend zu verbessern, daß sie es dem Benutzer ermöglichen, ein konsistentes Modell der dem System zugrundeliegenden Domäne zu entwickeln, und ihm zudem explizite und verständliche Rückmeldung zu geben. Ob die Benutzer solcher Systeme aber tatsächlich ein Training benötigen, wie die Autoren betonen, ist zumindest auf der Grundlage der Untersuchung von Ogden und Sorknes (1987) nicht zu entscheiden.

Auch Miller (1981) kommt zu dem Ergebnis, daß natürliche Sprache keineswegs für einfachere Bedienbarkeit sorgen muß, und macht Vorschläge für die Begrenzung ihres Einsatzes. Miller untersucht die Möglichkeiten (unbeschränkte), natürliche Sprache als Programmiersprache einsetzen zu können, und stellt in einem Versuch, bei dem computerunerfahrene Versuchspersonen einfache Handlungsanweisungen (Algorithmen) für andere Menschen schreiben sollen, spezifische Schwierigkeiten eines natürlichsprachlichen Ansatzes fest.

Neben der Tendenz der Nutzer, Weltwissen vom System zu erwarten (s. S.244), stellt Miller die häufige Verwendung von Präsuppositionen und Anaphern, einen eingeschränkten Wortschatz und zugleich viele Synonyme fest (s. S.241). Miller kommt zu dem Schluß, daß natürlichsprachliche Schnittstellen nur mit Beschränkungen und in bestimmten Einsatzgebieten Sinn machen. Die Einschränkungen, die Miller (im Szenario seiner Untersuchung) vorschlägt, sind:

- Beschränkung der vom Interface unterstützten Anwendungsfunktionalität auf Expertenniveau, damit nur solche Benutzer das natürlichsprachliche Interface verwenden, die ausreichend Wissen über die Anwendungsdomäne haben.
- Extreme Beschränkungen im Bereich des Vokabulars durch einerseits definitiven Ausschluß jeder Art von Ambiguität und andererseits eine Beschränkung der Gesamtmenge des verfügbaren Vokabulars.
- Moderate Beschränkungen der Syntax, in Hinblick auf Satzlänge, Anzahl der Modifikationen pro Nomen, Ausschluß komplexerer Konstruktionen usw.

Napier und Guadango (1989) zeigen in einem Vergleichstest klare Vorteile hinsichtlich Lernbarkeit und Akzeptanz für eine beschränkt natürlichsprachliche Schnittstelle, bei

Nutzern ohne Vorkenntnisse. Napier und Guadango vergleichen zwei Versionen des Tabellenkalkulationsprogrammes „Lotus 1-2-3“, die herkömmliche Version und eine Version mit beschränkt natürlichsprachlicher Schnittstelle („Lotus HAL“).

Die Akzeptanz¹⁰⁴ der natürlichsprachlichen Version war signifikant höher als die der herkömmlichen Version. Die im Rahmen des Versuches zu lösenden Aufgaben wurden von 70 % der Nutzer der natürlichsprachlichen Schnittstelle erfolgreich gelöst, während nur 36 % der Nutzer der herkömmlichen Version Erfolg hatten. Allerdings ist fraglich, ob der Effekt tatsächlich auf die Verwendung natürlicher Sprache zurückzuführen ist. Die natürlichsprachliche Version „Lotus HAL“ erlaubt, im Gegensatz zur herkömmlichen Version, die Verwendung von Kontextinformation, d.h. daß der Wirkungsbereich einer Formel nicht explizit für alle gewünschten Zellen angegeben werden muß, sondern vom Programm aus dem Kontext erschlossen wird. So ist es möglich, daß ein Benutzer die erste Zelle einer Spalte markiert, und für diese ENTER 1000 THEN PROJECT AT 12 % ACROSS angibt – das System ist dann in der Lage, aus dem Kontext, den mit „across“ gemeinten Wirkungsbereich zu erschließen. Dies ist, verglichen mit dem Aufwand, der für die herkömmliche Version „Lotus 1-2-3“ zu treiben ist, eine enorme Erleichterung. Inwieweit damit auch erfahrenen Nutzern von Tabellenkalkulationsprogrammen geholfen wird und ob sich bei dieser Gruppe andere Ergebnisse für die Akzeptanz der natürlichsprachlichen Schnittstelle ergeben, bleibt offen.

Das Fazit von Napier und Guadango lautet entsprechend: „Our research indicates that there are circumstances in which a restricted natural language interface can be beneficial. [...] the present research does not shed light on exactly *why* the performance with the restricted natural language interface was superior.“(Napier und Guadango, 1989, :1195)¹⁰⁵

Die Untersuchung von Walker und Whittaker (1990) (s. S. 227) kam zu einem ähnlichen Ergebnis, daß einerseits das Verhalten der Nutzer eine deutliche Vorliebe für die natürlichsprachliche Schnittstelle erkennen ließ, dies aber, mindestens teilweise, auf die erweiterten Möglichkeiten des natürlichsprachlichen Systems zurückzuführen war.

Anläßlich der Evaluation von OSIRIS kommt Moranz (2000) zu dem Schluß:

„Die [...] weit verbreiteten Vorstellungen bezüglich der Einfachheit und unmittelbaren Verständlichkeit von NL-Systemen konnte auch diese Untersuchung nicht bestätigen. Die Tatsache allein, daß es sich um ein natürlichsprachliches System handelt und daß infolgedessen auch 'natürliche' Eingaben möglich sind, reicht zur adäquaten Bedienung solcher Systeme nicht aus. Vielmehr muß auch das NL-System seine 'Angebote' vermitteln und deutlich machen, was möglich ist und was nicht.“(Moranz, 2000, :87f.)

¹⁰⁴ Antworten der Versuchspersonen auf die Fragen, wie ihnen das Produkt gefallen hat und ob es geeignet für ihre Bedürfnisse war.

¹⁰⁵ Hervorhebung im Original.

Bei aller (bereits mehrfach angemahnten) Vorsicht gegenüber der Literaturlage scheint sich dennoch das Fazit, das Moranz (2000) anlässlich der OSIRIS Evaluation zieht, zu bestätigen. Natürlichsprachliche Schnittstellen bereiten den Benutzern offenbar mehr Schwierigkeiten, als dies nach den Vorstellungen ihrer Entwickler möglich sein sollte, da natürliche Sprache an sich ja für einfache Bedienbarkeit sorgen soll. Zwei Begründungen sind für diese Situation möglich:

- Einerseits könnten natürlichsprachliche Schnittstellen einfach noch nicht den Stand erreicht haben, den andere Oberflächen hinsichtlich ihrer Bedienbarkeit auch erst im Laufe der Zeit durch kontinuierliche Benutzerrückmeldungen erlangen haben. Wenn natürlichsprachlichen Schnittstellen auch noch eine größere Komplexität zugebilligt werden kann, wäre dies eine Erklärung für die beobachtete Situation. Diese Erklärung würde zugleich auch zu der Erwartung Anlaß geben, daß durch weitere Entwicklungsarbeit, über kurz oder lang, natürlichsprachliche Schnittstellen mit herkömmlichen Schnittstellen in Hinblick auf ihre Bedienbarkeit gleichziehen und von diesem Zeitpunkt an die spezifischen Vorteile natürlicher Sprache als Anfragesprache zu tragen kommen.
- Andererseits könnten die beobachteten Schwierigkeiten ein spezifisches Problem natürlichsprachlicher Schnittstellen sein. Dadurch aber könnte das Konzept einer solchen Schnittstelle in Frage gestellt werden.

6.2.5 Erleichterung und Steuerung der Bedienung

Gustafson (1997) thematisiert das Problem, daß Menschen an natürlichsprachlichen Schnittstellen eine große Variabilität in der Ausdrucksweise zeigen. Im Gegensatz zu Menü-Systemen oder Kommandozeilensystemen mit festem, begrenzten Vokabular, können Benutzer einer NL-Schnittstelle ein und denselben Sachverhalt mit verschiedenen syntaktischen Konstruktionen und unterschiedlichem Vokabular ausdrücken. Der letztere Punkt, die lexikalische Variabilität, könnte im Mensch-Maschine-Dialog durch die Systemgestaltung beeinflußt werden, da die Teilnehmer eines zwischenmenschlichen Dialoges sich offenbar ihrer gemeinsamen Überzeugungen durch übereinstimmendes Vokabular vergewissern (Gustafson, 1997). Gustafson referiert dazu die einschlägige Literatur und stellt die Ergebnisse eines eigenen Wizard-of-Oz Experimentes vor. Die Ergebnisse dieses Experimentes bestätigen die zitierten Arbeiten, d.h. daß Benutzer einer natürlichsprachlichen Schnittstelle das vom System verwendete Vokabular aufgreifen.

Chin (1984) untersucht die Logfiles eines Wizard-of-Oz Experimentes mit Versuchspersonen, die über Computervorerfahrung verfügten, und vergleicht diese Ergebnisse mit den Transkriptionen einer Kontrollgruppe, die dieselben Aufgaben im direkten Gespräch mit menschlichen Experten zu lösen hatten. Chin stellt fest, daß die Teilnehmer des Wizard-of-Oz Experimentes in über 25 % der Eingaben mit Ellipsen, Anaphern, indirekten Sprechakten oder unvollständigen Sätzen auf den Kontext Bezug nehmen.

Die Transkriptionen der Kontrollgruppe zeigen dagegen doppelt so häufig kontextabhängige Konstruktionen. Daraus schließt Chin, daß der reduzierte Anteil an Kontextbezügen in den linguistischen Konstruktionen der Interaktionsprotokolle den Versuch der Teilnehmer demonstriert, sich im Sprachgebrauch der Maschine anzupassen.

Miller (1981) stellt in einem Papiertest-Versuch fest, daß, obwohl die natürlichsprachlichen Ausdrucksmöglichkeiten der Versuchspersonen im Versuch in keinerlei Hinsicht beschränkt wurden, dennoch nur ein beschränkter Wortschatz verwendet wurde, der sich durch eine große Anzahl von Synonymen auszeichnete.

Churcher et al. (1997) zitieren Guindon (1988) und Kennedy et al. (1988), die beide ebenfalls zu dem Schluß kommen, daß menschliche Äußerungen kürzer sind und eine geringere lexikalische Variationsbreite aufweisen, wenn sie an Computer gerichtet sind. Auch Krause (1980) geht davon aus, daß Benutzer sich in ihrem sprachlichen Verhalten den Möglichkeiten des Computers¹⁰⁶ anpassen:

„Man könnte auf den Gedanken kommen, das fehlende Wissen [über die natürlichsprachliche Mensch-Maschine-Kommunikation] durch eine Übertragung der Kenntnisse über den Mensch-Mensch-Dialog zu kompensieren, d.h. also, menschliche Kommunikation zu simulieren. Es gibt jedoch einige Hinweise dafür, daß dies nicht möglich ist, zumindest nicht pauschal. Der Benutzer weiß, daß er mit einer Maschine kommuniziert; er hat eine bestimmte Einstellung der Maschine gegenüber und seine Vermutungen, wie die Maschine reagieren wird.“ Krause (1980, :205)

Krause interpretiert entsprechend ungrammatische Eingaben der Nutzer als deren Versuch, sich einem (als abweichend empfundenen) Sprachgebrauch der Maschine anzugleichen.

Zoltan-Ford (1984) führt ein Wizard-of-Oz Experiment mit zwei Systemvarianten durch: Variante eins erlaubt unbeschränkte, natürlichsprachliche Eingaben, und Variante zwei beschränkt die möglichen Nutzereingaben (ohne Wissen der Nutzer) auf genau die Konstruktionen, die auch vom System selbst in den Antworten an den Nutzer verwendet werden. Zudem variiert Zoltan-Ford systematisch die syntaktische Komplexität der Systemreaktionen.

Dabei zeigt sich, daß signifikant mehr Nutzer der restringierten Version die Systemreaktionen imitieren (s.a. Dybkjær et al. (1995)). Unter diesen wiederum paßten sich diejenigen besonders gut den Systemausgaben an, die syntaktisch einfache Ausgaben erhielten. Als Ergebnis hält sie fest, daß die Variabilität in den Benutzereingaben deutlich reduziert werden kann, wenn die Systemausgaben sorgfältig und konsistent formuliert werden, da die Nutzer diese imitieren würden. Dabei werden die Ausgaben des Systems umso stärker imitiert, je einfacher der Satzbau ist. Dies macht, insbesondere in Bezug auf die Fehlermeldungen, Sinn, wenn nämlich dem Benutzer nicht

¹⁰⁶Oder auch: ihren Vorstellungen der Möglichkeiten des Computers.

nur erklärt wird, daß ein Fehler aufgetreten ist, und z.B. seine Eingabe nicht analysiert werden konnte, sondern ihm durch die Gestaltung der Fehlermeldung bereits die Systemfähigkeiten demonstriert werden.

Ogden und Sorknes (1987) zitieren eine Reihe an Untersuchungen, die zu dem Ergebnis kommen, daß Menschen sich bemühen, sich syntaktisch besonders einfach auszudrücken, wenn sie glauben, mit einer Maschine über eine natürlichsprachliche Schnittstelle zu kommunizieren. Diese Ergebnisse werden durch Ogden und Brooks (1983) und Burton und Steward (1993) ebenfalls bestätigt.

Capindale und Crawford (1990) kommen zu ähnlichen Ergebnissen: Viele der Befragten gaben an, das Systemecho ihrer Eingabe als Vorlage für die Formulierung nachfolgender Anfragen zu benutzen.

Jarke et al. (1985) stellen als wichtigste Beobachtung ihrer Untersuchung des Nutzerverhaltens an einer beschränkt natürlichsprachlichen Schnittstelle fest, daß die Rückmeldungen des Systems an den Benutzer entscheidend für die Performanz der Benutzer sind (ebenso Turner et al. (1984) und Chapman (1981)).

Auch Slator et al. (1986) führen ein Wizard-of-Oz Experiment mit zwei Gruppen durch, wovon die eine mehr und bearbeitetes Feedback erhält und die Kontrollgruppe lediglich sporadisch normale Systemmeldungen. Slator et al. stellen fest, daß sorgfältig bearbeitete Systemmeldungen den Benutzer in seinem Eingabeverhalten erziehen. Wenn dieses Mittel konsequent eingesetzt werde, sei dies eine gute Möglichkeit, die Mensch-Maschine-Interaktion zu beeinflussen.

Abschließend muß, wie schon im Fazit für 6.2.4, festgestellt werden, daß die Erwartungen der Entwickler natürlichsprachlicher Schnittstellen bezüglich der Wirkungsweise ihrer Systeme unzutreffend zu sein scheinen. Offenbar empfinden die Benutzer natürlichsprachlicher Schnittstellen nicht das große Freiheitsgefühl, nun endlich mit der Maschine so kommunizieren zu können wie mit dem menschlichen Gegenüber. Damit ist aber die grundsätzliche Annahme, Menschen würden an natürlichsprachlichen Schnittstellen „unverfälscht“ kommunizieren, zumindest für die aktuelle Schnittstellengeneration hinfällig. Ob dies am technischen Stand der Systeme liegt, oder aber andere Gründe hat, ist noch zu klären.

6.2.6 Benutzerverhalten an natürlichsprachlichen Schnittstellen

In diesem letzten Abschnitt zur Untersuchung der Beurteilung natürlichsprachlicher Schnittstellen in der Literatur geht es um die Beobachtungen, die zum Verhalten von Menschen an diesen Schnittstellen gemacht wurden.

Wie bereits in 6.2.2 dargestellt, vergleichen Bell und Rowe (1992) drei verschiedene Datenbank-Schnittstellen miteinander: eine graphische Oberfläche für SQL, eine SQL-basierte Schnittstelle und eine natürlichsprachliche Schnittstelle. Sie kommen zu dem Ergebnis, daß die natürlichsprachliche Schnittstelle sich qualitativ anders verhält als die beiden anderen, einander sehr ähnlichen Schnittstellen (ausführlicher auf S. 230).

Burton und Steward (1993) (ausführlicher auf S. 231) und Ogden und Brooks (1983) weisen die häufig geübte Kritik, natürlichsprachliche Schnittstellen seien besonders aufwendig zu bedienen, weil sie sehr viel Tipparbeit erfordern würden, zurück. Wenn der Kontext einer Äußerung für die weitere Interaktion verwendet werden kann, können aufeinanderfolgende Anfragen sehr knapp gehalten werden. Benutzer natürlichsprachlicher Schnittstellen benötigen dann nicht mehr Zeit als die Benutzer herkömmlicher Schnittstellen. In der Untersuchung von Jarke et al. (1985) (s. S. 225) sind die Benutzer natürlicher Sprache gegenüber denen, die SQL verwenden, sogar deutlich im Vorteil.

Hendrix schreibt über die Erwartungshaltung der Nutzer gegenüber einer natürlichsprachlichen Schnittstelle:

„Perhaps the main disadvantage of NL systems ist that people tend to assume they are 'smart'. For example, if a system can provide NL access to a data base of information, users will tend to believe it can deduce other facts from that information - facts that, although not explicitly encoded, would be obvious to anyone with common sense. More formal systems are not expected to perform common sense reasoning, because their functionality, and therefore their inherent limitations, is readily apparent to the user.“(Hendrix, 1982, :58)

Slator et al. (1986) sehen es ebenfalls als ein Problem natürlichsprachlicher Schnittstellen, daß deren Nutzer offensichtlich Weltwissen und Inferenzen über das Domänenwissen erwarten. Diese Erwartungen seien so hoch, daß sie auch von den anspruchsvollsten Systemen¹⁰⁷ nicht zu befriedigen seien.

Dekleva (1994) (s. S. 232) bestätigt diese Einschätzung, indem er an seinen Versuchspersonen eine Tendenz zu überhöhten Erwartungen an das System feststellt. Die Versuchspersonen erwarteten nicht nur Einträge aus der zugrundeliegenden Datenbank, sondern auch Beratungsleistungen, die Inferenzen über der Domäne erfordert hätten. Auch Capindale und Crawford (1990) erkennen diese überhöhten Erwartungen, und erklären sie durch die im Vergleich enormen Fähigkeiten der Anfragesprache:

„Yet, it is the flexibility of the language which entices people to expect much functionality.“(Capindale und Crawford, 1990, :360)

Ob dies allerdings für alle neuen, flexiblen Eingabemöglichkeiten, oder nur für natürliche Sprache gelten soll, lassen Capindale und Crawford offen.

Copestake und Sparck-Jones argumentieren ähnlich, sind in ihrer Begründung aber präziser. Für sie liegt das spezifische Problem natürlichsprachlicher Schnittstellen, überhöhte Erwartungen zu induzieren, in der Modellbildung der Nutzer begründet:

¹⁰⁷Stand 1986.

„ ... the obvious way for a user to model something that accepts natural language input is as though it were an intelligent, rational agent.“(Copestake und Sparck-Jones, 1990, :227)

Auch wenn diese Begründung auf den ersten Blick einleuchtend erscheint, so steht sie doch partiell in Widerspruch zu den in 6.2.5 referierten Erkenntnissen, daß Nutzer sehr genau wissen, ob sie es mit einem Menschen oder einer Maschine zu tun haben. Wenn aber die Nutzer natürlichsprachlicher Systeme sich in ihrem Eingabeverhalten auf die Maschine einstellen, dann sollten sie auch eine andere Erwartungshaltung gegenüber der Maschine haben als gegenüber einem menschlichen Gesprächspartner.

Miller (1981) schreibt, daß jeder Versuch, natürlichsprachliche Schnittstellen nur durch eine Erweiterung der syntaktischen Eingabemöglichkeiten zu verbessern, zu kurz greift. Eine solche syntaktische Erweiterung müsse stets von einer Erweiterung der Semantik und des Weltwissens begleitet sein. Die Möglichkeiten, ein System um Weltwissen zu erweitern, sind aber z.Zt. beschränkt.¹⁰⁸ Copestake und Sparck-Jones warnen zu Recht davor, in Hinblick auf das Weltwissen Systemerweiterungen ad hoc vorzunehmen: „The difficulty is that to expand the front end so that it could answer some of these questions [e.g. answer anticipated questions for advise] would give a misleading impression to the user about the capabilities of the system as a whole.“ Copestake und Sparck-Jones (1990, :230)

Shneiderman (1978) problematisiert das Ziel, natürlichsprachliche Schnittstellen zu entwickeln, deren Nutzer von der Struktur des zugrundeliegenden Datenbestandes nichts mehr wissen müssen. Dies könnte, nach Shneidermans Ansicht, beim Nutzer zu unrealistischen Erwartungen in Bezug auf die Fähigkeiten des Systems führen. Androutsopoulos et al. (1995) bestätigen in ihrem Literaturüberblick die Tendenz der Nutzer, aus der Möglichkeit eines Systems, natürliche Sprache zu verarbeiten, auf dessen Fähigkeit zu schliessen, mit Hilfe von Weltwissen zu rasonieren.

Diaper (1986) stellt in den Mensch-Maschine-Dialogen seines Wizard-of-Oz-Experimentes einen hohen Anteil an Weltwissen fest. Ziel der Untersuchung von Diaper war es, für die auf Seiten des Systems benötigten Wissensquellen (Domänenwissen, linguistisches Wissen, Weltwissen) den Umfang und die Voraussetzungen genauer beschreiben zu können, um die Entwicklung von natürlichsprachlichen Dialogschnittstellen insgesamt voranzubringen. In der Durchführung überzeugt das Experiment allerdings nicht: Auf Befragen erklärten die Versuchspersonen im Anschluß an die Untersuchung, sie seien aufgrund von Tippfehlern in der Systemausgabe sowie der deutlichen Verzögerung bei der Ausgabe von Tabellen nicht immer sicher gewesen, wirklich mit einer Maschine zu kommunizieren. Auch erscheint die Analyse der Dialoge in Form einfacher Adjazenz-Paare, -Tripel und -Quadrupel sowie deren Zuordnung zu einer endlichen Liste vorgefertigter Themen nicht tiefgehend genug.

Interessant, wenn auch vor dem Hintergrund der eben genannten Kritik mit Vorsicht zu

¹⁰⁸Zu den Problemen, s. 4.1.

betrachten, ist der Anteil der verschiedenen Wissensarten am Dialog. Im Schnitt wurde 37.2 mal pro Dialog¹⁰⁹ das vom eigentlichen Expertensystem angebotene Wissen benötigt. 40.2 mal pro Dialog wurde explizit auf im Verlauf des Dialoges eingeführtes Wissen Bezug genommen, und immerhin 19.4 mal pro Dialog wurde Weltwissen benötigt. 91.7 % dieses Weltwissens stand dabei in direktem Zusammenhang mit den Systembenutzern und deren Eigenschaften als menschliche Wesen.

Angesichts des hohen Anteils an Weltwissen in den (simulierten) Mensch-Maschine-Dialogen stellt sich die Frage, ob die Wirkung der Schnittstelle auf den Benutzer durch Systemeigenschaften beeinflusst wird, die normalerweise als menschliche Eigenschaften betrachtet werden. Die Wahrnehmung solcher (anthropomorpher) Eigenschaften (z.B. Sprache) könnte Benutzer dazu verleiten, weitere Fähigkeiten – z.B. Schlußfolgerungen ziehen zu können – vom System zu erwarten (Maass, 1983).

Diese Annahme wird gestützt von Untersuchungen, die Clifford Nass seit dem Anfang der 90er Jahre in Stanford durchführt. Nass et al. (1993) und Nass et al. (1994) berichten von einem Experiment, in dem Versuchspersonen gegenüber Computern ein Verhalten zeigen, das normalerweise nur Menschen gegenüber angemessen erscheint. Hatte die Maschine eine an Menschen erinnernde Eigenschaft wie z.B. Sprachausgabe,¹¹⁰ bewerteten die Versuchspersonen das Verhalten der Maschinen nach sozialen Normen wie „Eine andere Person zu loben ist besser, als sich selbst zu loben“ und „Kritik an anderen ist weniger freundlich als Kritik an sich selbst“. Zudem wurden unterschiedliche Stimmen als unterschiedliche Agenten betrachtet. Dabei war allen Versuchspersonen vollkommen klar, mit einer Maschine zu arbeiten, und alle Benutzer gaben in begleitenden Fragebögen an, daß sie es unangemessen finden, Maschinen nach menschlichen Maßstäben zu behandeln.

Nass et al. (1997) zeigen wiederum, daß die Interaktion zwischen Mensch und Maschine sozialen Regeln gehorcht und daß dies nicht auf Mißverständnisse im Experiment oder Fehlannahmen der Versuchspersonen über die Natur von Computern zurückzuführen ist. Durch die systematische Verwendung männlicher und weiblicher menschlicher Stimmen¹¹¹ konnten Nass et al. (1997) zeigen, daß die Versuchspersonen gegenüber der Maschine geschlechtsspezifische Stereotypen anwendeten, wie bspw. „Eine Beurteilung durch einen Mann ist zuverlässiger als eine Beurteilung durch eine Frau“ oder „Dominantes Verhalten ist für Männer angemessener als für Frauen“.

Offenbar werden Computer also unter bestimmten Voraussetzungen (z.B. bei Anwesenheit von Sprache) automatisch, ungewollt und systematisch von Menschen als soziale Wesen wahrgenommen und behandelt.

De Angeli et al. (1999) bestätigen die Ergebnisse von Nass et al. durch eigene Experimente und weisen nachdrücklich darauf hin, daß in den Augen der Versuchspersonen

¹⁰⁹Ein Dialog dauerte im Schnitt 40 Minuten und wurde durchschnittlich als aus 165 Segmenten bestehend analysiert.

¹¹⁰Mit Hilfe qualitativ hochwertiger Aufzeichnungen menschlicher Stimmen.

¹¹¹Bei sorgfältigem Ausschluß anderer Hinweise auf ein Geschlecht oder ein menschliches Bewußtsein auf Seiten der Maschine.

ganz offenbar keine Verwechslung zwischen Mensch und Maschine stattfindet: Auch gegenüber solchen Systemen, die (beabsichtigt) menschliche Eigenschaften imitieren, verwenden Benutzer nur ein eingeschränktes Vokabular und einfache Strukturen, wie dies bereits für ältere natürlichsprachliche Schnittstellen festgestellt wurde (s. 6.2.5). De Angeli et al. kommen daher zu dem Schluß, daß Computer für den menschlichen Benutzer soziale Objekte seien. Sie unterscheiden sich vom Menschen und werden von Menschen auch so unterschiedlich wahrgenommen, nehmen aber eine Sonderstellung zwischen Werkzeugen oder Objekten einerseits und belebten Wesen andererseits ein. Daß diese soziale Rolle des Computers nicht abwegig ist, bestätigen schon Geser (1989) und Hartmann (1992). De Angeli und Petrelli (2000) fordern daher, daß die Forschung im Bereich Human-Computer-Interfaces (HCI) die Annahme aufgebe, der Computer habe für die Benutzer den Charakter eines Werkzeuges. Ansonsten seien echte Verbesserungen der Nutzerzufriedenheit und der Akzeptanz der Schnittstellen nicht zu erreichen.

Wichtig ist aber der Hinweis in De Angeli et al. (1999) und De Angeli und Petrelli (2000) auf die möglichen Ursachen der Wahrnehmung des Computers als soziales Wesen. Die schon von Nass et al. angesprochenen menschlichen Eigenschaften des Systems, die der Auslöser für die Behandlung der Maschine nach sozialen Normen sein sollen, sind in mancher Hinsicht noch unverstanden (s. 7.2). Jedoch scheint, aufgrund der Erkenntnisse von Nass et al., De Angeli et al. (1999) sowie De Angeli und Petrelli (2000) sicher zu sein, daß natürliche Sprache nicht der alles bestimmende Faktor ist. Obwohl Sprache ganz offenbar die beschriebenen Reaktionen der menschlichen Benutzer verursachen kann (s. Nass et al. (1994)), sind dieselben Effekte aber auch in Umgebungen ohne Sprache nachzuweisen. In 7.2 wird auf dieses Thema weiter eingegangen.

6.2.7 Fazit

Die zitierten Untersuchungen konkreter Systeme sind nur schwer oder gar nicht zu vergleichen, da sie überwiegend alt sind und dadurch beträchtliche Unterschiede im Entwicklungsstand aufweisen.

Daher ist es nur mit einer gewissen Vorsicht möglich, folgende Schlüsse zu ziehen:

Natürlichsprachliche Schnittstellen schneiden im Vergleich zu SQL- oder anderen Schnittstellen nicht so erfolgreich ab, daß ihre bloße Existenz einen positiven Effekt auf die Nutzer haben würde. Im besten Fall erreichen natürlichsprachliche Schnittstellen dieselben Akzeptanzwerte beim Nutzer wie herkömmliche Schnittstellen. Das liegt an den vorhandenen technischen Möglichkeiten, aber auch an den spezifisch anderen Problemen, die natürlichsprachliche Schnittstellen aufwerfen (z.B. der Verlagerung der Verantwortung im Fehlerfall vom Benutzer zum System).

Viele Entwickler gehen davon aus, daß durch natürlichsprachliche Schnittstellen den Benutzern ein kommunikativer Umgang mit dem System ermöglicht wird, wie sie ihn auch anderen Menschen gegenüber pflegen. Im Gegensatz zu dieser Intuition steht die

Tatsache, daß die Eingaben an natürlichsprachlichen Schnittstellen hinsichtlich des Sprachumfanges und der Diskursphänomene stark eingeschränkt sind. Offenbar werden natürlichsprachliche Schnittstellen von Benutzern deutlich anders wahrgenommen als menschliche Gesprächspartner.

Die Vorstellung, natürlichsprachliche Schnittstellen seien besonders einfach zu bedienen, weil die Benutzer im Umgang mit natürlicher Sprache besonders vertraut seien, ist ebenfalls weit verbreitet. Hingegen zeigte sich, daß zumindest aktuell realisierte Systeme den Benutzern ebenso große Schwierigkeiten bereiten wie herkömmliche Schnittstellen. Prinzipiell könnte dies sowohl am immer noch unzulänglichen Entwicklungsstand dieser Systeme liegen, als auch eine Folge der Verwendung natürlicher Sprache sein, deren Auswirkungen noch unzureichend untersucht sind. Zu diesen Auswirkungen gehört vermutlich auch die festgestellte Erwartungshaltung der Benutzer, die von natürlichsprachlichen Systemen Wissen über die Welt und Inferenzen über den Datenbestand des Systems verlangen.

Neben den Einschränkungen zur Aussagekraft der zitierten Untersuchungen muß aber auch betont werden, daß manches, was von Systementwicklern als natürliche Sprache bezeichnet wird, doch sehr formal und z.T. lediglich als mit natürlichsprachlichen Einschüben versehen wirkt. Angesichts solcher Eingabemöglichkeiten ist es wenig verwunderlich, daß die Benutzer mit den zugehörigen Schnittstellen Schwierigkeiten hatten. Diese Schwierigkeiten sind aber natürlich kein prinzipielles Argument gegen natürliche Sprache an Schnittstellen.

Ogden und Bernick sehen zwei Möglichkeiten, die für natürlichsprachliche Schnittstellen genannten Probleme zu lösen: Einerseits bieten sich ihrer Ansicht nach Techniken z.B. der Künstlichen Intelligenz an, um die Nutzererwartungen besser als bislang befriedigen zu können. Dies bedeutet die Modellierung von Weltwissen und die Ausweitung der linguistischen Eingabemöglichkeiten. Neben den in 4.1 referierten Schwierigkeiten bleibt aber das Dilemma, daß Nutzer mit steigenden Systemfähigkeiten auch steigende Ansprüche entwickeln und damit die Fortschritte des Systems wieder aufzehren (vgl. Miller (1981)). Denn letztlich ist eine Schnittstelle, die den Turing Test besteht, derzeit nicht ernsthaft in Sicht. Und selbst wenn es eine solche gäbe, würde das noch lange nicht bedeuten, daß sie auch in einer längeren Interaktion mit einem Nutzer bestehen würde, der auf seinem Spezialgebiet Auskünfte sucht: Eine überzeugende Alltagskonversation im Turing Test, und der Dauereinsatz, an einem das Wissen eines Menschen übersteigenden Datenbestand, haben nicht notwendig dieselben Voraussetzungen.

Der zweite mögliche Ansatz, die genannten Probleme zu lösen, besteht für Ogden und Bernick (1997) in verstärktem Nutzertraining, Verbesserung des Systemfeedbacks und der Entwicklung eines konsistenten Systemmodells. Nachteil dieser, ganz auf den Benutzer gerichteten Verbesserungen, ist natürlich der Verlust des eigentlich mit natürlichsprachlichen Schnittstellen inhärent verknüpften Anspruchs, eine Schnittstelle für Benutzer ohne Vorkenntnisse zu sein.

Ogden und Bernick schlagen daher eine Kombination beider Lösungswege vor und

machen einige konkrete Angaben, wie zukünftig bessere Systeme beschaffen sein sollten. Erfolgreiche natürlichsprachliche Schnittstellen sollten laut Ogden und Bernick durch eine klare Definition der Anwendungsdomäne und einer Zielgruppe gekennzeichnet sein: Zu einem gewissen Teil verabschiedeten sich Ogden und Bernick hier von der Utopie der für alle jederzeit und ohne Vorkenntnisse erfolgreich zu bedienenden Schnittstelle, ohne allerdings explizit die Forderung nach Training der Benutzer aufzustellen. Weiterhin müsse man für natürlichsprachliche Schnittstellen mit einer inkrementellen Adaptionphase rechnen, in der die Schnittstelle kontinuierlich verändert werde. Schließlich müßten zukünftige natürlichsprachliche Schnittstellen Zugriff auf Meta-Information geben und sich durch eine deutlich größere syntaktische Abdeckung, sowie ein besseres, auch zur Gestaltung der Eingaben des Nutzers geeignetes Feedback auszeichnen.

Unklar sind bislang die Implikationen, die aus der Wahrnehmung von Computern als soziale Objekte erwachsen, wie es in 6.2.6 angedeutet wurde. Auf dieses Thema wird in Kapitel 7.2 weiter eingegangen. Es scheint aber bereits jetzt klar zu sein, daß es sich hierbei um ein bislang vernachlässigtes Problem der Wirkung von Computerprogrammen im allgemeinen auf Menschen handelt, aber nicht unbedingt um ein spezifisches Ausschlußkriterium für natürlichsprachliche Schnittstellen. Entsprechend werden die Ergebnisse aus 7.2 wichtig sein, für die Weiterentwicklung von (auch) natürlichsprachlichen Schnittstellen, in Kombination mit den von Ogden und Bernick (1997) vorgeschlagenen Verbesserungen.

7 Ausblick

Waltz (1982) schrieb schon vor 18 Jahren, daß man erst damit begonnen habe zu verstehen, wie Menschen sich an natürlichsprachlichen Schnittstellen verhalten. In gewisser Weise könnte man feststellen, daß sich bis heute an diesem Fazit nichts geändert hat und wir immer noch keinen Schritt weiter sind. Man muß dabei aber natürlich bedenken, wie sehr sich seit 1982 die Maschinen und Programme, und damit auch unsere Vorstellung einer natürlichsprachlichen Schnittstelle, verändert haben: Während früher vielfach bereits die Vermengung von SQL-Befehlen mit Wörtern z.B. des Englischen als natürlichsprachliche Eingabe betrachtet wurde (s. 6.2), wird heute schon über die Verarbeitung von Spontansprache in Echtzeit nachgedacht (s. 4.1). Andererseits ist es aber auch richtig, daß die grundsätzlichen Schwierigkeiten¹¹² natürlichsprachlicher Schnittstellen, die bereits vor 20 Jahren beschrieben wurden, auch heute noch dieselben sind. Positiv betrachtet sind wir heute in der Lage, Systeme zu bauen, die in keiner Weise mehr mit denen zu vergleichen sind, die Waltz (1982) beschrieb. Daß diese Systeme trotz ihrer Leistungsfähigkeit offenbar immer noch nicht den (hochgesteckten) Erwartungen der Entwickler in Hinblick z.B. auf eine Bedienbarkeit ohne Training gerecht werden, sollte aber Ansporn sein, weiter zu forschen. Denn es darf nicht übersehen werden: Weder ist es gelungen, eine alternative Interaktionsform zu entwickeln, die die in natürlichsprachliche Schnittstellen gesteckten Erwartungen im allgemeinen Fall erfüllen würde, noch ist es – trotz verschiedener Versuche – gelungen, den Ansatz der natürlichsprachlichen Schnittstelle an sich zu diskreditieren.

Eine wichtige Frage, angesichts der in 5 vorgestellten Verbesserungsmöglichkeiten von OSIRIS, sowie des in 4.2 formulierten Verbesserungsbedarfs für natürlichsprachliche Schnittstellen allgemein, ist die Verfügbarkeit ausreichender linguistischer Ressourcen. Wie in 3.1 bereits für das bestehende OSIRIS angeführt, gilt natürlich auch für dessen mögliche Erweiterungen, daß angesichts der zu bearbeitenden Datenmengen für flächendeckende, linguistische Informationen manuelle Arbeit nicht in Frage kommt.

7.1 Informationen zur Systemerweiterung

Konkret stellt sich die Frage, wo die, für eine Erweiterung von OSIRIS oder vergleichbarer Systeme benötigten, Informationen herkommen sollen. Woher z.B. soll die semantische Information kommen, die für eine substantielle Erweiterung des Systems, und nicht nur für eine Technologiestudie reicht? In diesem Abschnitt können dazu keine endgültigen Antworten gegeben werden. Es sollen jedoch einige vielversprechende Ansätze zur (semi)automatischen Erstellung solcher Ressourcen vorgestellt werden um zu motivieren, daß die vorgeschlagenen Erweiterungen von OSIRIS oder ähnlichen Systemen auch in der nötigen Breite möglich sind.

¹¹²Z.B. die Notwendigkeit von Weltwissen bei gleichzeitiger Unmöglichkeit, dieses derzeit flächendeckend zu gewährleisten.

Soderland et al. (1995) gehen auf das Problem der für NL-Schnittstellen wie für intelligente Information Retrieval Systeme dringend benötigten semantischen Informationen ein. Sie stellen mit CRYSTAL ein System vor, das auf der Grundlage eines von Experten per Hand annotierten Korpus automatisch ein Lexikon mit konzeptuellen Knoten erzeugt. Das Trainingskorpus wird maschinell syntaktisch analysiert und in Subjekt, Prädikat und Objekt entsprechende Teile zerlegt. Das Wortmaterial jeder mittels SGML-artiger Tags annotierten Einheit wird mittels eines semantischen Lexikons Klassen einer Hierarchie zugeordnet, so daß im ersten Schritt Konzepte entstehen, deren Typ einem SGML-Tag entspricht und die beispielsweise Informationen folgender Art enthalten: „Wurde aus einer Präpositionalphrase extrahiert; der Kontext enthielt kein Verb; das Subjekt hieß 'unremarkable'; die Präposition hieß 'with'; die in der Präpositionalphrase enthaltenen Wörter lauten ...“.

Ziel von CRYSTAL ist es nun, durch Generalisierung der Konzepte (d.h. systematische Rücknahme der Bedingungen) die minimale Menge konzeptueller Knoten zu finden, die alle positiven Trainingsinstanzen abdeckt. Weiterhin wird für jedes modifizierte Konzept anhand des Trainingskorpus überprüft, ob die Fehlerrate innerhalb der vorgegebenen Toleranzgrenzen bleibt.

Futrelle und Zhang (1994) versuchen, ohne Rückgriff auf Lexika oder andere externe Quellen aus dem Volltext von Dokumenten automatisch Informationen zu gewinnen, um die Suche über dem Dokumentenbestand zu verbessern. Dazu wird der direkte Kontext der Wörter untersucht, um verwandte Wörter zu entdecken, d.h. solche, die in (annähernd) gleichen Kontexten auftreten. Für die 1.000 häufigsten Wörter eines Korpus wird dann für ein Fenster von z.B. vier Wörtern der Kontext in Form eines Vektors repräsentiert, wobei für jedes der Kontextwörter dessen Häufigkeit relativ zum Gesamtkorpus betrachtet wird. Im Anschluß werden aus der Menge der Vektoren über ein nicht näher genanntes Verfahren die beiden einander ähnlichsten Kontextvektoren identifiziert und die beiden dahinterstehenden Wörter zu einer Einheit der Ebene 0 zusammengefaßt. Auch die beiden Kontexte werden zu einer Einheit verschmolzen und es werden die Ähnlichkeiten im Korpus neu berechnet und der Prozeß wiederholt. In den gezeigten Beispielen ist die semantische Hierarchie sehr einfach, da auf einer höheren Ebene immer Ähnlichkeiten zwischen zuvor für ähnlich befundenen und deshalb zusammengefaßten Elementen und einem neuen Element entdeckt werden – es entsteht also ein linearer Baum, der einfach und automatisch zu einer Klassifikation des verwendeten Wortmaterials beiträgt. Warum dabei keine Ähnlichkeiten zwischen zwei Vektoren auftreten können, ohne daß der im Schritt zuvor aus der Verschmelzung der beiden ähnlichsten Kontexte entstandene Vektor beteiligt wäre (so daß ein Baum mit einer Breite größer zwei entsteht) bleibt offen.

Ein Problem bei diesem Vorgehen ist natürlich, daß ambige Wörter durch das statistische Vorgehen „disambiguiert“ werden, d.h. es wird nur die häufigste Verwendungsweise (in Form ähnlicher Kontexte) erkannt und die anderen werden ignoriert. Deshalb schlagen die Autoren vor, in einer Weiterentwicklung *alle* Wörter (d.h. jedes token) samt ihres Kontextes in einer Datenbank zu speichern und mit Clustering Verfahren

und einem Ähnlichkeitsmaß zu bearbeiten.

Auch wenn die Verfahren tatsächlich so robust und ohne Vorbearbeitung der Daten einzusetzen sind, stellt sich doch die Frage, ob nicht eine gezielte Kombination mit lexikonbasierten Verfahren zu sehr viel besseren Ergebnissen führt, als der Versuch, langfristig nur mit Clustering sogar (domänenabhängig) Diskursstrukturen erkennen zu wollen.

Paek und Jeon (1997) verweisen auf die Möglichkeit, einen Thesaurus für einfache, nicht durch extensive Regelanwendung modellierbare Klassifikationsaufgaben zu verwenden. Daß ein Thesaurus an sich für Klassifikationsaufgaben eingesetzt werden kann ist nicht überraschend – Paek und Jeon (1997) geht es aber um die Anreicherung des Thesaurus mit weiteren Informationen. Der Effekt den sie dabei erzielen überzeugt nicht unbedingt, da ihre Anwendung sehr speziell erscheint (Klassifikation von Objekten, die an einem Budgetplanungsprogramm für einen Privathaushalt natürlichsprachlich mit Hilfe mehrerer Nomen bezeichnet wurden) und die Erweiterungen des Thesaurus nicht trivial sind (u.a. Vorkommenshäufigkeit einzelner Lesarten ambiger Ausdrücke).

Ansätze wie die von Paek und Jeon (1997) könnten durch den Versuch von Rada und Bicknell (1989) unterstützt werden, die an einer semiautomatischen Verbesserung und auch Anreicherung von Thesauri am Beispiel des MeSH¹¹³-Thesaurus arbeiten.

Riloff und Shepherd (1997) stellen fest, daß auch umfangreiche Versuche semantisches Wissen zu sammeln (wie z.B. in WordNet oder Cyc) auf Dauer nur für bestimmte Applikationen nützliche Ergebnisse liefern werden (s.a. Winiwarter (1996)). Den generellen Bedarf an Semantik für natürlichsprachliche Systeme werden solche Projekte nach Meinung von Riloff und Shepherd nicht decken können, weil viele Anwendungen stark domänenabhängig sind und daher durch eine Subsprache mit Fachtermini oder sogar fachspezifischen Jargon gekennzeichnet sind. Diese Subsprachen vollständig in einer Wissensbasis zusammenzuführen halten die Autorinnen weder für möglich noch in Hinblick auf Ambiguitäten für sinnvoll. Riloff und Shepherd schlagen als Ausweg ein einfaches, semiautomatisches Verfahren vor, das aus großen Textkorpora einen Entwurf für ein semantisches Lexikon erstellt, der nachträglich von einem Menschen kontrolliert werden muß.

Die Grundidee von Riloff und Shepherd besagt, daß Wörter einer bestimmten semantischen Kategorie häufig gemeinsam auftreten, z.B. in Aufzählungen, Konjunktionen oder in Nominalkomposita. Ausgehend von nur fünf per Hand identifizierten Mitgliedern je Kategorie, den Startwörtern, verfahren Riloff und Shepherd wie folgt auf einem Textkorpus:

1. Identifiziere alle Sätze des Korpus, in denen mindestens eines der Startwörter vorkommt. Übergebe diese Sätze an einen Parser, der eine Zerlegung in Nominal-, Verbal- und Präpositionalphrasen zurückliefert.

¹¹³Medical Subject Headings – National Library of Medicine, USA.

2. Betrachte die Nominalphrasen, in denen das Startwort als Kopf auftritt. Identifiziere in diesen Fällen den Kontext des Startwortes, als die links bzw. rechts jeweils nächststehenden Nomen (soweit vorhanden und ohne eine Satzgrenze zu überschreiten).
3. Berechne für jedes der so gefundenen Wörter den „category score“ bestehend aus der Häufigkeit, mit der das in Frage stehende Wort in einem Kontextfenster auftritt, dividiert durch die Häufigkeit, mit der das Wort im gesamten Korpus auftritt.
4. Entferne mit Hilfe einer Liste Stopwörter sowie alle Wörter, die weniger als sechs Belege im Korpus haben. Die Stopwortliste wird manuell erstellt und enthält u.a. Zahlen, Pronomen und Artikel. Die verbleibenden Nomen werden nach ihrem category score sortiert, so daß das Wort mit dem höchsten Wert (d.i. das Wort mit der größten Ähnlichkeit zur Kategorie) an höchster Stelle steht.
5. Füge die fünf am höchsten notierten Nomen, die keine Startwörter sind, der Liste der Startwörter der Kategorie hinzu und beginne den Prozeß von vorn.

Am Ende des letzten Zyklus¹¹⁴ steht als Ergebnis die nach dem category score geordnete Liste der Nomen, die nun von einem Menschen abschliessend manuell auf Zugehörigkeit zur Kategorie überprüft werden müssen. Riloff und Shepherd schreiben, daß nach ihren Erfahrungen innerhalb von acht Zyklen gute Ergebnisse erzielt wurden – zuviele Iterationen können dazu führen, daß ein nicht zur Kategorie gehöriges, aber dennoch im Verlauf fälschlich den Startwörtern hinzugefügtes Wort das Ergebnis verdirbt.

Neben der Anzahl der Iterationsschritte spielt die Qualität des verwendeten Korpus eine entscheidende Rolle: Ist eine manuell mit Startwörtern versehene semantische Kategorie nicht ausreichend im Korpus vertreten, werden bereits bei sehr wenigen Iterationen abwegige Wörter als Startwörter angenommen, was zu einer unbrauchbaren Ergebniswortliste führt. Versuche der Autorinnen mit einer größeren Zahl manuell vorgegebener Startwörter (zehn statt fünf) zeigten keinen Effekt hinsichtlich der Qualität des Ergebnisses, jedoch wurden weniger Iterationen benötigt, um in etwa dieselbe Ergebnisqualität zu erreichen.

Der Ansatz von Riloff und Shepherd wirkt auf den ersten Blick bestechend, unter anderem weil er auf einfachem Text arbeitet, ohne eine Vorverarbeitung oder externes Wissen vorauszusetzen. Auch die einfache Herangehensweise und der Verzicht auf aufwendige Automatismen erscheinen vorteilhaft. Neben dem vertretbaren Anteil manueller Arbeit ist aber zu beachten, daß möglichst homogene, große Korpora Voraussetzung für qualitativ gute Ergebnisse sind. Ob diese Korpora für beliebige semantische Kategorien zur Verfügung stehen und wieviel Arbeit eventuell in die Prüfung oder gar Homogenisierung eines Korpus fließt bleibt offen.

¹¹⁴Die Anzahl der Wiederholungen wird dem System explizit vorgegeben.

Neben der Analyse von Text liegt eine andere Möglichkeit, (semi)automatisch semantisches Wissen bereit zu stellen, in der Extraktion impliziten Wissens in strukturierten Quellen wie z.B. Wörterbüchern. Chodorow et al. (1985) stellen ein Verfahren vor, das semantische Informationen aus den Einträgen herkömmlicher (maschinenlesbarer) Wörterbücher extrahieren soll. Dazu wird ein solcher Eintrag geparkt und in den Lexikoneintrag und die Lexikondefinition zerlegt. Innerhalb der Lexikondefinition wird dann mit einer Heuristik versucht, das dem Lexikoneintrag übergeordnete Konzept zu identifizieren. Der Rest des Eintrages wird dann als die für den aktuellen Eintrag charakteristische Spezifikation des übergeordneten Konzeptes betrachtet. Lautet der Eintrag z.B.

car: (n) a *vehicle* moving on wheels

dann wird „vehicle“ als der Kopf der Phrase „a *vehicle* moving on wheels“ als dem Lexikoneintrag „car“ übergeordnetes Konzept betrachtet. Die Verbalphrase „moving on wheels“ bezeichnet dann das spezifische Charakteristikum des Lexikoneintrages „car“ im Unterschied zu „vehicle“. Als Ergebnis entsteht eine semantische Hierarchie mit sich überschneidenden Abhängigkeiten, da aufgrund von Ambiguität einzelne Wörter mehrere übergeordnete Konzepte haben können.

Véronis und Ide (1991) kritisieren den Ansatz von Chodorow et al. (1985) als unzureichend, weil ein einzelnes Wörterbuch keine verlässliche Quelle für einen automatischen Extraktionsprozeß sei. Véronis und Ide berichten von Fehlerraten in Höhe von 55 % bis 70 % für automatisch aus nur einem Wörterbuch erzeugte semantische Hierarchien. Diese unzumutbare Fehlerrate wird durch die Modifikation des Verfahrens durch Véronis und Ide nach deren eigenen Angaben auf nur noch 6 % gesenkt. Auch wenn einzelne Wörterbücher keine zuverlässige Quelle für den automatischen Aufbau lexikalisch-semantischer Wissensbasen sind, weil die aus ihnen extrahierbaren Hierarchien unvollständig sind, so könnten doch viele Wörterbücher gemeinsam durchaus eine wichtige Rolle in diesem Prozeß spielen, weil es sehr unwahrscheinlich ist, daß Information, die in einem Wörterbuch nicht vorhanden ist, systematisch auch in den anderen Wörterbüchern fehlt.

Von dieser Idee ausgehend haben Véronis und Ide fünf¹¹⁵ ganz unterschiedliche Wörterbücher als Grundlage genommen. Wird ein Lexikoneintrag wie „cup“ in allen fünf Quellen als Hyperonym zu „vessel“ verzeichnet, so kann diese Information mit sehr großer Sicherheit übernommen werden. Wird „cup“ nur in drei von fünf Quellen als Hyperonym zu „vessel“ bezeichnet, ansonsten aber nicht erwähnt, dann wird die Information, daß „cup“ ein Hyperonym zu „vessel“ sei übernommen. Wird „cup“ hingegen in einer Quelle als Hyperonym zu „vessel“, in einer anderen aber als Hyperonym zu „container“ angegeben, dann kann, wenn bereits taxonomische Information zum Zusammenhang von „container“ und „vessel“ vorhanden ist, die Angabe von „cup“

¹¹⁵Collins English Dictionary, Oxford Advanced Learner's Dictionary, COBUILD Dictionary, Longman's Dictionary of Contemporary English und Webster's 9th Dictionary.

als Hyperonym zu „container“ als Bestätigung der Angabe von „cup“ als Hyperonym zu „vessel“ gesehen werden. Ist noch nichts über das Verhältnis von „container“ zu „vessel“ bekannt, wird die Bearbeitung dieses Eintrages bis zum nächsten Rekursionsschritt zurückgestellt, da dann aus der Bearbeitung anderer Einträge eventuell neue Informationen zum Verhältnis der beiden Konzepte vorliegen.

Am Ende mehrerer Rekursionsschritte werden dann immer noch einige ungeklärte Fälle übrig bleiben, zumeist bei den Zuordnungen in der oberen, abstrakteren Schicht der Konzepte („tool“, „utensil“, „instrument“ etc.). Diese Fälle können am besten vermutlich nur durch manuelle Kontrolle bearbeitet werden, fallen aber angesichts der eigentlichen Aufgabe, eine umfangreiche semantische Hierarchie zu entwerfen nach Aussagen von Véronis und Ide nicht ins Gewicht.

Knight und Luk (1994) gehen in eine ähnliche Richtung, wenn sie die (semiautomatische) Vereinigung von fünf¹¹⁶ lexikalischen Ontologien beschreiben. Dabei sehen Knight und Luk vor, die Vereinigung sequentiell in vier Schritten und mit unterschiedlichen Mitteln zu betreiben. Je nach Quelle wird die Vereinigung manuell (Penman und Ontos), überwacht (ontologisch angereicherte Lexikoneinträge des Spanischen und des Englischen zu Übersetzungspaaren) oder automatisch (WordNet und Longman) vorgenommen, ausgehend von der Überzeugung, daß die Kontrolle eines solchen Verfahrens unter einer geeigneten Oberfläche für Menschen wesentlich weniger aufwendig ist, als die Ergebnisse selbst zu erstellen. Angesichts des Ergebnisses mit 400 (manuell aus Penman und Ontos zusammengeführten) generellen Konzepten, einer Mittelschicht mit 50.000 Konzepten (meist auf Wortebene) und einer zahlenmäßig nicht näher bestimmten unteren Schicht für stark domänenabhängige Konzepte erscheint der Aufwand für die menschliche Kontrolle des Prozesses aber nicht unerheblich.

Hötker (1997) behandelt ebenfalls die Frage der Wiederverwertung lexikalischer Informationen aus bereits bestehenden Quellen. Allerdings geht er über die bereits genannten Ansätze hinaus, indem er lexikalische Quellen unterschiedlicher Art wie z.B. Wörterbücher, maschinenlesbare Lexika oder Korpora gleichermaßen berücksichtigt. Außerdem ist die Arbeit von Hötker richtungweisend, da er nicht nur den Import lexikalischer Informationen aus verschiedenen Quellen und Formaten in ein internes Repräsentationsformat berücksichtigt, sondern für die Struktur dieses Repräsentationsformates auch den möglichen Export lexikalischer Information für unterschiedliche Anwendungsgebiete berücksichtigt. Hötker geht also sehr viel genereller vor und leistet nicht nur beispielhafte Arbeit für den Wissenserwerb eines konkreten Systems, sondern stellt eine generelle Möglichkeit der Wiederverwertung lexikalischer Informationen vor.

Eines der Probleme, die Hötker in diesem Zusammenhang angeht, sind die bekannten theorieabhängigen Eigenarten lexikalischer Quellen, die bei naiver Vereinigung entweder zu Inkonsistenzen oder aber zu unvermeidbar hohem Informationsverlust führen können. Zudem sollte eine neue lexikalische Wissensbasis Möglichkeiten vorsehen,

¹¹⁶Penman Upper Model, Ontos Model, Longman, Collins Spanisch-Englisch und WordNet.

Weiterentwicklungen im lexikologischen Erkenntnisprozeß inkrementell zu übernehmen, um auch zukünftig den Im- und Export theoretisch aktueller lexikalischer Quellen gewährleisten zu können.

Weiterhin sind bestimmte Aspekte eines existierenden Lexikons anwendungsabhängig und sollten beim Import in ein generelles Repräsentationsformat als anwendungsabhängig erkennbar bleiben. So haben ein auf Fremdsprachenlerner im ersten Jahr spezialisiertes Lexikon und eines für die Rechtschreibkorrektur einer Standardsoftware ganz andere Ansprüche an die Sprachabdeckung und die Modellierungstiefe der Einträge.

Und schließlich reicht es nicht, daß in einem Lexikon, das aus verschiedenen Quellen gewonnen wurde, die strukturelle Information der Einträge konsistent ist. Ein solches Lexikon läßt sich nur nutzen, wenn die Bedeutung der strukturellen Merkmale bekannt ist und diese aufeinander abgebildet werden können. Dazu ist in der Regel aber eine Inspektion und eine syntaktische Normalisierung der einzelnen Quellen nötig.

Um diesen Ansprüchen gerechtwerden zu können, schlägt Hötger (1997) eine dynamische Organisation lexikalischen Wissens mit Hilfe der Wissenspakete von Wachsmuth vor (Wachsmuth, 1987, 1989; Wachsmuth und Gängler, 1992). Dadurch werden spezifische Informationen in einer Baumstruktur im Bereich der Blätter vorgehalten und durch die Definition von Sichtbarkeitsbereichen unterschiedliche, konsistente Sichten auf einen Datenbestand ermöglicht, der möglicherweise insgesamt inkonsistente Informationen enthält.

Mahesh und Nirenburg (1996) bemerken, daß es sinnvoll ist, für mehrsprachige Anwendungen die lexikalische Repräsentation sprachabhängig in eigenen Lexika zu belassen. Unter dieser Voraussetzung stellt sich aber die Frage nach der Verknüpfung der lexikalischen Information mit dem Weltwissen. Statt für jedes (sprachspezifische) Lexikon eine eigene Wissensbasis zu erstellen schlagen Mahesh und Nirenburg vor, das Weltwissen in einer sprachunabhängigen Ontologie zu repräsentieren, die von allen Lexika gemeinsam genutzt werden kann. Dabei kommen sie zu der Einsicht, daß unter den Rahmenbedingungen ihrer Aufgabe eine möglichst eingeschränkte Ausdrucksmächtigkeit des Repräsentationsformalismus erstrebenswert ist, da sonst die resultierenden Repräsentationen undurchschaubar würden.

Mahesh und Nirenburg leisten mit ihrem Vorgehen, das letztlich auf die Verbesserung maschineller Übersetzungen durch die Integration von linguistischem Wissen und Weltwissen abzielt, keinen Beitrag zur (semi)automatischen Erstellung dieser Wissensressourcen. Sie tragen aber dazu bei, daß die vorhandenen oder auch zukünftig (semi)automatisch konstruierten Ressourcen gemeinsam und damit besser eingesetzt werden können.

Automatische Verfahren um andere Informationen wie z.B. Dialogakte zu finden stehen erst am Anfang. Dialogakte manuell zu bestimmen (wie in VERBMOBIL geschehen, siehe S. 113) ist eine enorme Arbeit. Ausreichend viele transkribierte Aufnah-

men¹¹⁷ oder (für Systeme, die nicht auf gesprochene Sprache abzielen) ausreichend große, für die Domäne einschlägige Korpora müssen von Menschen analysiert und annotiert werden. Diese Annotationen müssen von verschiedenen Personen vorgenommen werden, um intersubjektive Übereinstimmungen feststellen zu können. Die Vorgabe theoretisch von Menschen angenommener Dialogakte ist oftmals sehr subjektiv und die Abbildung dieser angenommenen Dialogakte auf die tatsächlich vorliegenden Daten sehr schwierig, als daß dieses Vorgehen einen vollwertigen Ersatz für die manuelle Extraktion darstellen würde. In der Menge der für eine Domäne identifizierten Dialogakte müssen dann intellektuell die einschlägigen Elemente selektiert werden.

Möller (1997) stellt im Anwendungskontext von VERBMOBIL einen interessanten Versuch vor, um automatisch (unüberwacht) Dialogakte aus den Daten lernen zu lassen. Dazu bedarf es einiger (für VERBMOBIL zum großen Teil schon vorhandener) domänenspezifischer Module für die Worterkennung, die Behandlung von Prosodie sowie die Syntax- und die Semantikanalyse. Als Lernalgorithmus wird der inkrementelle *ClassItAll* eingesetzt, der sich durch die mögliche Kombination verschiedener Wissensrepräsentationen (symbolisch, numerisch, strukturiert und mit Unsicherheit behaftet) auszeichnen soll.

Die mit diesem Verfahren für die VERBMOBIL Domäne erkannten Dialogakte unterscheiden sich von denen, die durch manuelle Bearbeitung der Daten entstanden, aber Möller zeigt im direkten Vergleich zu den VERBMOBIL Dialogakten, daß die automatisch gelernten Dialogakte mindestens ebenso geeignet sind hinsichtlich Erkennbarkeit, Vorhersagbarkeit und Verwendbarkeit in einem Dialogplaner.

Eugenio et al. (1998) kommen in ihrer Untersuchung der Möglichkeiten, Dialogakte zu erkennen, zu weniger befriedigenden Ergebnissen. Im Rahmen der „Discourse Resource Initiative (DRI)“, die es sich zum Ziel gesetzt hat, unterschiedliche Korpora durch eine einheitliche, unterspezifizierte, individuell zu verfeinernde Annotierung international verfügbar zu machen, wurde die Leistung von Menschen bei der manuellen Annotation untersucht. Eugenio et al. ließen dazu 24 verschriftete, computervermittelte, zwischenmenschliche Verhandlungsdialoge¹¹⁸ von Menschen annotieren, um die Grenzen von Dialogakten, die Rückbezüge auf vorhergehende Dialogakte und Objekte im Dialog und die Art des Dialogaktes zu bestimmen. Gewünscht wurde eine empirisch stichhaltige Annotation eines Dialogabschnittes mit bspw. folgender (hier natürlichsprachlich umschriebener) Information: „Dialogabschnitt 25 ist die Ablehnung durch Sprecher B des in Abschnitt 16 von Sprecher A getätigten Vorschlages“. Dabei zeigte sich, daß die menschlichen Kodierer nicht übereinstimmend in der Lage waren, auf Grundlage der verschrifteten Dialoge zu bestimmen, ob es sich um eine Zustimmung, eine Ablehnung oder eine Weiterverhandlung handelte. Auch hatten die

¹¹⁷Der Mangel an ausreichenden, einschlägigen Aufnahmen kann dazu führen, daß mit Hilfe von Studenten, die Terminabsprachen zwischen Geschäftsleuten mimen, Aufnahmen erst generiert werden müssen.

¹¹⁸Die Dialoge wurden von zwei Menschen in verschiedenen Räumen über eine Computertastatur geführt.

menschlichen Bearbeiter große Schwierigkeiten zu bestimmen, worauf im vorangehenden Dialog sich genau dieser Dialogakt bezog.

Eugenio et al. schließen daraus, daß die Chancen für eine automatische, zuverlässige Erkennung von Dialogakten gering sind, da Computer über sehr viel weniger Hintergrundwissen beim Sprachverstehen verfügen als Menschen.

7.2 Computer als soziale Agenten

In 6.2.6 wurde bereits angesprochen, daß die Wirkung von Computern auf ihre Benutzer durch die Verwendung von Merkmalen, die normalerweise mit Menschen assoziiert werden, verändert werden kann.

Zahlreiche Untersuchungen kommen zu dem Schluß, daß z.B. durch Sprachausgabe das System von Benutzern nicht mehr nur als unbelebtes Werkzeug, sondern als ein Objekt behandelt wird, gegenüber dem soziale Normen zu wahren sind und auf das zwischenmenschliche Stereotypen projiziert werden können. Angesichts dieser Erkenntnisse fällt es nicht schwer zu glauben, daß Benutzer unter diesen Umständen auch von Computern erwarten, daß diese über Weltwissen verfügen und Inferenzen ziehen können.

Nass et al. (1993) und Nass et al. (1994) zeigen, daß Menschen systematisch und unbewußt dazu gebracht werden können, Maschinen gegenüber ein Verhalten zu zeigen, das auch nach Ansicht der Versuchspersonen selbst Maschinen gegenüber unangemessen ist. Ihr Vorgehen beschreiben Nass et al. wie folgt: Nimm ein aus der Soziologie oder Psychologie bekanntes, experimentell nachweisbares Verhaltensmuster, das zwischen Menschen wirksam ist. Ersetze in der Theorie wie im Versuchsaufbau einen von zwei oder mehr menschlichen Beteiligten durch einen oder mehrere Computer. Statte diese Computer mit Merkmalen aus, die normalerweise nur Menschen zugeschrieben werden (Sprachausgabe, menschliches Rollenverhalten etc.) und prüfe, ob die ursprüngliche Verhaltensregel auch in diesem Szenario noch gilt.

Nass et al. wenden dieses Vorgehen in verschiedenen Studien an, um zu klären, ob Benutzer von Computern

- Regeln allgemeiner Höflichkeit gegenüber Maschinen beachten.
- die Konzepte „Selbst“ und „Gegenüber“ auf Maschinen übertragen und falls ja, ...
- ... auf welcher Grundlage sie diese Unterscheidung vornehmen (Stimme der Ausgabe oder Gehäuse).
- geschlechtsspezifische Stereotypen auf die Maschine übertragen.
- meinen, mit der Maschine selbst, oder mit einem anderen Agenten wie z.B. dem Programmierer der Maschine zu interagieren, wenn sie sozial auf Maschinen reagieren.

Um diese Fragen zu klären, wurden nur Versuchspersonen mit Computervorerfahrung herangezogen, um die Verfälschung der Untersuchungsergebnisse durch realitätsferne Vorstellungen von Anfängern über die Funktionsweise eines Computers zu verhindern. Der generelle Versuchsaufbau sah zu Beginn eine Vermittlungsphase vor, in der 25 bis 30 Fakten aus einer, den Versuchspersonen vertrauten, Domäne (Massenmedien, Computer, Sozialverhalten von Menschen ...) erklärt wurden. Darauf folgte eine Testphase, in der die Versuchspersonen 15 Fragen beantworten sollten. Im Anschluß daran wurde von einem Computer der Erfolg der ursprünglichen Vermittlungsphase bewertet. Zum Schluß mußten die Versuchspersonen einen Fragebogen mit Bewertungen und persönlichen Einstellungen zu allen vorangegangenen Phasen des Experimentes ausfüllen. Nach dem Experiment wurden die Versuchspersonen über den Versuchsaufbau aufgeklärt und nach ihren Einstellungen zu Maschinen und der Angemessenheit sozialer Reaktionen gegenüber Maschinen befragt.

Um der Frage der Höflichkeit nachzugehen, wurde für 33 Versuchspersonen die Vermittlungs-, Test- und Evaluationsphase auf ein und demselben Computer durchgeführt (Nass et al., 1994). Im Anschluß daran wurde das Ausfüllen des Fragebogens variiert: Ein Teil der Versuchspersonen füllte den Fragebogen auf demselben Computer aus, der auch für die Vermittlungs-, Test- und Evaluationsphase verwendet wurde, eine andere Gruppe verwendete einen zweiten, zuvor nicht benutzten Computer, und eine dritte Gruppe füllte den Fragebogen in Papierform aus.

Personen, die das vorangegangene Experiment auf derselben Maschine bewerteten, die zuvor die Vermittlungs-, Test- und Evaluationsphase durchgeführt hatte, äußerten sich signifikant homogener und positiver über die Ansprache und die Kompetenz dieser Maschine, als diejenigen, die auf Papier oder auf einer anderen Maschine den Fragebogen beantworteten. Die Antworten der beiden Gruppen, die auf Papier bzw. einer anderen Maschine gearbeitet hatten, unterschieden sich nicht signifikant untereinander. Nass et al. schließen daraus, daß Menschen elementare Höflichkeitsregeln auch gegenüber Maschinen befolgen, und sich an ein und demselben Computer positiver über dessen Leistungen äußern, als sie dies in dessen Abwesenheit tun würden. Zudem scheint das Medium, mit dem die Bewertung vorgenommen wird, keinen Einfluß zu haben, wie das gleichartige Abschneiden der beiden anderen Gruppen zeigt.

Obwohl die Versuchspersonen in der Aufklärungsphase nach dem Experiment bestätigten, daß Höflichkeit ein gegenüber Maschinen unangebrachtes Verhalten ist, wendeten sie dennoch dieses Verhalten unbewußt im Versuch an. Nass et al. (1999) können in einem weiteren Experiment diese Ergebnisse bestätigen.

Um zu überprüfen, ob die Versuchspersonen die Konzepte von „Selbst“ und „Gegenüber“ auf Computer übertragen, wurden 44 Versuchspersonen mit zwei bzw. drei Maschinen konfrontiert, die über Sprachausgabe mittels Aufzeichnungen menschlicher Stimmen, sowie zwei deutlich ausgeprägten Grundeinstellungen („kritisierend“ vs. „lobend“) verfügten (Nass et al., 1993, 1994). Die eine Gruppe arbeitete in der Vermittlungs- und der Evaluationsphase mit demselben Computer, der auch konstant dieselbe Stimme verwendete. Die andere Gruppe arbeitete in der Vermittlungs- und

in der Evaluationsphase mit (physikalisch) verschiedenen Computern, die auch unterschiedliche Stimmen in der Ausgabe verwendeten. Beide Gruppen interagierten in der Testphase mit jeweils anderen Maschinen, so daß die erste Gruppe mit zwei verschiedenen Rechnern, die zweite Gruppe aber mit drei verschiedenen Rechnern in Berührung kamen. Zusätzlich wurde, quer zu den genannten Gruppen, die Evaluationsphase manipuliert: Für einen Teil der Versuchspersonen bewertete der Computer die Vermittlungsphase überwiegend kritisch, gegenüber den anderen Versuchspersonen fiel die Bewertung überwiegend positiv aus. Alle Versuchspersonen füllten den abschließenden Fragebogen in Papierform aus.

Die Versuchspersonen befolgten verschiedene, aus dem zwischenmenschlichen Verhalten bekannte Regeln gegenüber den Computern und verhielten sich konsistent zu der Annahme, daß physikalisch unterschiedliche Maschinen mit unterschiedlichen Stimmen unterschiedliche soziale Wesen sind. Die Versuchspersonen beachteten im Experiment offenbar die folgenden Regeln:

- Eigenlob ist weniger objektiv und weniger freundlich als das Lob anderer Personen.
- Kritik anderer Personen erfordert mehr Intelligenz als Lob anderer Personen.
- Eine kritisierte Leistung ist einer mit Lob bedachten Leistung unterlegen.

Im Experiment bedeutete dies, daß die Versuchspersonen eine positive Beurteilung der Vermittlungsphase in der Evaluationsphase als freundlicher und korrekter bewerteten, wenn diese Bewertung von einem Computer vorgenommen wurde, der physikalisch eigenständig und mit einer anderen Stimme ausgestattet war, als der Rechner, der die Vermittlungsphase durchgeführt hatte. Eine negative Beurteilung der Vermittlungsphase in der Evaluationsphase wurde als weniger freundlich, aber intelligenter bewertet, wenn diese Bewertung von einem Computer vorgenommen wurde, der physikalisch eigenständig und mit einer anderen Stimme ausgestattet war, als der Rechner, der die Vermittlungsphase durchgeführt hatte. Wurde die Bewertung der Vermittlungsphase von einer anderen Maschine mit anderer Stimme vorgenommen, betrachteten die Versuchspersonen die Bewertung als präziser, und die Vermittlungsphase, die bewertet wurde, als fairer. Die Versuchspersonen bewerteten in einem Fragebogen (auf Papier) auch selbst die Leistung derjenigen Computer aus der Vermittlungsphase besser, deren Leistung zuvor in der Evaluationsphase gelobt worden war. Nass et al. (1994) berichten zudem, daß die Versuchspersonen rückblickend der (objektiv falschen) Ansicht waren, daß sie mehr Fragen korrekt beantwortet hatten in den Testphasen, die auf in der Evaluation gelobte Vermittlungsphasen folgten.

In einer Erweiterung der vorangegangenen Untersuchung um zusätzliche Kombinationen aus Rechnern und zugehörigen Stimmen stellten Nass et al. (1994) fest, daß die Identifikation der sozialen Agenten über die Stimme und nicht über das (physikalische) Gerät erfolgt. Verschiedene Stimmen wurden als unterschiedliche Akteure und gleiche Stimmen als ein einziger Akteur identifiziert.

Um zu überprüfen, inwieweit Menschen geschlechtsspezifische Stereotypen auf Maschinen übertragen, wurden 24 weibliche und 24 männliche Versuchspersonen mit je einem anderen Rechner in der Vermittlungsphase, der Testphase und der Evaluationsphase konfrontiert. In der Vermittlungs- und der Evaluationsphase wurden männliche oder weibliche Stimmen in allen möglichen Kombinationen¹¹⁹ eingesetzt, die Testphase wurde ohne Sprachausgabe durchgeführt. Nass et al. (1994) berichten, daß die Versuchspersonen sich in Übereinstimmung mit geschlechtsspezifischen Stereotypen wie „Lob von einem Mann ist überzeugender als Lob von einer Frau“ und „Frauen wissen mehr über Gefühle und zwischenmenschliche Verhältnisse als Männer“ verhielten. Nass et al. (1997) können in einem weiteren Experiment diese Ergebnisse bestätigen. Lediglich aufgrund der geschlechtsspezifischen Stimmen (bei sorgfältigem Ausschluß aller anderen möglichen Hinweise auf eine Persönlichkeit oder ein Geschlecht von Seiten des Computers) werden geschlechtsstereotypische Reaktionen der Versuchspersonen provoziert.

Zu ähnlichen Ergebnissen kommen auch Nass et al. (1995): 48 Versuchspersonen wurden mit einem psychologischen Standardtest als dominant oder sich unterordnend charakterisiert und mit Computersystemen konfrontiert, von denen einige dominantes Verhalten zeigten, andere aber sich dem Benutzer eher unterordneten. Dieses Systemverhalten wurde durch die Systemreaktionen erzeugt, das bspw. in der dominanten Variante aus Anweisungen und Aussagesätzen bestand und stets die Initiative in der Benutzerinteraktion hatte. Die dem Versuch zugrundeliegenden Hypothesen lauteten:

- Die Benutzer nehmen die beschriebenen Charakteristiken in der Ausgabe des Systems wahr und identifizieren sie mit menschlichen Eigenschaften wie „Dominanz“ bzw. „Fügsamkeit“.
- Insgesamt steht die Bewertung der Systeme hinsichtlich z.B. Kompetenz in keiner Abhängigkeit zu den genannten Charakteristika in den Systemausgaben.
- Benutzer ziehen es aber vor, mit „ähnlichen“ Systemen zu arbeiten: Beispielsweise sollten Benutzer, denen der erwähnte Standardtest eine dominante Persönlichkeit bescheinigte, mit einem dominanten System zufriedener sein, als mit einem, das sich ihnen unterordnet (similarity-attraction relationship).

In der Versuchsdurchführung unterschieden sich die Systemausgaben der beiden Varianten nicht hinsichtlich des Informationsgehaltes, sondern lediglich hinsichtlich des Stils (Wortwahl etc.). Alle Systemreaktionen waren antizipiert und es waren keine natürlichsprachlichen Eingaben möglich.

Der Versuch bestätigte alle drei genannten Hypothesen. Moon und Nass (1996) können diese Ergebnisse ebenfalls bestätigen. Nass et al. schließen daraus, daß Persönlichkeit eines Computersystems ein sehr wirksames und leicht zu manipulierendes Mittel für die Mensch-Maschine-Interaktion ist, das auch mit einfachsten Mitteln komplexe

¹¹⁹M/M, M/F, F/M, F/F

soziale Reaktionen beim Benutzer hervorruft. Insbesondere betonen Nass et al. den Punkt, daß weder natürliche Sprache, noch Methoden der Künstlichen Intelligenz oder animierte Agenten eine notwendige Voraussetzung für die Wahrnehmung der Systeme als soziale Agenten durch die Benutzer ist.

Bei der Diskussion der möglichen Erklärungen für die gezeigten Phänomene betonen Nass et al. noch einmal den Punkt, daß es sich bei allen Versuchspersonen um Grundstudiumsstudenten mit Computervorerfahrung handelte, die mit Sicherheit sich der Tatsache bewußt waren, daß Computer unbelebte Maschinen sind, denen gegenüber Verhaltensweisen wie z.B. Höflichkeit unangemessen sind. Diese Sicht wurde auch in den auf das Experiment folgenden Befragungen von den Versuchspersonen bestätigt. Eine andere Erklärungsmöglichkeit, wie es zu den beschriebenen Reaktionen kommt, besagt, daß die Versuchspersonen sehr wohl wissen, daß Verhalten wie z.B. Höflichkeit unangemessen gegenüber Computern ist, daß sich die Versuchspersonen aber gedanklich nicht an die Maschine, sondern an eine hinter der Maschine gedachte Person, z.B. den Programmierer wenden.

Nass et al. (1994) (sowie später auch Nass et al. (1999)) weisen diese Erklärungsmöglichkeit zurück und führen die Ergebnisse einer Untersuchung mit 33 Versuchspersonen an, die in dem bereits zuvor für Nass et al. (1994) dargestellten Versuchsszenario mit zwei Computern interagierten. Diesmal durchliefen die Versuchspersonen den Zyklus aus Vermittlung, Test und Evaluation allerdings zweimal, wobei im ersten Durchgang die Evaluation generell positiv und im zweiten Durchgang generell negativ ausfiel. Vermittlung und Evaluation fanden auf demselben Rechner statt, der Test wurde auf einem zweiten Rechner durchgeführt. Variiert wurde die Art, wie in der Systemausgabe sowie durch den Versuchsleiter auf den Computer referiert wurde: Für die erste Gruppe referierten der Versuchsleiter und die Systemausgaben mit „der Computer“ oder „dieser Computer“. Für die zweite Gruppe wurde die Referenz in der Systemausgabe mit „Ich“ und von Seiten des Versuchsleiters mit „der Computer“ vorgenommen. Die dritte Gruppe bekam als Referenz auf den Computer in der Systemausgabe wiederum „Ich“ zu hören, von Seiten des Versuchsleiters aber wurde mit „der Programmierer“ auf den Computer referiert. Im Vergleich mit den Personen der dritten Gruppe beurteilten die Versuchspersonen der ersten und zweiten Gruppe den Computer als freundlicher und leichter zu handhaben. Das Verhalten der Angehörigen der ersten beiden Gruppen zeigte untereinander keinen Unterschied. Ganz offensichtlich reagieren die Versuchspersonen direkt auf den Computer und nicht auf eine angenommene Person wie den Programmierer.

Nass et al. erklären ihre Ergebnisse damit, daß soziale Reaktionen von Menschen auf Computer automatisch und unbewußt erfolgen, da die Mensch-Maschine-Kommunikation eine ganze Reihe von Eigenschaften der zwischenmenschlichen Kommunikation zeige. Welche Merkmale oder Kombinationen von Merkmalen genau die Benutzer dazu veranlassen, soziale Normen aus dem zwischenmenschlichen Bereich auf die Mensch-Maschine-Kommunikation zu übertragen, können Nass et al. nicht klären.

Lester und Stone (1997) stellen eine Lernsoftware für Kinder vor, die über einen animierten Agenten (Persona) verfügt, der mit vorgefertigten Texten Informationen zu verschiedenen Wissensgebieten vermittelt. Außerdem hilft der Agent bei der Bearbeitung von Übungsaufgaben und überwacht dazu das Problemlösungsverhalten des Lerners. Das Verhalten des animierten Agenten wird von einer Steuerungskomponente geplant, die auf der Grundlage von konkurrierenden Plänen versucht, ein lebensechtes und pädagogisch angemessenes Verhalten zu erzeugen. Lester et al. (1997a) stellen für das beschriebene System den sogenannten „Persona-Effekt“ fest, d.i. die Steigerung der Motivation der Lernenden durch die Anwesenheit des animierten pädagogischen Agenten.

Auch van Mulken et al. (1998) untersuchen den Persona-Effekt und kommen zu dem Ergebnis, daß animierte Charaktere von den Lernern als unterhaltend empfunden werden und dafür sorgen, daß die von diesen Charakteren begleiteten Aufgaben subjektiv als einfacher wahrgenommen werden. Dies scheint insbesondere auch materialabhängig zu sein, d.h. Aufgaben aus einem eher technischen Bereich zeigten diesen Effekt stärker als andere Aufgaben. Allerdings konnten van Mulken et al. keinen objektiven Einfluß feststellen.

Wie schon für die Untersuchungen von Nass et al. festgestellt, geben auch die Untersuchungen von Lester et al. und van Mulken et al. keinen Aufschluß über den Einfluß natürlicher Sprache auf den Benutzer, da keines der betrachteten Systeme natürliche Sprache als Eingabe akzeptierte. Auch auf der Ausgabeseite wurde nur mit herkömmlichen, vorgefertigten Texten gearbeitet.

Schweizer et al. (2000) hingegen kommen in einer Langzeituntersuchung von 101 Studenten zu dem Schluß, daß ausgeprägte soziale Merkmale einer virtuellen Dozentin (Sprachausgabe und Standbilder) direkt mit einer Steigerung des Wissenserwerbs korrespondierten. Dabei war der Effekt umso stärker, je mehr von diesen Merkmalen (inklusive Sprachausgabe) vorhanden waren. Abgesehen von diesem, für die Zukunft animierter Agenten in virtuellen Lehr- und Lernumgebungen wichtigen Ergebnis, erhoffen sich die Autoren durch die Integration sozialer Kontexthinweise (Status und Positionsmerkmale) und Dramaturgie (Gestik, Stimmhebung) eine Verbesserung der, ihrer Ansicht nach degenerierten,¹²⁰ Mensch-Maschine-Kommunikation. Diese sei, durch die Abwesenheit der genannten Merkmale, aber auch von z.B. regulierendem Feedback (Nicken, Augenkontakt), durch Depersonalisierungserscheinungen, die Verlagerung der Konzentration des Rezipienten vom Gesprächspartner weg und hin zur Aufgabe, sowie den leichteren Bruch kommunikativer Konventionen gekennzeichnet. Auch Carberry et al. (1999) betonen in diesem Zusammenhang die Notwendigkeit, Gesten und Rückversicherungen wie Kopfnicken auf Seiten des Systems zu ermöglichen, um nach dem Vorbild des menschlichen Gesprächspartners das System zu einem vollwertigen Partner in einem Problemlösungsdialog zu machen. Bemerkenswert ist der Ansatz von Carberry et al. deshalb, weil die Gesten des animierten Charak-

¹²⁰Im Vergleich zur zwischenmenschlichen Kommunikation, die die Autoren als Vorbild der Mensch-Maschine-Kommunikation nehmen.

ters nicht nur auf Grundlage der sprachlichen Systemausgaben generiert werden sollen, sondern in direkter Reaktion auf Mimik und Gestik des menschlichen Benutzers möglich werden sollen. Dazu schlagen Carberry et al. (1999) vor, Videoaufnahmen zwischenmenschlicher Dialoge zu annotieren. Auf der Grundlage eines solchen Korpus wollen Carberry et al. mit Hilfe von Bilderkennungsalgorithmen Zusammenhänge zwischen Dialogakten wie „Zweifel“ oder „Zustimmung“ und mimischen oder auch gestischen Auffälligkeiten (z.B. Bewegung der Augenbrauen) systematisieren.

Walker et al. (1994) können einen Effekt ihrer mit einem menschlichen Gesicht versehenen Schnittstelle auf den Nutzer feststellen. Walker et al. vergleichen zwei Schnittstellen miteinander, an denen Versuchspersonen Fragen beantworten sollten. An einer der beiden Schnittstellen wurde ein synthetisches, mit Sprachausgabe versehenes Gesicht eingesetzt, die andere Schnittstelle präsentierte nur Texte. Die Benutzer der anthropomorphen Schnittstelle verbrachten mehr Zeit am Rechner, machten weniger Fehler bei der Beantwortung der Fragen und produzierten mehr Kommentare zum Versuch. Die um das Gesicht angereicherte Schnittstelle wurde in zwei Ausführungen eingesetzt: Eine Variante zeigte ein neutrales Gesicht und die zweite Variante ein strenges Gesicht. Die Benutzer der Variante mit dem strengen Gesicht verbrachten im Vergleich zu denen, die ein neutrales Gesicht vorfanden, wiederum mehr Zeit am Rechner, machten weniger Fehler bei der Beantwortung der Fragen und produzierten mehr Kommentare zum Versuch. Diese Gruppe empfand die Interaktion mit der Schnittstelle und mit dem Gesicht aber auch als deutlich unangenehmer als die anderen Versuchspersonen.

Benutzer reagieren offenbar stärker auf Gesichter als auf reinen Text, allerdings erstaunlicherweise stärker negativ. Walker et al. versuchen dies positiv zu deuten („längere Verweildauer“ etc.), aber das zugrundeliegende Phänomen scheint tatsächlich Furcht oder etwas ähnliches zu sein. Leider untersuchen Walker et al. kein ausgesprochen freundliches Gesicht. Auch wird nicht thematisiert, wie sehr die Mimik eines menschlichen Gesichtes, gerade im Zusammenspiel mit Sprachausgabe, eigentlich verstanden ist. Die Bezeichnung „neutraler“ Gesichtsausdruck sowie „strenger“ Gesichtsausdruck gibt die Intention der Programmierer wider – die (sprechenden) Gesichter können allerdings eventuell ganz fremdartig und abstoßend wirken.¹²¹

Auch Koda und Maes (1996) können nachweisen, daß Gesichter an der Mensch-Maschine-Schnittstelle die Interaktion beleben und zu längeren Verweilzeiten führen. Sie stellen zudem fest, daß (photo-)realistische Gesichter von den Benutzern positiver empfunden werden als künstliche Gesichter. Allerdings stellen Koda und Maes auch fest, daß (realistische, menschliche) Gesichter, die einen intelligenten Eindruck machten, den Benutzern unsympathisch waren, während Tierkarikaturen als weniger intelligent, aber sympathischer empfunden wurden.

In Hinblick auf die Ergebnisse von Walker et al. (1994) und Koda und Maes (1996) wäre es durchaus denkbar, daß eine niedlich anzusehende, nicht übermäßig intelli-

¹²¹Ein Effekt, der auf Photos durch die vertikale Drehung der Augenpartie oder, im zwischenmenschlichen Bereich, durch die Beobachtung der Mimik eines sehr vertrauten Gesichtes um 180 Grad gedreht (z.B. im Liegen) hervorgerufen werden kann.

gent wirkende, animierte Figur (z.B. ein Papagei oder auch die berühmt-berüchtigte, sprechende Büroklammer) positiver aufgenommen wird, da von diesen (im Vergleich zu einem menschlichen Gesicht) nicht in ernstzunehmendem Maße Intellekt erwartet werden kann.¹²² Dies würde sich mit den Erkenntnissen von Nass et al. decken, die für den Computer eine soziale Rolle identifiziert haben, die durchaus zu Konkurrenzsituationen, aber auch zu Teambildung führen kann (Nass et al., 1996).

De Angeli et al. (1999) bestätigen grundsätzlich die Ergebnisse der Untersuchungen von Nass et al., weisen jedoch auf der Grundlage eigener Untersuchungen auf einen weiteren bedeutenden Aspekt hin: Obwohl Menschen eindeutig in der oben geschilderten Weise sozial auf Computer reagieren, behandeln sie die Maschine dennoch anders als z.B. Menschen. Auch an einer (simulierten) Schnittstelle, die absichtlich menschliche Eigenschaften imitiert, verwenden Benutzer, wie in 6.2.5 bereits dargestellt, nur ein eingeschränktes Register. De Angeli et al. meinen daher, daß Computer und insbesondere solche Computer, die mittels Avataren oder anderen anthropomorphen Schnittstellen dem Nutzer begegnen, ein neues Objekt mit besonderen Eigenschaften in der sozialen Umwelt des Menschen sind. Dieses neue soziale Objekt müsse von der Forschung zur Kenntnis genommen und auf seine Besonderheiten für die Nutzerzufriedenheit und Bedienbarkeit untersucht werden (De Angeli und Petrelli, 2000).

Gong (2000) kommt zu einer in diesem Zusammenhang wichtigen Feststellung. Sie warnt davor, isolierte Aspekte eines Systems auszubauen, sondern kommt auf der Grundlage einer eigenen Untersuchung zu dem Schluß, daß Konsistenz als soziale Norm eine wichtige Eigenschaft auch von Computersystemen ist. Gong beschreibt die Wirkung einer Schnittstelle mit menschlichem Gesicht und Sprachausgabe auf die Nutzer. Sprachausgabe kann zur Zeit in sehr hoher Qualität mit Aufzeichnungen menschlicher Stimmen oder mit deutlich geringerer Qualität in Echtzeit synthetisch generiert werden. Technisch sind Videoaufzeichnungen von Gesichtern an Schnittstellen derzeit nur für genau den Text einsetzbar, der bei der Aufzeichnung gesprochen wurde. In allen anderen Fällen ist die Lippsynchronisation nicht möglich, so daß derzeit Gesichter, die über Mimik verfügen sollen, erfolgreich nur synthetisch eingesetzt werden können. Gong untersucht nun die Wirkung der verschiedenen Kombinationsmöglichkeiten (synthetische Sprache, synthetisches Gesicht vs. realistische Sprache, synthetisches Gesicht) auf den Benutzer und stellt fest, daß die konsistente Schnittstelle, obwohl hinsichtlich der Sprachausgabe qualitativ schlechter, sehr viel positiver aufgenommen wird. Gong schließt, daß die technischen Aspekte natürlichsprachlicher Schnittstellen immer in Zusammenhang mit den (bislang zumeist übersehenen) sozialen Aspekten betrachtet werden müssen, um langfristig erfolgreiche Systeme entwickeln zu können.

Dabei ist bereits jetzt abzusehen, daß der Bedarf an einem tieferen Verständnis des Zusammenhangs von technischem Fortschritt und sozialer Wirkung von Schnittstellen kontinuierlich steigen wird. Zahlreiche Erweiterungs- und Verbesserungsvorschläge für animierte Charaktere und natürlichsprachliche Schnittstellen setzen gezielt auf die

¹²²Lester und André, personal communications.

Wirkung menschlicher Attribute an Computersystemen. Walker et al. (1996) und Walker et al. (1997) wollen mit linguistischen Mitteln die Ausprägung eines Sprachstils ermöglichen, der die Glaubwürdigkeit artifizierlicher Agenten erhöhen soll, indem diese den Anschein einer individuellen Persönlichkeit bekommen. Towns et al. (1998) bemühen sich ebenfalls um die Glaubwürdigkeit animierter Agenten und wollen dazu Gesten und andere Bewegungen mit der Sprachausgabe koordinieren, um einen stimmigen Gesamteindruck zu erreichen. Moldt und von Scheve (2000) sprechen bereits von Hybridgemeinschaften zwischen Agenten und Benutzern. Dazu soll der Agent explizit, mittels anthropomorpher Merkmale, eine soziale Beziehung zum Nutzer aufbauen.

Milde (2000) plant die Umsetzung eines teilautonomen Agenten, der in intelligenten, virtuellen Umgebungen natürlichsprachlich mit dem Benutzer kommuniziert. Die Diskussion um diese sogenannten Software Agenten zeigt ebenfalls, daß der Einfluß komplexer, als soziale Wesen akzeptierter Systeme wachsen wird.

Software Agenten sind autonome, adaptive und proaktive Programme: Sie haben zu einem gewissen Grad Kenntnis von den Interessen der Nutzer, passen sich diesen Interessen an und werden zu deren Vorteil selbständig aktiv, z.B. bei der Suche im Netz (Shneiderman und Maes, 1997). Ihre Entwicklung erfolgte unter dem Eindruck der sich ändernden Anforderungen an Computerprogramme, durch die Verschiebung der Nutzerschichten, weg von den Experten, hin zu breiten Anwenderkreisen ohne spezielle Ausbildung (s. a. 2.1). Software Agenten sind daher explizit gedacht für unerfahrene Nutzer, die erfolgreich in den stark wachsenden, dynamischen und unstrukturierten Datenbeständen suchen wollen, die verteilt im Netz vorliegen.

Software Agenten werden von Vertretern der Direct Manipulation Interfaces (als deren prominentester Shneiderman gelten kann) heftig kritisiert, weil sie häufig mit einer anthropomorphen Oberfläche, z.B. in Form von Avataren, einhergehen. Dadurch kommt es nach Meinung von z.B. Shneiderman zu Mißverständnissen auf Seiten der Nutzer, die dem System mehr Funktionalität zuschreiben würden, als dieses zur Verfügung stellen könne – zudem würden solche Schnittstellen die Verantwortlichkeit des Nutzers für sein Handeln und seine Kontrollmöglichkeiten über das System reduzieren.

Maes hält dieser Argumentation allerdings entgegen (Shneiderman und Maes, 1997), daß Software Agenten nicht notwendigerweise eine anthropomorphe Oberfläche haben müssen, sondern sehr gut auch mit einer Direct Manipulation Oberfläche kombiniert werden können. Jedes System – auch ein Software Agent – brauche eine gute Oberfläche. Die Art der Schnittstelle werde nicht durch das zugrundeliegende System bestimmt, sondern müsse dem Datenbestand und der Zielgruppe angemessen sein. Wenn es sich um erfahrene Nutzer handle, die in einer überschaubaren, wohlstrukturierten und daher gut visualisierbaren Domäne suchten, seien Direct Manipulation Oberflächen auch für Software Agenten gut geeignet.

Ähnlich wie bei der Unterscheidung für den Einsatz linguistischer Analysemethoden im Retrieval System selbst und an der Schnittstelle zum Benutzer (s. 6.2), zeigt sich auch hier die notwendige Trennung zwischen den systeminternen Charakteristika, wie

Proaktivität und Autonomie, und den Anforderungen an die Schnittstelle. Dies ist wiederum eine Bestärkung der bereits vertretenen Ansicht, daß die generelle Diskussion über den Einsatz von z.B. natürlicher Sprache verfehlt ist. Es muß deutlich unterschieden werden zwischen der Funktionalität des Kernsystems und der Gestaltung der Schnittstelle. Dabei müssen Entscheidungen, die für das Kernsystem getroffen werden (z.B. linguistische Analyse zu betreiben oder proaktiv mit Hilfe von Lernverfahren zu sein), nicht zwangsläufig Auswirkungen auf die Oberfläche haben. Aus den Untersuchungen von Nass et al., Lester et al. oder Walker et al. ist aber auch ersichtlich, daß die Behandlung von Computern als soziale Wesen offenbar einer Wirkung des Gesamtsystems entspringt, und nicht einzelnen Merkmalen der Schnittstelle zuzuschreiben ist. In diesem Sinne ist eine Argumentation, wie sie z.B. Shneiderman gegen anthropomorphe oder auch unbeschränkt natürlichsprachliche Schnittstellen führt, verfehlt. Unter dem Aspekt der Systemwirkung macht es offenbar keinen Sinn, zwischen verschiedenen Arten von Schnittstellen zu unterscheiden, sondern es zählt der gesamte Umfang der Systemfunktionalität und die insgesamt wahrgenommenen Systemattribute.

Angesichts der in 2.1 beschriebenen wachsenden Bedeutung von Online Informationen und den wachsenden Datenmengen ist hier aber eine Entscheidung längst gefallen: In Zukunft wird, aufgrund steigender Anforderungen an die Systeme, deren Funktionalität weiter zunehmen. Die neue Rolle von Computern als soziale Agenten muß deshalb von den Systementwicklern wie von den Schnittstellenexperten zur Kenntnis genommen werden. Dabei steht, angesichts des deutlich anderen, neuen Status von Computern zwischen Werkzeugen und Mitmenschen, das Fernziel der Entwicklung natürlichsprachlicher Schnittstellen neu zur Diskussion: Eine natürlichsprachliche Schnittstelle mit unbeschränkter Eingabemöglichkeit, die explizit die zwischenmenschliche Kommunikation zum Modell nimmt, scheint vor diesem Hintergrund noch schwieriger realisierbar zu sein als bislang angenommen. Alle Vorstufen auf dem Weg zu so einer unbeschränkten Schnittstelle,¹²³ werden vermutlich, auch bei steigender Perfektion, auf deutliche Ablehnung durch den Benutzer stoßen: Menschen unterscheiden in der Kommunikation offenbar sehr genau zwischen Menschen und Maschinen und sind nicht bereit, beide gleich zu behandeln.

Eingeschränkt natürlichsprachliche Schnittstellen, die deutlich ihren eigenständigen Charakter als Maschine betonen, scheinen daher geeigneter, für die Benutzer akzeptable und verständliche Schnittstellen zu sein. Allerdings ist auch für eingeschränkt natürlichsprachliche Schnittstellen (wie ausführlich in 4 und 5 dargestellt), weitere Entwicklung nötig. An der generellen Richtung dieser Entwicklung (größerer Sprachumfang, mehr Semantik, Anreicherung um ontologische Informationen) ändert sich aber nichts: Auch ohne den Anspruch, ein mit dem menschlichen Gesprächspartner verwechselbares System bauen zu wollen, muß der verfügbare Sprachumfang der Systeme ausgebaut werden, um beispielsweise die Fehlerrate bei der Analyse der Nutzereingaben zu minimieren.

¹²³Ob überhaupt jemals Maschinen gebaut werden, die dieses Ziel erreichen, sei dahingestellt.

In diesem Sinne sind wir heute doch einen Schritt weiter als Waltz dies 1982 war: Die Systeme haben sich verändert und mit ihnen ihre Wirkung auf den Benutzer. Wir haben diese Wirkung identifiziert und können uns darauf einstellen.

Literatur

- R. Agarwal. Towards a PURE Spoken Dialogue System for Information Access. In: J. Hirschberg, Hrsg., *Proceedings of the ACL/EACL Workshop on Interactive Spoken Dialog Systems: Bringing Speech and NLP Together in Real Applications*, Seiten 90–97, Sommerset, NJ, 1997. Universidad Nacional de Educación a Distancia, Madrid, Spain, ACL.
- C. Ahlberg und B. Shneiderman. Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In: *Proceedings of CHI 94*, Seiten 313–317, Boston, MA USA, 4 1994. ACM.
- J. Alexandersson. Some Ideas for the Automatic Acquisition of Dialogue Structure. In: S. LuperFoy, A. Nijholt, und G. Veldhuijzen van Zanten, Hrsgg., *Dialogue Management in Natural Language Systems. Proceedings of the 11th Twente Workshop on Language Technology (TWLT 11)*, Seiten 149–158, Enschede, 1996. University of Twente.
- J. Alexandersson, B. Buschbeck-Wolf, T. Fujinami, E. Maier, N. Reithinger, B. Schmitz, und M. Siegel. Dialogue Acts in VERBMOBIL-2. Technical Report 204, DFKI Saarbrücken, Universität Stuttgart, Technische Universität Berlin, Universität des Saarlandes, 5 1997a.
- J. Alexandersson, E. Maier, und N. Reithinger. A Robust and Efficient Three-Layered Dialog Component for a Speech-to-Speech Translation System. Technical Report 50, DFKI GmbH, 12 1994.
- J. Alexandersson, N. Reithinger, und E. Maier. Insights into the Dialogue Processing of VERBMOBIL. Technical Report 191, DFKI Saarbrücken, 1997b.
- J. Allen und C.R. Perrault. Analyzing Intention in Utterances. *Artificial Intelligence*, 15(3):143–178, 1980.
- A. Altmann. Direkte Manipulation: Empirische Befunde zum Einfluß der Benutzeroberfläche auf die Erlernbarkeit von Textsystemen. *Zeitschrift für Arbeits- und Organisationspsychologie*, 31(3):108–114, 1987.
- R. Amant und P. Cohen. Interaction with a Mixed-Initiative System for Exploratory Data Analysis. In: *Proceedings of IUI 97*, Seiten 15–22, Orlando, FL USA, 1997. ACM.
- E. André und T. Rist. Von Textgeneratoren zu Intellimedia-Präsentationssystemen. *Künstliche Intelligenz. Organ des Fachbereichs 1 Künstliche Intelligenz der Gesellschaft für Informatik e.V.*, 7(2):40–48, 1993.

- J.-M. Andreoli, U. Borghoff, und R. Pareschi. Signed Feature Constraint Solving. In: *Proceedings of the 3rd International Conference on the Practical Application of Constraint Technology (PACT '97)*, Seiten 35–46, Blackpool, UK, April 23–25 1997a. The Practical Application Ltd.
- J.-M. Andreoli, U. M. Borghoff, R. Pareschi, S. Bistarelli, U. Montanari, und F. Rossi. Constraints and Agents for a Decentralized Network Infrastructure. In: E. C. Freuder, Hrsg., *Proc. Int'l. AAAI-97 Workshop on Constraints and Agents*, Seiten 39–44, Menlo Park, CA, July 27 1997b. AAAI Press.
- I. Androutsopoulos, G. Ritchie, und P. Thanisch. Natural Language Interfaces to Databases – An Introduction. *Journal of Language Engineering*, 1(1):29–81, 1995.
- N. Anikina, V. Golender, S. Kozhukhina, L. Vainer, und B. Zagatsky. REASON: NLP-based Search System for the WWW. In: Kavi Mahesh, Hrsg., *Natural Language Processing for the World Wide Web. Papers from the 1997 AAAI Spring Symposium*, Stanford, USA, 3 1997. AAAI.
- N. Asher und A. Lascarides. Questions in Dialogue. *Linguistics and Philosophy*, 21 (4):237–309, 1998.
- D. Austin. Automatisierung in der Sacherschließung der British Library. *Bibliothek: Forschung und Praxis*, 8(1):45–57, 1984.
- J.L. Austin. *How to Do Things With Words*. Oxford University Press, Oxford, 1962.
- G. Ball und J. Breese. Modeling the Emotional State of Computer Users. In: 'Attitude, Personality and Emotions in User-Adapted Interaction.' *Proceedings of the UM'99 Workshop*, 1999.
- S. Banerjee und V.O. Mittal. On the Use of Linguistic Ontologies for Accessing and Indexing Distributed Digital Libraries. In: *Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*, 1994.
- M. Barja, T. Bratvold, J. Myllymaki, und G. Sonnenberger. Informia: a mediator for integrated access to heterogeneous information sources. In: *CIKM '98. Proceedings of the 1998 ACM 7th international conference on information and knowledge management*, Seiten 234–241. ACM, 1998.
- A. Barth, M. Breu, A. Endres, und A. de Kemp. *Digital Libraries in Computer Science: The MeDoc Approach*, Vol. 1392: *Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, New York, 1998.
- M. Bates. There is still gold in the database mine. *COLING-84*, Seiten 184–185, 1984.

- R. Bäuerle und T. Zimmermann. Fragesätze. In: A. Stechow, Hrsg., *Semantik: ein internationales Handbuch der zeitgenössischen Forschung*, Vol. 6: *Handbücher zur Sprach- und Kommunikationswissenschaft*, Kapitel 15, Seiten 333–348. de Gruyter, Berlin, 1991.
- R. Bayer, P. Vogel, und S. Wiesener. OMNIS/Myriad Document Retrieval and Its Database Requirements. In: D. Karagiannis, Hrsg., *Database and Expert Systems Applications – Proceedings of DEXA '94*, Vol. 856: *Lecture Notes in Computer Science*, Seiten 550–564, Berlin, 1994. Springer.
- M. Beaulieu. Experiments on Interfaces to Support Query Expansion. *Journal of Documentation*, 53:8–19, 1997.
- J. E. Bell und L. A. Rowe. An Exploratory Study of Ad Hoc Query Languages to Databases. In: F. Golshani, Hrsg., *8th International Conference on Data Engineering*, Seiten 606–613, Los Alamitos, Ca., USA, February 1992. IEEE Computer Society Press.
- T.C. Bell, A. Moffat, und I.H. Witten. Compressing the Digital Library. In: *Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*, 1994.
- R. Beun. The DENK Program: Modelling Pragmatics in Natural Language Interfaces. In: *Proceedings of the Twente Workshop on Language Technology (TWLT5)*, Seiten 75–79, University of Twente, 1993.
- A.W. Biermann, B.W. Ballard, und A.H. Sigmon. An Experimental Study of Natural Language Programming. *International Journal of Man-Machine Studies*, 18(1):71–87, 1983.
- J.-L. Binot, L. Debille, D. Sedlock, und B. Vandecapelle. Natural Language Interfaces: A New Philosophy. *SunExpert Magazine*, January 1991.
- W.P. Birmingham, K.M. Drabenstott, C.O. Frost, A.J. Warner, und K. Willis. The University of Michigan Digital Library: This is Not Your Father's Library. In: *Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*, 1994.
- A. Black, J. van de Plassche, und Briony Williams. Analysis of Unknown Words through Morphological Decomposition. In: *Proceedings of the 5th Conference of the European Chapter of the Association for Computational Linguistics*, Seiten 101–106, Morristown, NJ, 4 1991. Association for Computational Linguistics.
- N. Borenstein. Is English a Natural Language. In: K. Hopper und I.A. Newman, Hrsgg., *Foundation for human-computer communication*, Vol. 3: *IFIP working conference on command languages*, Seiten 60–72, Amsterdam, 1986. International Federation for Information Processing, North-Holland.

- U. M. Borghoff, P.-Y. Chevalier, und J. Willamowski. Adaptive Refinement of Search Patterns for Distributed Information Gathering. In: A. Verbraeck, Hrsg., *Proceedings of the 1st International Conference EuroMedia/WEBTEC '96*, Seiten 5–12, San Diego, CA, December 19-21 1996a. The Society for Computer Simulation.
- U. M. Borghoff, E. R. Hilf, R. Pareschi, T. Severiens, Stamerjohanns, und J. H., Willamowski. Agent-Based Document retrieval for the European Physicists: A Project Overview. In: *Proceedings of the 2nd International Conference on Practical Applications of Intelligent Agents and Multi-Agent Technology (PAAM '97)*, Seiten 271–285, London, UK, April 1997. The Practical Application Company Ltd.
- U. M. Borghoff, R. Pareschi, H. Karch, M. Nöhmeier, und J. H. Schlichter. Constraint-Based Information Gathering for a Network Publication System. In: *Proceedings of the 1st International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology (PAAM'96)*, Seiten 45–59, London, UK, April 22-24 1996b. The Practical Application Company Ltd.
- C. Borgman. Mental Models: Ways of Looking at a System. *Bulletin of the American Society for Information Science*, 9(2):38–39, 1982.
- C. Borgman. Why are online catalogs hard to use? Lessons learned from information retrieval studies. *Journal of the American Society for Information Science*, 37(6): 387–400, 1986.
- P. Bosch. *Agreement and Anaphora: A Study of the Role of Pronouns in Syntax and Discourse*. Cognitive Science Series. Academic Press, London, 1983.
- N. Brinkhoff. Gegenwart und Zukunft der EU-Förderung von Sprachtechnologien (5. Forschungsrahmenprogramm). In: *Forum Sprache ohne Grenzen*, Seiten 155–163. DLR, Projektträger Informationstechnik, 1997.
- J.S. Brown, R.R. Burton, und A.G. Bell. SOPHIE: A Step Toward Creating a Reactive Learning Environment. *International Journal of Man-Machine Studies*, 7(5):675–696, 1975.
- T. Brückner. The text categorization system TEKLIS at TREC-6. In: Ellen M. Voorhees und Donna K. Harman, Hrsg., *Proceedings of TREC-6, 6th Text Retrieval Conference*, Seiten 619–621, Gaithersburg, USA, 1997. National Institute of Standards and Technology, Gaithersburg, USA.
- H. Bunt, R. Ahn, R.-J. Beun, und T. Borghuis. Multimodal Cooperation with the DenK System. In: H. Bunt, Hrsg., *Multimodal human-computer communication: systems, techniques, and experiments*, Vol. 1374: *Lecture Notes in Computer Science*, Seiten 39–67, Berlin, 5 1998. International Conference on Cooperative Multimodal Communication (CMC), 1995, Springer.

- A. Burton und A.P. Steward. Effects of Linguistic Sophistication on the Usability of a Natural Language Interface. *Interacting with Computers*, 5(1):31–59, 1993.
- E. Busch. Das Problem der VP-Ellipsen im Rahmen einer Theorie der Sachverhaltsanaphern. Technical Report 123, Institut für Wissensbasierte Systeme, IBM Deutschland GmbH, Stuttgart, 1990.
- B. Buxton, M. Krueger, B. Laurel, und L. Vertelny. Drama and Personality in User Interface Design (Panel). In: *Proceedings of ACM CHI'89 Conference on Human Factors in Computing Systems*, Interacting with Computer Advisors, Seiten 105–108, 1989.
- J.P. Callan und W.B. Croft. An Evaluation of Query Processing Strategies Using the TIPSTER Collection. In: *Proceedings of the Sixteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Query Processing and Evaluation, Seiten 347–355, 1993.
- R. Capindale und R. Crawford. Using a Natural Language Interface with Casual Users. *International Journal of Man-Machine Studies*, 32(3):341–361, 1990.
- S. Carberry, C. Kambhamettu, und L. Schroeder. Gestural Evidence and the Recognition of Attitude and Intention. In: *'Attitude, Personality and Emotions in User-Adapted Interaction.'* *Proceedings of the UM'99 Workshop*, 1999.
- S. Carberry und L. Lambert. A Process Model for Recognizing Communicative Acts and Modeling Negotiation Subdialogues. *Computational Linguistics*, 25(1):1–53, 1999.
- J.G. Carbonell. Is there natural language after databases? *COLING-84*, Seiten 186–187, 1984.
- J.G. Carbonell, W.M. Boggs, M.L. Mauldin, und P.G. Anick. The XCALIBUR Project: A Natural Language Interface to Expert Systems. In: A. Bundy, Hrsg., *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Seiten 653–656, Karlsruhe, August 1983. William Kaufmann.
- T. Catarci, S.K. Chang, M.F. Costabile, S. Levialdi, und G. Santucci. A Multiparadigmatic Visual Environment for Adaptive Access to Databases. In: *Proceedings of ACM INTERCHI'93 Conference on Human Factors in Computing Systems – Adjunct Proceedings*, Short Papers (Talks): Visual Languages and Virtual Reality, Seiten 7–8, 1993.
- A. Chakravarthy und K. Haase. NetSerf: using semantic knowledge to find Internet information archives. In: *SIGIR '95. Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval*, Seiten 4–11. ACM, 1995.

- J. Chapman. A State Transition Analysis of Online Information-Seeking Behaviour. *Journal of the American Society for Information Science*, 32(9):325–333, 1981.
- E. Charniak. Jack and Jane in search of a theory of knowledge. In: B.J. Sparck-Jones, K. Grosz und B.L. Webber, Hrsgg., *Readings in Natural Language Processing*, Seiten 331–337. Morgan Kaufmann, Los Altos, 1986.
- B. Chidlovskii und U. M. Borghoff. Query Translation for Distributed Information Gathering on the Web. In: B. Eaglestone, B. C. Desai, und J. Shao, Hrsgg., *Proceedings 2nd IEEE International Database Engineering and Applications Symposium (IDEAS '98)*, Seiten 214–223, Los Alamitos, CA, July 8-10 1998. IEEE Computer Society Press.
- B. Chidlovskii, U. M. Borghoff, und P.-Y. Chevalier. Boolean Query Translation for Brokerage on the Web. In: A. Verbraeck, Hrsg., *Proceedings of the 2nd International Conference EuroMedia/WEBTEC '98*, Seiten 37–44, San Diego, CA, January 5-7 1998. The Society for Computer Simulation.
- D. Chin. An Analysis of Scripts Generated in Writing Between Users and Computer Consultants. *AFIPS Conference Proceedings*, 53:637–642, 1984.
- M. Chodorow, R. Byrd, und G. Heidorn. Extracting semantic hierarchies from a large on-line dictionary. In: *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, Seiten 299–304, Chicago, Ill., 1985. Association for Computational Linguistics, University of Chicago.
- G. Churcher, E. Atwell, und C. Souter. Dialogue Management Systems: a Survey and Overview. Technical Report 97.06, School of Computer Studies, Univ. of Leeds, 2 1997.
- C. Cleverdon. Optimizing convenient online access to bibliographic databases. *Information Services & Use*, 4:37–47, 1984.
- P. Cohen. *Survey of the State of the Art of Human Language Technology*, Kapitel 6.3: Dialogue Modelling. Cambridge University Press, Cambridge, MA, 1996.
- P. Cohen und H. Levesque. Rational Interaction as the Basis for Communication. In: P. Cohen, J. Morgan, und M.E. Pollack, Hrsgg., *Intentions in Communication*, System Development Foundation Benchmarks Series. Bradford, Cambridge, MA, 1990.
- P.R. Cohen, M. Dalrymple, D.B. Moran, F.C.N. Pereira, J.W. Sullivan, R.A. Gargan, J.L. Schlossberg, und S.W. Tyler. Synergistic Use of Direct Manipulation and Natural Language. In: *Proceedings of ACM CHI'89 Conference on Human Factors in Computing Systems, Gesture and Language*, Seiten 227–233, 1989.

- R. Comino, R. Gemello, G. Guida, C. Rullent, L. Sisto, und M. Somalvico. Understanding Natural Language Through Parallel Processing of Syntactic and Semantic Knowledge: An Application to Data Base Query. In: Alan Bundy, Hrsg., *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Seiten 663–667, Karlsruhe, August 1983. William Kaufmann.
- A. Copestake und K. Sparck-Jones. Natural Language Interfaces to Databases. *The Knowledge Engineering Review*, 5(4):225–249, 1990.
- A. Cremers. Object Reference in Task-Oriented Keyboard Dialogues. In: H. Bunt, Hrsg., *Multimodal human-computer communication: systems, techniques, and experiments*, Vol. 1374: *Lecture Notes in Computer Science*, Seiten 279–293, Berlin, 5 1998. International Conference on Cooperative Multimodal Communication (CMC), 1995, Springer.
- I. Dagan und K.W. Church. TERMIGHT: Identifying and Translating Technical Terminology. In: *Proceedings of the Fourth ACL Conference on Applied Natural Language Processing (13–15 October 1994, Stuttgart)*. Association for Computational Linguistics, October 1994.
- V. Dahl und M.C. McCord. Treating coordination in logic grammars. *American Journal of Computational Linguistics*, 9(2):69–91, 1983.
- N. Dählback und A. Jönsson. Empirical Studies of Discourse Representations for Natural Language Interfaces. In: *Fourth Conference of the European Chapter of the Association for Computational Linguistics (EACL'89)*, Seiten 291–298. European Chapter of the Association for Computational Linguistics (EACL), Association for Computational Linguistics, 1989.
- M. Dalrymple, S. Shieber, und F. Pereira. Ellipsis and Higher-Order Unification. *Linguistics and Philosophy*, 14:399–452, 1991.
- P. Dalrymple und R. Cox. An examination of the effects of non-boolean enhancements to an information retrieval system. In: M. Williams, Hrsg., *Proceedings of the 13th National Online Meeting*, Seiten 75–81, Medford, NJ, 1992. Learned Information.
- P. Das-Gupta. Boolean Interpretation of Conjunctions for Document Retrieval. *Journal of the American Society for Information Science*, 38(4):245–254, 1987.
- C. Davenport und G. Weir. Plan Recognition for Intelligent Advice and Monitoring. In: M. D. Harrison und A. F. Monk, Hrsgg., *People and Computers: Designing for Usability*, Seiten 296–315. Cambridge University Press, Cambridge, U.K., 1986.
- A. De Angeli, W. Gerbino, E. Nodari, und D. Petrelli. From tools to friends: where is the borderline? In: *'Attitude, Personality and Emotions in User-Adapted Interaction.'* *Proceedings of the UM'99 Workshop*, 1999.

- A. De Angeli und D. Petrelli. Bridging the gap between NLP and HCI: A new synergy in the name of the user. In: *Proceedings of the CHI 2000 Workshop on Natural Language Interfaces*, 4 2000.
- L. de Lima, A. Laender, und B. Ribeiro-Neto. A hierarchical approach to the automatic categorization of medical documents. In: *CIKM '98. Proceedings of the 1998 ACM 7th international conference on information and knowledge management*, Seiten 132–139. ACM, 1998.
- A. De Roeck, R. Ball, K. Brown, C. Fox, M. Groefsema, N. Obeid, und R. Turner. Helpful Answers to Modal and Hypothetical Questions. In: *Fifth Conference of the European Chapter of the Association for Computational Linguistics*, Seiten 257–262, Morristown, NJ, 1991. Association for Computational Linguistics.
- M. Debevc, B. Meyer, und R. Svecko. An Adaptive Short List for Documents on the World Wide Web. In: *Proceedings of IUI 97*, Seiten 209–211, Orlando, FL, CA USA, 1997. ACM.
- S.M. Dekleva. Is Natural Language Querying Practical? *Data Base*, Seiten 24–36, May 1994.
- S. Detemple. PRECIS. Ein computerunterstütztes System zur Herstellung alphabetischer Sachregister und Schlagwortkataloge. *Bibliothek: Forschung und Praxis*, 6 (1/2):4–46, 1982.
- D. Diaper. Identifying the Knowledge Requirements of an Expert System's Natural Language Processing Interface. In: M. D. Harrison und A. F. Monk, Hrsgg., *People and Computers: Designing for Usability*, Seiten 263–280. Cambridge University Press, Cambridge, U.K., 1986.
- J. Dickson. An Analysis of User Errors in Searching an Online Catalog. *Cataloging and Classification Quarterly*, 4(3):19–38, 1984.
- C. Djeraba und M. Bouet. Digital Information Retrieval. In: *Proceedings of CIKM 97*, Seiten 185–192, Las Vegas, 1997. ACM.
- G. Dreis. *Benutzerverhalten an einem Online-Publikumskatalog für wissenschaftliche Bibliotheken: Ergebnisse und Erfahrungen aus dem OPAC-Projekt der Universitätsbibliothek Düsseldorf*. No. 57 in *Zeitschrift für Bibliothekswesen und Bibliographie*, Sonderhefte. Klostermann, 1994.
- H. Dreyfus. *What Computers Can't Do: The Limits on Artificial Intelligence*. Harper and Row, New York, 1979.
- V. Droste, W. Paulus, und K. Weishaupt. Der Bibliotheks-Goliath Aleph 500 im Kampf gegen BIS-LOK, den bibliothekarischen David. *Bibliotheksdienst*, 32(11):1959–1965, 1998.

- D. Dryer. Wizards, Guides, and Beyond: Rational and Empirical Methods for Selecting Optimal Intelligent User Interface Agents. In: *Proceedings of IUI 97*, Seiten 265–268, Orlando, FL USA, 1997. ACM.
- S.T. Dumais, T.K. Landauer, und M.L. Littman. Automatic Cross-Linguistic Information Retrieval using Latent Semantic Indexing. In: G. Grefenstette, Hrsg., *Working Notes of the Workshop on Cross-Linguistic Information Retrieval*. ACM SIGIR, August 1996.
- S.T. Dumais, T.A. Letsche, M.L. Littman, und T.K. Landauer. Automatic Cross-Language Retrieval Using Latent Semantic Indexing. In: *AAAI Symposium on Cross-Language Text and Speech Retrieval*. American Association for Artificial Intelligence, March 1997.
- H. Dybkjær, L. Dybkjær, und N. Bernsen. Design, formalization and evaluation of spoken language dialogue. In: J. Andernach, S. van de Burgt, und G. van der Hoeven, Hrsgg., *Corpus-based Approaches to Dialogue Modelling. Proceedings of the 9th Twente Workshop on Language Technology (TWLT9)*, University of Twente, 1995.
- K. Eckert und A. Schulz. SABINE: OPAC oder opak? Kein Durchblick beim neuen Online Public Access Catalogue der Universität des Saarlandes. *Bibliotheksdienst*, 29(6):979–984, 1995.
- E.N. Efthimiadis. A User-Centred Evaluation of Ranking Algorithms for Interactive Query Expansion. In: *Proceedings of the Sixteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Query Expansion, Seiten 146–159, 1993.
- T. Einsporn. *Strategische Nutzung des Wirtschaftsfaktors Information am Beispiel von 2500 mittelständischen Unternehmen*. No. 202=4/1995 in Beiträge zur Gesellschafts- und Bildungspolitik. Deutscher Instituts-Verlag, Köln, 1995.
- D. Embley, D. Campbell, R. Smith, und S. Liddle. Ontology-based extraction and structuring of information from data-rich unstructured documents. In: *CIKM '98. Proceedings of the 1998 ACM 7th international conference on information and knowledge management*, Seiten 52–59. ACM, 1998.
- A. Endres und N. Fuhr. Student Access Books and Journals through MeDoc. *ACM Communications*, 41(4):76–77, 1998.
- S. Ervin-Tripp. Children's Verbal Turn-Taking. In: E. Ochs und B. B. Schieffelin, Hrsgg., *Developmental Pragmatics*, Seiten 391–414. Academic Press, New York, 1979.
- B. Eugenio, J. Moore, P. Jordan, und R. Thomason. An Empirical Investigation of Proposals in Collaborative Dialogues. In: *Proceedings of the 36th Annual Meeting*

- of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, Seiten 325–329, Los Altos, CA, 1998. Morgan Kaufmann Publishers.
- C. Fenichel. Online Searching: Measures that Discriminate among Users with Different Types of Experiences. *Journal of the American Society for Information Science*, 32(1):23–32, 1981.
- FhG-ISI, Hrsg. *Delphi: Studie zur globalen Entwicklung von Wissenschaft und Technik*. Fraunhofer-Institut für Systemtechnik und Innovationsforschung im Auftrag des Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie (BMBF), FhG-ISI, Karlsruhe, 1998.
- D.P. Flickinger. Panel on natural language and databases. *COLING-84*, Seiten 188–189, 1984.
- V. Florance und G. Marchionini. Information Processing in the Context of Medical Care. In: *Proceedings of the Eighteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, User Studies, Seiten 158–163, 1995.
- C. Fluhr, D. Schmit, F. Elkateb, P. Ortet, und K. Gurtner. Multilingual Database and Crosslingual Interrogation in a Real Internet Application. In: *AAAI Symposium on Cross-Language Text and Speech Retrieval*. American Association for Artificial Intelligence, March 1997.
- C. Fluhr, D. Schmit, P. Ortet, F. Elkateb, K. Gurtner, und V. Semanova. Distributed Multilingual Information Retrieval. In: *Proceedings of the first Workshop on Multilinguality in Software Engineering: The AI Contribution (MULSAIC)*. European Coordinating Committee for Artificial Intelligence, August 1996.
- E. Fox. From Information Retrieval to Networked Multimedia Information Access. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 116–124, Konstanz, 1993. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- E.A. Fox, R.K. France, E. Sahle, A. Daoud, und B.E. Cline. Development of a Modern OPAC: From REVTOLC to MARIAN. In: *Proceedings of the Sixteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Interface Issues, Seiten 248–259, 1993.
- K. Franzén und J. Karlgren. Project Presentation REPTILE Retrieval, Extraction, Presentation and Translation using Language Engineering. In: *AAAI Symposium on Cross-Language Text and Speech Retrieval*. American Association for Artificial Intelligence, March 1997.

- C. Frappaolo. Artificial Intelligence and Text Retrieval: A Current Perspective on the State-of-the-art. In: M. Williams, Hrsg., *Proceedings of the 13th National Online Meeting*, Seiten 113–114, Medford, NJ, 1992a. Learned Information.
- C. Frappaolo. What do users want from an online system? In: M. Williams, Hrsg., *Proceedings of the 13th National Online Meeting*, Seiten 115–116, Medford, NJ, 1992b. Learned Information.
- H.-P. Frei und Y. Qiu. Effectiveness of Weighted Searching in an Operational IR Environment. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 41–54, Konstanz, 1993. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- W. Frey, U. Reyle, und C. Rohrer. Automatic Construction of a Knowledge Base by Analysing Texts in Natural Language. In: A. Bundy, Hrsg., *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Seiten 727–729, Karlsruhe, August 1983. William Kaufmann.
- R.P. Futrelle und X. Zhang. Large-Scale Persistent Object Systems for Corpus Linguistics and Information Retrieval. In: *Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*, 1994.
- B.R. Gaines und M.L.G. Shaw. Dialog Engineering. In: M.E. Sime und M.J. Coombs, Hrsgg., *Designing for Human-Computer Communication*, Computers and People, Seiten 23–53. Academic Press, New York, NY, 1983.
- P. Gallmann. Fugemorpheme als Nicht-Kasus-Suffixe. In: M. Butt und N. Fuhrhop, Hrsgg., *Variation und Stabilität in der Wortstruktur*, Germanistische Linguistik, Seiten 73–86. Olms, Hildesheim, Zürich, New York, 1999.
- B. Gattung. Der OPAC der Universitätsbibliothek Düsseldorf – Benutzerperspektive. In: E. Plassmann, Hrsg., *Bibliotheken in Europa. 80. Deutscher Bibliothekartag in Saarbrücken 1990.*, No. 53; Sonderheft in Zeitschrift für Bibliothekswesen und Bibliographie, Seiten 103–108, Frankfurt a.M., 1991. Klostermann.
- G. Gediga und K.-C. Hamborg. Heuristische Evaluation und IsoMetrics: Ein Vergleich. In: R. Liskowsky, B. Velichkovsky, und W. Wüschmann, Hrsgg., *Software-Ergonomie '97: Usability Engineering: Integration von Mensch-Computer-Interaktion und Software-Entwicklung*, Vol. 49: *Berichte des German Chapter of the ACM*, Seiten 145–155, Stuttgart, 3 1997. German Chapter of the ACM, Gesellschaft für Informatik (GI), TU Dresden, Teubner.
- H. Geser. Der PC als Interaktionspartner. *Zeitschrift für Soziologie*, 18(3):230–243, Juni 1989.

- B. Geurts. Natural Language Understanding in LILOG: An Intermediate Overview. Technical Report 137, Institut für Wissensbasierte Systeme, IBM Deutschland GmbH, 1990a.
- B. Geurts. Constraints on presupposition projection. In: R.A. Van der Sandt, Hrsg., *Presupposition and Anaphora*, Lisbon, 1993. Fifth European summer school in logic, language, and information. Workshop Reader.
- B. Geurts. *Presupposing*. PhD thesis, Universität Stuttgart, Stuttgart, 1995.
- B. Geurts. Triggers. In: *QUD WWW-Homepage*, <http://turing.wins.uva.nl/~pdekker/QUD.html>, 7 1997a.
- B. Geurts. Why the description theory of names won't go away. 7 1997b.
- B. Geurts. The mechanisms of denial. *Language*, Seiten 274–307, 1998.
- B. Geurts und R. Van der Sandt. Domain Restriction. In: P. Bosch und R. Van der Sandt, Hrsgg., *Focus: linguistic, cognitive, and computational perspectives*. Cambridge University Press, Cambridge, 1997.
- J. Gilarranz, J. Gonzalo, und F. Verdejo. An approach to Conceptual Text Retrieval Using the EuroWordNet Multilingual Semantic Database. In: *AAAI Symposium on Cross-Language Text and Speech retrieval*. American Association for Artificial Intelligence, March 1997.
- J. Ginzburg. Dynamics and the semantics of Dialogue. In: J. Seligman, Hrsgg., *Language, Logic and Computation, Volume 1*, CSLI Lecture Notes. CSLI, Stanford, 1996.
- A. Glöckner-Rist. Suche und Du wirst vielleicht finden: Die Formulierung von Suchproblemen und ihre Transformation in Suchfragen. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 183–198, Konstanz, 1993a. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- A. Glöckner-Rist. *Suchfragen im Information Retrieval: eine empirische Untersuchung zum Rechercheverhalten von Informationsvermittlern und Endbenutzern*, Vol. 8: *Schriften zur Informationswissenschaft*. Universitäts-Verlag, Konstanz, 1993b.
- A. Glöckner-Rist, W. Lehmler, und M. Wettler. Endnutzerrecherchen in bibliographischen CD-ROM-Datenbanken. *nfd Information – Wissenschaft und Praxis*, 40: 151–159, 1989.
- W. Gödert und M. Liebig. Maschinelle Indexierung auf dem Prüfstand. Ergebnisse eines Retrievaltests zum MILOS II Projekt. *Bibliotheksdienst*, 31(1):59–68, 1997.

- L. Gong. What makes a Natural Language Interface natural? Consistency as a Social Rule. In: *Proceedings of the CHI 2000 Workshop on Natural Language Interfaces*, 4 2000.
- A. Grasso, M. Koch, und D. Snowdon. Campiello – New user interface approaches for community networks. In: *Proc. Workshop Designing Across Borders: The Community Design of Community Networks (PDC98/CSCW98)*, Seattle, WA, November 1998.
- G. Grewendorf, F. Hamm, und W. Sternefeld. *Sprachliches Wissen: Eine Einführung in moderne Theorien der grammatischen Beschreibung*. No. 695 in suhrkamp taschenbuch wissenschaft. Suhrkamp, Frankfurt a.M., 1988.
- H. Grice. *Studies in the Way of Words*, Kapitel 'Logic and conversation', Seiten 22–57. Harvard University Press, Cambridge MA, 1989.
- A. Grieszl. Ziele, Pläne oder Topics? Diskursstrukturtheorien – ein Vergleich. Master's thesis, Universität Osnabrück, Institut für Semantische Informationsverarbeitung, 3 1995.
- J.-M. Griffiths und K.K. Kertis. Access to Large Digital Libraries of Scientific Information Across Networks. In: *Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*, 1994.
- J. Groenendijk und M. Stokhof. Questions. In: J. van Benthem und A. ter Meulen, Hrsgg., *Handbook of Logic and Language*, Kapitel 19, Seiten 1055–1124. Elsevier, Amsterdam, 1997.
- B. Grosz. *Survey of the State of the Art of Human Language Technology*, Kapitel 6.1: Overview. Cambridge University Press, Cambridge, MA, 1996.
- B. J. Grosz und C. L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.
- B.J. Grosz. The representation and use of focus in a system for understanding dialogs. In: K. Sparck-Jones, B.J. Grosz, und B.L. Webber, Hrsgg., *Readings in Natural Language Processing*, Seiten 353–362. Morgan Kaufmann, Los Altos, 1986.
- J. Gu, U. Thiel, und J. Zhao. Efficient Retrieval of Complex Objects: Query Processing in a Hybrid DB and IR System. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 67–81, Konstanz, 1993. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- R. Guindon, Hrsg. *Cognitive science and its applications for human-computer interaction*. Lawrence Erlbaum Associates, Hillsdale, N.J., 1988.

- J. Gustafson. How do system questions influence lexical choices in user answers? In: R. Bannert, M. Heldner, K. Sullivan, und P. Wretling, Hrsgg., *FONETIK-97*, No. 4 in Reports from the Department of Phonetics, 5 1997.
- M. Haapalainen und A. Majorin. GERTWOL und Morphologische Disambiguierung für das Deutsche. In: *Proceedings of NODALIDA-95*, 1995.
- M. Haase und K. Heitmann. Die Erweiterte Kölner Phonetik. 526, 2000.
- C. Habel. Cognitive Linguistics: The Processing of Spatial Concepts. Technical Report 45, Institut für Wissensbasierte Systeme, IBM Deutschland GmbH, 1988.
- W. v. Hahn, W. Hoepfner, A. Jameson, und W. Wahlster. The anatomy of the natural language dialogue system HAM-RPM. In: L. Bolc, Hrsg., *Natural Language Based Computer Systems*, Seiten 119–253. Hanser, München, 1980.
- K.-C. Hamborg. Evaluationsbericht OSIRIS (Heuristische Evaluation). Technical report, Universität Osnabrück, Fachbereich Psychologie, 1997.
- K.-C. Hamborg. Evaluationsbericht OSIRIS (IsoMetrics)^L. Technical report, Universität Osnabrück, Fachbereich Psychologie, 1998.
- K.-C. Hamborg, G. Gediga, und C. Moranz. Evaluationsstudie zur softwareergonomischen Bewertung des OPAC und OSIRIS Systems der Universitätsbibliothek Osnabrück. Technical report, Universität Osnabrück, Fachbereich Psychologie, 1998.
- K.-C. Hamborg und S. Greif. Heterarchische Aufgabenanalyse. In: H. Dunkel, Hrsg., *Psychologische Arbeitsanalyse*. VDF-Verlag, 1998.
- S. Hanks und D. McDermott. Default Reasoning, Nonmonotonic Logics, and the Frame Problem. In: *Proceedings of the AAAI-86: 5th international conference on AI*, Seiten 328–333. AAAI, 1986.
- C. Hartmann. *Technische Interaktionskontexte. Aspekte einer sozialwissenschaftlichen Theorie der Mensch-Computer-Interaktion*. Sozialwissenschaft. Deutscher Universitäts-Verlag, Wiesbaden, 1992.
- Alexander G. Hauptmann und Bert F. Green. A Comparison of Command, Menu-Selection and Natural-Language Computer Programs. *Behaviour and Information Technology*, 2(2):163–178, 1983.
- P. J. Hayes. The Frame Problem and Related Problems in Artificial Intelligence. In: *Readings in Artificial Intelligence*, Seiten 223–230. Morgan Kaufmann, San Mateo, California, 1981.

- H. Helbig, C. Gnörlich, und D. Menke. Realization of a User-friendly Access to Networked Information Retrieval Systems. In: Kavi Mahesh, Hrsg., *Natural Language Processing for the World Wide Web. Papers from the 1997 AAAI Spring Symposium*, Stanford, USA, 3 1997.
- H. Helbig und A. Mertens. Word Agent Based Natural Language Processing. In: A. Nijholt und L. Boves, Hrsgg., *Speech and Language Engineering — Proc. of the 8th Twente Workshop on Language Technology. Enschede, Dec 1994*, Enschede, The Netherlands, 1994.
- H. Helbig und M. Schulz. Knowledge Representation with MESNET - A Multilayered Extended Semantic Network. In: *Proceedings of the AAAI Spring Symposium on Ontological Engineering*, Stanford, CA, 1997.
- M. Hemmje. Eine inhaltsorientierte, intuitive 3D-Benutzerschnittstelle für Information-Retrieval-Systeme. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 82–97, Konstanz, 1993. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- G. Hendrix. Natural language interface. *American Journal of Computational Linguistics*, 8(2):56–61, 1982.
- M. Herfurth, P. Mutschke, und H. Ohly. AKCESS: Konzept-orientiertes Retrieval mit bibliographischem Kontextwissen. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 199–207, Konstanz, 1993. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- T. Hermes. *Texturen: Analyse, Beschreibung und Synthese*. PhD thesis, TZI, Universität Bremen, Bremen, 1999.
- T. Hermes und A. Dammeyer. DiVA - intelligente Videoanalyse. *Künstliche Intelligenz. Organ des Fachbereichs 1 Künstliche Intelligenz der Gesellschaft für Informatik e.V.*, 13(1):45–48, 1 1999.
- T. Hermes, A. Miene, und O. Moehrke. Automatic Texture Classification by Visual Properties. In: *Proceedings of the 23rd Annual Conference of the Gesellschaft für Klassifikation e.V.*, Bielefeld, 10.-12. März 1999. Gesellschaft für Klassifikation e.V.
- O. Herzog, A. Miene, T. Hermes, und P. Alshuth. Graphics in/for Digital Libraries — Integrated information mining for texts, images, and videos. *Computers and Graphics*, 22(6):675–685, December 1998.
- O. Herzog und C. R. Rollinger, Hrsgg. *Text Understanding in LILOG*. Springer, Berlin, Heidelberg, New York, 1992.

- C. R. Hildreth. To Boolean or Not To Boolean? *Information technology and libraries*, 2(9):235–237, September 1983.
- C.R. Hildreth. Beyond Boolean: Designing the Next Generation of Online Catalogs. *Library trends*, 35:647–667, 1987.
- I.D. Hill. Natural Language versus Computer Language. In: M.E. Sime und M.J. Coombs, Hrsgg., *Designing for Human-Computer Communication*, Computers and People, Seiten 55–72. Academic Press, New York, NY, 1983.
- G. Hirst. *Anaphora in Natural Language Understanding: A Survey*. No. 119 in Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, New York, 1981.
- G.J. Hitch, A.G. Sutcliffe, J.M. Bowers, und L.M. Eccles. Empirical Evaluation of Map Interfaces: A Preliminary Study. In: M. D. Harrison und A. F. Monk, Hrsgg., *People and Computers: Designing for Usability*, Seiten 565–585, Cambridge, U.K., 1986. Cambridge University Press.
- L. Hitzengerger. Möglichkeiten bei der Anwendung phonologischer Verfahren bei der Namenssuche in Datenbanken. In: P. Hellwig und H. Lehmann, Hrsgg., *Trends in der Linguistischen Datenverarbeitung*, Seiten 71–85. Olms, Hildesheim, 1986.
- J.R. Hobbs. On the Coherence and Structure of Discourse. Technical Report CSLI–85–37, Center for the Study of Language and Information, Stanford University, Stanford, California, 1985.
- J.R. Hobbs. Resolving Pronoun References. In: K. Sparck-Jones, B.J. Grosz, und B.L. Webber, Hrsgg., *Readings in Natural Language Processing*, Seiten 339–352. Morgan Kaufmann, Los Altos, California, 1986.
- T. N. Höhle. Über Komposition und Derivation: zur Konstituentenstruktur von Wortbildungsprodukten im Deutschen. *Zeitschrift für Sprachwissenschaft*, 1(1):76–112, 1982.
- K. Höök. Evaluating the Utility and Usability of an Adaptive Hypermedia System. In: *Proceedings of IUI 97*, Seiten 179–186, Orlando, FL USA, 1997. ACM.
- L. Horn. *On the Semantic Properties of the Logical Operators in English*. Ph.d. dissertation, University of California at Los Angeles, Los Angeles, California, 1972.
- L. Horn. *The Natural History of Negation*. Chicago University Press, 1989.
- P. Horsman. Accessibility of Archival Documents. In: F.M.G. de Jong und A. Nijholt, Hrsgg., *TWLT 5: Natural Language Interfaces. From Laboratory to Commercial and User Environment*, Seiten 37–41, Enschede, 1993. Universiteit Twente.

- W. Hötker. *Dynamische Modularisierung lexikalischer Informationen in einem Wiederverwertungsszenario*. No. 151 in *Dissertationen zur künstlichen Intelligenz*. Infix, Sankt Augustin, 1997.
- C. Huls und E. Bos. Edward: A Multimodal User Interface. In: *Proceedings of the Twente Workshop on Language Technology (TWLT5)*, Seiten 89–98, University of Twente, 1993.
- C. Huls und E. Bos. Studies into Full Integration of Language and Action. In: H. Bunt, Hrsg., *Multimodal human-computer communication: systems, techniques, and experiments*, Vol. 1374: *Lecture Notes in Computer Science*, Seiten 312–325, Berlin, 5 1998. International Conference on Cooperative Multimodal Communication (CMC), 1995, Springer.
- J. Hulstijn, R. Steetskamp, H. ter Doest, S. van de Burgt, und A. Nijholt. Topics in SCHISMA Dialogues. In: S. LuperFoy, A. Nijholt, und G. Veldhuijzen van Zanten, Hrsgg., *Dialogue Management in Natural Language Systems. Proceedings of the 11th Twente Workshop on Language Technology (TWLT 11)*, Seiten 89–99, Enschede, 1996. University of Twente.
- H. Isahara, K. Uchimoto, und H. Ozaku. Intelligent Network News Reader with Visual User Interface. In: *Proceedings of the COLING-ACL'98 Workshop on Content Visualization and Intermedia Representations (CVIR'98)*, 1998.
- A. Jameson. Adapting to the User's Time and Working Memory Limitations: New Directions of Research. In: U. J. Timm, Hrsg., *Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen (ABIS-98)*, Erlangen; München; Passau, 1998.
- A. Jameson, W. Hoepfner, und W. Wahlster. The natural language system HAM-RPM as a hotel manager: Some representational prerequisites. In: R. Wilhelm, Hrsg., *GI-10: Jahrestagung*, Saarbrücken, 1980. Springer-Verlag.
- B. Janosky, P.J. Smith, und C.R. Hildreth. Online Library Catalog Systems: An Analysis of User Errors. *International Journal of Man-Machine Studies*, 25(5):573–592, 1986.
- M. Jarke, J. A. Turner, E. A. Stohr, Y. Vassiliou, N. H. White, und K. Michielsen. A Field Evaluation of Natural Language for Data Retrieval. *IEEE Transactions on software engineering*, SE-11(1):97–114, January 1985.
- S. Jekat, A. Klein, E. Maier, I. Maleck, M. Mast, und J. Quantz. Dialogue Acts in VerbMobil. Technical Report 65, Universität Hamburg, DFKI GmbH, Universität Erlangen, TU Berlin, 4 1995.
- S. Jekat, H. Tappe, H. Gerlach, und T. Schöllhammer. Dialogue Interpreting: Data and Analysis. Technical Report 189, Arbeitsbereich Natürlichsprachliche Systeme, Fachbereich Informatik, Universität Hamburg, 3 1997.

- C. Johnson und J. Bachenko. Applied Computational Linguistics in Perspective: Proceedings of the Workshop. *American Journal of Computational Linguistics*, 8(2): 55–84, 1982.
- S. Jones, M. Gatford, und S. Walker. Transaction Logging. *Journal of Documentation*, 53:35–50, 1997a.
- S. Jones, S. Walker, M. Gatford, und T. Do. Peeling the Onion: OKAPI System Architecture and Software Design Issues. *Journal of Documentation*, 53:58–68, 1997b.
- T. Jörding und K. Meissner. Intelligent multimedia presentations in the Web: fun without annoyance. *Computer Networks and ISDN Systems*, 30(1–7):649–650, April 1998.
- T. Jörding, S. Michel, und M. Popella. TELLIM – Ein System für adaptive multimediale Produktpräsentation im World Wide Web. In: U. J. Timm, Hrsg., *Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen (ABIS-98)*, Seiten 29–40, Erlangen; München; Passau, 1998.
- B. Jung, S. Kopp, M.E. Latoschik, T. Sowa, und I. Wachsmuth. Virtuelles Konstruieren mit Gestik und Sprache. *Künstliche Intelligenz. Organ des Fachbereichs 1 Künstliche Intelligenz der Gesellschaft für Informatik e.V.*, 14(2):5–11, 4 2000.
- U. Junger. Möglichkeiten und Probleme automatischer Erschließungsverfahren in Bibliotheken. Bericht vom KASCADE-Workshop in der Universitäts- und Landesbibliothek Düsseldorf. *Bibliothek*, 23(1), 1999.
- F. Junginger. Die Eignung von Precis für deutsche Bibliotheken. Bemerkungen zu Bernd Maaßen: 'Precis: Erfahrungen mit einem Projekt'. *Zeitschrift für Bibliothekswesen und Bibliographie*, 31(1):31–40, 1984.
- H. Kamp. A theory of truth and semantic representation. In: J.A.G. Groenendijk, T.M.V. Janssen, und M.B.J. Stokhof, Hrsgg., *Formal Methods in the Study of Language*, No. 135 in Mathematical Centre Tracts, Seiten 277–322. Mathematical Centre, Amsterdam, 1981.
- H. Kamp und U. Reyle. *From Discourse to Logic*. Kluwer Academic Publishers, Dordrecht, 1990.
- U. Kampffmeyer. Multilinguale Dokumenten-Retrievalsysteme. Implementierung und Beispiele. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 223–241, Konstanz, 1993. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- S. Kanngießner. Strukturen der Wortbildung. In: C. Schwarze und D. Wunderlich, Hrsgg., *Handbuch der Lexikologie*, Seiten 134–183. Athenäum, Königstein/Ts, 1985.

- H.-U. Kasper, W. Krieger. Integration of Prosodic/Grammatical Information in the Analysis of Dialogs. Technical Report 141, DFKI Saarbrücken, 7 1996.
- A. Kennedy, A. Wilkes, L. Elder, und W. Murray. Dialogue with machines. *Cognition*, 30:73–105, 1988.
- G. Kikui, Y. Hayashi, und S. Suzaki. Cross-lingual Information Retrieval on the WWW. In: *Proceedings of the First Workshop on Multilinguality in Software Engineering: The AI Contribution (MULSAIC)*. European Coordinating Committee for Artificial Intelligence, August 1996.
- R. Kittredge. Sublanguages. *American Journal of Computational Linguistics*, 8(2): 79–84, 1982.
- P. Klahr. Conditional Answers in question-answering systems. In: *Proceedings of the Sixth International Joint Conference on Artificial Intelligence*, Seiten 481–483, Los Altos, CA, 1979. Kaufmann.
- J. Klavans und P. Schauble. NSF-EU Multilingual Information Access. *Communications of the ACM*, 41(4):69, 1998.
- G. Klose, E. Lang, und T. Pirlein. Die Ontologie und Axiomatik der Wissensbasis von LEU/2: Erfahrungen, Probleme, Ausblicke. Technical Report 171, Institut für Wissensbasierte Systeme, IBM Deutschland GmbH, 1991.
- W. Kneale. Modality, de dicto and de re. In: E. Nagel, P. Suppes, und A. Tarski, Hrsgg., *Logic, Methodology and the Philosophy of Science. Proceedings of the 1960 International Congress.*, Seiten 622–633. Stanford University Press, 1962.
- K. Knight und S. Luk. Building a Large-Scale Knowledge Base for Machine Translation. In: *Proceedings of the 12th National Conference on Artificial Intelligence*, Vol. 1, Seiten 773–778, Menlo Park, CA, USA, July 31–August 4 1994. AAAI Press.
- A. Knoll, I. Glöckner, H. Helbig, und S. Hartrumpf. A System for the Content-Based Retrieval of Textual and Non-Textual Documents Using a Natural Language Interface. *Informatik Berichte 232/1998*, FernUniversität Hagen, Fachbereich Informatik, Hagen, 7 1998.
- T. Koda und P. Maes. Agents with Faces: The Effects of Personification of Agents. In: *Proceedings of HCI'96*, London, 1996.
- J. Krause. Mensch-Maschine-Interaktion in natürlicher Sprache. Zur Bewertung eines natürlichsprachigen Frage-Antwort-Systems. In: R. Kuhlen, Hrsgg., *Datenbasen, Datenbanken, Netzwerke: Praxis des Information Retrieval. Bd. 3: Nutzung und Bewertung von Retrievalsystemen*, Seiten 199–229. Saur, München, 1980.

- S. Kripke. Naming and Necessity. In: Gilbert Harman und Donald Davidson, Hrsgg., *Semantics of Natural Language*, Seiten 253–355. D. Reidel Publishing Co., Dordrecht, 1972.
- A. Krüger und M. Ronthaler. Intelligente Systemreaktionen in CALL-Software. Technical Report 25, Scientific Centre, Institute for Logic and Linguistics, IBM Deutschland GmbH, Heidelberg, 4 1997.
- P. Kugel. Dirty Boole? *Journal of the American Society for Information Science*, 22: 293–294, 1971.
- H. Kunisch. Die Anzahl der Treffer ist oft Null. *Informationweek*, (22):87, 10 1999.
- C. Ladewig. Ähnlichkeitsmessung mit und ohne aspektische Indexierung. *nfd Information – Wissenschaft und Praxis*, 49:459–462, 1998a.
- C. Ladewig. 'Information Retrieval ohne Linguistik?' Erwiderung zu dem Artikel von Gerda Ruge und Sebastian Goeser, *NfD* 49(1998)6, S.361-369. *nfd Information – Wissenschaft und Praxis*, 49:476–478, 1998b.
- T. K. Landauer, M. K. Galotti, und S. Hartwell. Natural Command Names and Initial Learning: a Study of Text-Editing Terms. *Communications of the ACM*, 26(7):495–503, July 1983.
- D.T. Langendoen und H. Savin. The projection problem for presuppositions. In: C. Fillmore und D.T. Langendoen, Hrsgg., *Studies in Linguistic Semantics*. Holt, New York, 1971.
- H. Langer. Disambiguierung von Präpositionalkonstruktionen mit einem Syntaktischen Parser. In: S. Mehl, A. Mertens, und M. Schulz, Hrsgg., *Präpositionalsemantik und PP-Anbindung*. Uni Duisburg, 1996.
- J. Lee und K. Stenning. Anaphora in Multimodal Discourse. In: H. Bunt, Hrsg., *Multimodal human-computer communication: systems, techniques, and experiments*, Vol. 1374: *Lecture Notes in Computer Science*, Seiten 250–263, Berlin, 5 1998. International Conference on Cooperative Multimodal Communication (CMC), 1995, Springer.
- W. Lehnert. Question answering in natural language processing. In: Leonard Bolc, Hrsg., *Natural Language Question Answering Systems*, Seiten 9–71. Hanser, München, 1980.
- K. Lepsky. Maschinelles Indexieren zur Verbesserung der sachlichen Suche im OPAC. *Bibliotheksdienst*, 28(8):1234–1242, 1994.
- K. Lepsky. Abschlußbericht zum Projekt MILOS II.
http://www.rz.uni-duesseldorf.de/ulb/mil/_ber.htm, 1996a.

- K. Lepsky. Automatische Indexierung und bibliothekarische Inhaltserschließung: Ergebnisse des DFG-Projekts MILOS I. In: E. Niggemann und K. Lepsky, Hrsgg., *Zukunft der Sacherschließung im OPAC: Vorträge des 2. Düsseldorfer OPAC-Kolloquiums am 21. Juni 1995*, Vol. 25: *Schriften der Universitäts- und Landesbibliothek Düsseldorf*. Univ.- und Landesbibliothek, Düsseldorf, 1996b.
- K. Lepsky. Sacherschließung ohne RSWK? Neue Praxis an der Universitäts- und Landesbibliothek Düsseldorf. *ProLibris*, (2):112–114, 1998.
- K. Lepsky. Automatische Indexierung zur Erschließung deutschsprachiger Dokumente. *nfd Information – Wissenschaft und Praxis*, 50(6), 1999.
- K. Lepsky, J. Siepmann, und A. Zimmermann. Automatische Indexierung für Online-Kataloge: Ergebnisse eines Retrievaltests. *Zeitschrift für Bibliothekswesen und Bibliographie*, 43(1):47–56, 1996.
- K. Lepsky und H. Zimmermann. Katalogerweiterung durch Scanning und Automatische Dokumenterschließung. Das DFG-Projekt KASCADE. *ABI-Technik*, 18(1): 56–60, 1998.
- A. Lesch. Question Answering Systems: A Survey of Major Accomplishments. In: I. S. Batori, W. Lenders, und W. Putschke, Hrsgg., *Computational Linguistics - Computerlinguistik*, Handbücher zur Sprach- und Kommunikationswissenschaft, Seiten 722–739. de Gruyter, Berlin, 1989.
- J. Lester, S. Converse, S. Kahler, B. Barlow, S. Stone, und R. Bhogal. The Persona Effect: Affective Impact of Animated Pedagogical Agents. In: *Proceedings of CHI 97*, Seiten 359–366, Atlanta, GA USA, 3 1997a. ACM.
- J. Lester, P. FitzGerald, und B. Stone. The Pedagogical Design Studio: Exploiting Artifact-Based Task Models for Constructivist Learning. In: *Proceedings of IUI 97*, Seiten 155–162, Orlando, FL, CA USA, 1997b. ACM.
- J. Lester und B. Stone. Increasing Believability in Animated Pedagogical Agents. In: *Proceedings of Autonomous Agents 97*, Seiten 16–21, Marina Del Rey, CA USA, 1997. ACM.
- D. Lewis. Scorekeeping in a language game. *Journal of Philosophical Logic*, 8, 1979.
- H. Lieberman. Letizia: An Agent that Assists Web Browsing. In: C. S. Mellish, Hrsgg., *Proceedings of 14th International Joint Conference on Artificial Intelligence (IJCAI-95)*, Vol. 1, Seiten 924–929, San Mateo, CA, 1995. Morgan Kaufmann.
- H. Lieberman. An Automated Channel-Surfing Interface Agent for the Web. <http://www.media.mit.edu/~lieber/>, 1996.

- J. Light. A Distributed, Graphical, Topic-oriented Document Search System. In: *CIKM '97. Proceedings of the 1997 ACM 6th international conference on Information and knowledge management*, Seiten 285–292. ACM, 1997.
- C. Lisetti. A User Model of Emotion-Cognition. In: *'Attitude, Personality and Emotions in User-Adapted Interaction.'* *Proceedings of the UM'99 Workshop*, 1999.
- D. Litman und J.F. Allen. A Plan Recognition Model for Subdialogues in Conversations. *Cognitive Science*, 11:163–200, 1987.
- N. V. Loukachevitch. Knowledge Representation for Multilingual Text Categorization. In: *AAAI Symposium on Cross-Language Text and Speech Retrieval*. American Association for Artificial Intelligence, March 1997.
- M. Luria. Dividing Up the Question Answering Process. In: *Proceedings of the The National Conference on Artificial Intelligence*, Seiten 71–74, 1982.
- B. Maaßen. Man sollte das alles nicht so eng sehen (Einige Bemerkungen zum 'engen' Schlagwort. *Zeitschrift für Bibliothekswesen und Bibliographie*, 28(5):353–357, 1981.
- B. Maaßen. PRECIS. Erfahrungen mit einem Projekt. *Zeitschrift für Bibliothekswesen und Bibliographie*, 30(4):293–301, 1983.
- S. Maass. Why Systems Transparency? In: T. Green, S. Payne, und G. van der Veer, Hrsgg., *The Psychology of Computer Use*, Computer and People Series, Seiten 19–28. Academic Press, London, New York, 1983.
- K. Mahesh und S. Nirenburg. Meaning Representation for Knowledge Sharing in Practical Machine Translation. In: *Proceedings of the 9th Florida Artificial Intelligence Research Symposium (FLAIRS)*, Seiten 182–186, 1996.
- C. Maienborn. Lokale Verben und Präpositionen: Semantische und konzeptuelle Verarbeitung in LEU II. Technical Report 119, Institut für Wissensbasierte Systeme, IBM Deutschland GmbH, 1990a.
- C. Maienborn. Processing Spatial Knowledge in LILOG. Technical Report 157, Institut für Wissensbasierte Systeme, IBM Deutschland GmbH, 1991.
- E. Maier. Context Construction as Subtask of Dialogue Processing - the VERBMOBIL Case. In: S. LuperFoy, A. Nijholt, und G. Veldhuijzen van Zanten, Hrsgg., *Dialogue Management in Natural Language Systems. Proceedings of the 11th Twente Workshop on Language Technology (TWLT 11)*, Seiten 113–122, Enschede, 1996. University of Twente.
- E. Maier. Evaluating a Scheme for Dialogue Annotation. Technical Report 193, DFKI GmbH Saarbrücken, 4 1997.

- W. C. Mann und S. A. Thompson. Rhetorical Structure Theory: Toward a functional theory of text organization. *TEXT*, (8 (3) (1988)):243–281, 1988.
- G. Mark. The Case against Cross-Situational Consistency in Human-Computer Interaction. In: 'Attitude, Personality and Emotions in User-Adapted Interaction.' *Proceedings of the UM'99 Workshop*, 1999.
- A. Mark Pejtersen. Semantic Information Retrieval. *Communications of the ACM*, 41 (4):90–92, 1998.
- J. Marks und E. Reiter. Avoiding Unwanted Conversational Implicatures in Text and Graphics. In: T. Dietterich und W. Swartout, Hrsgg., *Proceedings of the 8th National Conference on Artificial Intelligence (AAAI-90)*, Vol. 1, Seiten 450–456, Cambridge, MA, 1990. MIT Press.
- J. Marx und M. Schudnagis. WING-M2: eine objektorientierte multimodale Benutzeroberfläche für Werkstoffdatenbanken. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 169–182, Konstanz, 1993. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- Europäische Kommission. Informationen des öffentlichen Sektors - Eine Schlüsselresource für Europa - Grünbuch über die Informationen des öffentlichen Sektors in der Informationsgesellschaft, 1998.
- J. McCarthy. Circumscription: A Form of Non-Monotonic Reasoning. *Artificial Intelligence*, (13/1980):27–39, 1979a.
- J. McCarthy und P. J. Hayes. Some Philosophical Problems from the Standpoint of Artificial Intelligence. In: B. Meltzer und D. Michie, Hrsgg., *Machine Intelligence 4*. Edinburgh University Press, Edinburgh, 1969.
- D. McDonald und S. Searing. Bibliographic Instruction and the Development of Online Catalogs. *College and Research Libraries*, Seiten 5–11, 1983.
- K.R. McKeown. Natural Language for Expert Systems: Comparisons with Database Systems. In: *Proceedings of the 10th International Conference on Computational Linguistics (COLING'84)*, Seiten 190–193, 1984.
- C. McKnight. Many Projects That Depend on Collaboration. *ACM Communications*, 41(4):86–87, 1998.
- R. McNab, L. Smith, I. Witten, C. Henderson, und S. Cunningham. Towards the digital music library: tune retrieval from acoustic input. In: *DL '96. Proceedings of the 1st ACM international conference on digital libraries*, Seiten 11–18. ACM, 1996.

- D. Menke, C. Gnörlich, und H. Helbig. Access to Bibliographic Databases - The Provider Agents in MeDoc. In: A. Barth, M. Breu, A. Endres, und A. de Kemp, Hrsgg., *Digital Libraries in Computer Science: The MeDoc Approach*, No. 1392 in Lecture Notes in Computer Science, Seiten 51–66. Springer, Berlin, Heidelberg, New York, 1998.
- A. Mertens, M. Schulz, und H. Helbig. Analyse mit Wortagenten im NLP-System LINAS. In: L. Hitzenberger, Hrsgg., *Angewandte Computerlinguistik. Vorträge im Rahmen der Jahrestagung 1995 der Gesellschaft für Linguistische Datenverarbeitung (GLDV) e.V.*, Vol. 15: *Sprache und Computer*, Seiten 63–75, Hildesheim, 1995. Gesellschaft für Linguistische Datenverarbeitung (GLDV) e.V., Georg Olms Verlag.
- J.-T. Milde. Mensch-Maschine-Kommunikation in intelligenten virtuellen Umgebungen: Der kommunikative Agent Lokutor. *Künstliche Intelligenz. Organ des Fachbereichs 1 Künstliche Intelligenz der Gesellschaft für Informatik e.V.*, 14(2):17–21, 4 2000.
- C. Miller. Computational Approaches To Interface Design: What Works, What Doesn't, What Should and What Might. In: *Proceedings of IUI 97*, Seiten 123–126, Orlando, FL, CA USA, 1997. ACM.
- L. A. Miller. Natural Language Programming: Styles, Strategies, and Contrasts. *IBM Systems Journal*, 20(2):184–215, 1981.
- D. Moldt und C. von Scheve. Soziologisch adäquate Modellierung emotionaler Agenten. In: M. Müller, Hrsgg., *Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen. Proceedings der ABIS 2000*. Gesellschaft für Informatik (GI), Universität Osnabrück, 9 2000.
- G. Möller, K. Carstensen, B. Diekmann, und H. Wätjen. Automatic Classification of the World-Wide Web using the Universal Decimal Classification. In: B. McKenna und C. Graham, Hrsgg., *Proceedings of the 23rd International Online Information Meeting (Online99)*, Seiten 231–237, 1999.
- J.-U. Möller. DIA-MOLE: An Unsupervised Learning Approach to Adaptive Dialogue Models for Spoken Dialogue Systems. In: *Proceedings of Eurospeech '97, 5th European Conference on Speech Communication and Technology*, Seiten 2271–2274, Grenoble, September 1997. European Speech Communication Association.
- C. Montgomery. Concept Extraction. *American Journal of Computational Linguistics*, 8(2):70–73, 1982.
- Y. Moon und C. Nass. Adaptive Agents and Personality Change: Complementarity versus similarity as forms of adaption. In: *Proceedings of CHI 96*, Seiten 287–288, Vancouver, BC Canada, 4 1996. ACM.

- J. D. Moore. Responding to 'Huh?': Answering Vaguely Articulated Follow-Up Questions. In: K. Bice und C. Lewis, Hrsgg., *Wings for the mind: Proceedings of ACM CHI'89 Conference on Human Factors in Computing Systems*, Interacting with Computer Advisors, Seiten 91–96, New York, 1989.
- C. Moranz. OPAC und OSIRIS. Vergleich einer strukturiert- und einer natürlichsprachlichen Eingabeschnittstelle für die Literaturrecherche. Technical report, Universität Osnabrück, Fachbereich Psychologie, 1998.
- C. Moranz. OPAC oder OSIRIS? Evaluationsstudie zum Vergleich der Benutzbarkeit eines strukturiert- und eines natürlichsprachlichen Bibliothekssystems. Diplomarbeit, Universität Osnabrück, Fachbereich Psychologie und Gesundheitswissenschaften, Fachgebiet Arbeits- und Organisationspsychologie, Osnabrück, März 2000.
- C. Moranz, K.-C. Hamborg, und M. Ronthaler. On the influence of natural language vs. structured command language systems on query strategies. In: H.-J. Bullinger und P.H. Vossen, Hrsgg., *Human-Computer Interaction: Proceedings of HCI International '99 (the 8th International Conference on Human-Computer Interaction)*, Seiten 199–200, Stuttgart, August 1999. Fraunhofer IRB Verlag.
- K. Morik. Demand and Requirements for Natural Language Systems-Results of an Inquiry. In: Alan Bundy, Hrsg., *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Seiten 647–649, Karlsruhe, August 1983. William Kaufmann.
- P. Moser. Die Recherche wird leichter. *Börsenblatt des Deutschen Buchhandels*, 80 (10):9–10, 10 1998.
- E.-M. Mueckstein. Q-TRANS: Query Translation Into English. In: Alan Bundy, Hrsg., *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Seiten 660–662, Karlsruhe, August 1983. William Kaufmann.
- N. Müller, S. Osswald, A. Pieper, und T. Kirsten. Auf SIERA ist immer Verlaß? *Bibliotheksdienst*, 32(11):1965–1969, 1998.
- K. Nagao und A. Takeuchi. Social Interaction: Multimodal Conversation with Social Agents. In: *Proceedings of the 12th National Conference on Artificial Intelligence. Volume 1*, Seiten 22–28, Menlo Park, CA, USA, 1994. AAAI Press.
- H.A. Napier und N.S. Guadango. The Impact of a Restricted Natural Language Interface on Ease of Learning and Productivity. *Communications of the ACM, CACM*, 32(10):1190–1198, October 1989.
- C. Nass, B. Fogg, und Y. Moon. Can computers be teammates? *International Journal of Human-Computer Studies*, 45:669–678, 1996.

- C. Nass, Y. Moon, and P. Carney. Are People Polite to Computers? Responses to Computer-Based Interviewing Systems. *Journal of applied social psychology: devoted to applications of experimental behavioral science research to problems of society*, 29(5):1093–1110, 1999.
- C. Nass, Y. Moon, B.J. Fogg, B. Reeves, and C. Dryer. Can Computer Personalities be Human Personalities? In: *Proceedings of CHI 95*, Seiten 228–229, Denver, Colorado, 5 1995. ACM.
- C. Nass, Y. Moon, and N. Green. Are Machines Gender Neutral? Gender Stereotypic Responses to Computers With Voices. *Journal of applied social psychology: devoted to applications of experimental behavioral science research to problems of society*, 27(10):864–876, 1997.
- C. Nass, J. Steuer, E. Tauber, and H. Reeder. Anthropomorphism, Agency, & Ethopoeia: Computers as Social Actors. In: *Proceedings of ACM INTERCHI'93 Conference on Human Factors in Computing Systems – Adjunct Proceedings*, Short Papers (Posters): Multimedia and Multiuser Interfaces, Seiten 111–112, 1993.
- C. Nass, J. Steuer, and E.R. Tauber. Computers are Social Actors. In: Beth Adelson, Susan Dumais, and Judith Olson, Hrsgg., *Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems*, Vol. 1: *Expressive Interfaces*, Seiten 72–78, New York, NY, USA, April 1994. ACM Press.
- G. Navarro und R. Baeza-Yates. A Language for Queries on Structure and Contents of Textual Databases. In: *Proceedings of the Eighteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Integrating Structured and Unstructured Information, Seiten 93–101, 1995.
- G. Navarro und R. Baeza-Yates. Proximal Nodes: A Model to Query Document Databases by Content and Structure. *ACM Transactions on Information Systems*, 15(4): 400–435, 1997.
- D. Novick und S. Sutton. Building on experience: managing spoken interaction through library subdialogues. In: S. LuperFoy, A. Nijholt, und G. Veldhuijzen van Zanten, Hrsgg., *Dialogue Management in Natural Language Systems. Proceedings of the 11th Twente Workshop on Language Technology (TWLT 11)*, Seiten 51–60, Enschede, 1996. University of Twente.
- F. Obermeier. Schlagwortsuche in einem lokalen OPAC am Beispiel der Universitätsbibliothek Eichstätt. *Bibliotheksforum Bayern (BFB)*, 27(3):296–319, 1999.
- W.C. Ogden und P. Bernick. Using Natural Language Interfaces. In: M. Helander, T. Landauer, und P. Prabhu, Hrsgg., *Handbook of Human-Computer Interaction*, Kapitel 7, Seiten 137–161. Elsevier, 2nd, revised edition, 1997.

- W.C. Ogden und S.R. Brooks. Query Languages for the Casual User: Exploring the ground between Formal and Natural Languages. In: *Proceedings of the Annual Meeting of the Computer Human Interaction of the ACM*, Seiten 161–65, 1983.
- W.C. Ogden und C. Kaplan. The use of AND and OR in a natural language computer interface. In: *Proceedings of the Human Factors Society – 30th annual meeting 1986*, Seiten 829–833, Santa Monica, CA, USA, 1986. Human Factors Society.
- W.C. Ogden und A. Sorknes. What do users say to their natural language interfaces? In: *Proceedings of IFIP INTERACT'87: Human-Computer Interaction*, 3. Human-Computer Interface Design: 3.3 Natural Language Dialogues, Seiten 561–566, North Holland, 1987. Elsevier Science Publishers B.V.
- D. Owen. Answers First, Then Questions. In: D. A. Norman und S. W. Draper, Hrsgg., *User Centered System Design: New Perspectives on Human-Computer Interaction*, Seiten 361–375. Erlbaum, Hillsdale, NJ, 1986.
- E. Paek und H. Jeon. Automating a Classification Task Based on an Augmented Thesaurus. In: *Proceedings of IUI 97*, Seiten 221–224, Orlando, FL, CA USA, 1997. ACM.
- C. Paice. An Evaluation Method for Stemming Algorithms. In: *Proceedings of the 17th Annual Int. ACM SIGIR Conference on Research and Development in Information Retrieval, Indexing*, Seiten 42–50, 1994.
- M. Pandit und S. Kalbag. The Selection Recognition Agent: Instant Access to Relevant Information and Operations. In: *Proceedings of IUI 97*, Seiten 47–52, Orlando, FL, CA USA, 1997. ACM.
- M.J. Pazzani und C. Engelman. Knowledge based question answering. In: *ACL Proceedings, Conference on Applied Natural Language Processing*, Seiten 73–80, Santa Monica, CA USA, 1983.
- C. R. Perrault. An Application of Default Logic to Speech Act Theory. In: P. Cohen, J. Morgan, und M. Pollack, Hrsgg., *Intentions in Communication*, Seiten 161–185. MIT Press, Cambridge, MA, 1990.
- C. R. Perrault und B. J. Grosz. Natural-language interfaces. In: H.E. Shrobe und AAI, Hrsgg., *Exploring Artificial Intelligence: Survey talks from the national conferences on Artificial Intelligence*, Kapitel 4, Seiten 133–172. Morgan Kaufman, San Mateo, CA, 1988.
- C. Peters und E. Picchi. Using Linguistic Tools and Resources in Cross-Language Retrieval. In: *AAAI Symposium on Cross-Language Text and Speech Retrieval*. American Association for Artificial Intelligence, March 1997.

- Th. Peters. When smart people fail: an analysis of the transaction log of an online public access catalog. *The journal of academic librarianship: articles, features, and book reviews for the academic library professionals*, 15(5):267–273, 1989.
- L. Pietzka. Techniken und Methoden wissensbasierter Systeme: Expertensysteme in Bibliothek, Information und Dokumentation. *Bibliothek*, 19(3):371–385, 1995.
- C. Piller. INNOVATION – Web Travelers Follow Beaten Paths to Similar Sites. *Los Angeles Times*, 23.8. 1999.
- P. Piwek. Information Flow and Gaps. Technical report, IPO - Center for Research on User-System Interaction, Eindhoven, Eindhoven, The Netherlands, 1999.
- A.S. Pollitt, G.P. Ellis, M.P. Smith, M.R. Gregory, C.S. Li, und H. Zangenberg. A Common Query Interface for Multilingual Document Retrieval from Databases of the European Community Institutions. In: *Proceedings of the Sixteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Demonstrations*, Seite 359, 1993.
- H.-J. Postel. Die Kölner Phonetik. *IBM-Nachrichten*, 19:928–931, 1969.
- R. Rada und E. Bicknell. Ranking Documents with a Thesaurus. *Journal of the American Society for Information Science*, 40(5):304–310, 1989.
- M. Rauterberg. MAUS versus Funktionstaste. Ein empirischer Vergleich einer Desktop- mit einer ASCII-orientierten Benutzungsoberfläche. In: S. Maaß und H. Oberquelle, Hrsgg., *Software-Ergonomie '89: Aufgabenorientierte Systemgestaltung und Funktionalität*, Vol. 32: *Berichte des German Chapter of the ACM*, Seiten 313–323, Stuttgart, 3 1989. German Chapter of the ACM, Gesellschaft für Informatik (GI), Univ. Hamburg, Teubner.
- M. Rauterberg. An Empirical Comparison of Menu-Selection (CUI) and Desktop (GUI) Computer Programs Carried Out by Beginners and Experts. *Behaviour and Information Technology*, 11(4):227–236, 1992.
- M. Rauterberg. *Ein Konzept zur Quantifizierung software-ergonomischer Richtlinien*. PhD thesis, Universität Zürich, Zürich, 1995.
- I. Recker, M. Ronthaler, und H. Zillmann. OSIRIS (Osnabrück Intelligent Research Information System) - ein Hyperbase Front End System für OPACs. *Bibliotheksdienst*, 30(5):833–848, 5 1996.
- N. Reithinger, R. Engel, M. Kipp, und M. Klesen. Predicting Dialogue Acts for a Speech-To-Speech Translation System. Technical Report 151, DFKI GmbH Saarbrücken, 8 1996.

- E. Riloff und J. Shepherd. A Corpus-Based Approach for Building Semantic Lexicons. In: *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, 1997.
- T. Rist, E. André, und J. Müller. Adding Animated Presentation Agents to the Interface. In: *Proceedings of IUI 97*, Seiten 79–86, Orlando, Florida, 1997. ACM.
- S.E. Robertson. Overview of the OKAPI Projects. *Journal of Documentation*, 53:3–7, 1997.
- S.E. Robertson und M. Beaulieu. Research and Evaluation in Information Retrieval. *Journal of Documentation*, 53:51–57, 1997.
- S.E. Robertson, S. Walker, und M. Beaulieu. Laboratory Experiments with OKAPI: Participation in the TREC Programme. *Journal of Documentation*, 53:20–34, 1997.
- S.P. Robertson, J.D. Ullman, und A. Mehta. Simultaneous Question Comprehension and Answer Retrieval. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 474–479, Konstanz, 1993. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- M. Ronthaler. Präsuppositionen, Anaphern und Ellipsen. Master's thesis, Universität Osnabrück, 1994.
- M. Ronthaler und H. Zillmann. Literaturrecherche mit OSIRIS: ein Test der OSIRIS-Retrievalkomponente. *Bibliotheksdienst*, 32(7):1203–1212, Juli 1998.
- S. Roppel. Intelligentes Faktenretrieval am Beispiel der Werkstoffinformation. In: G. Knorz, J. Krause, und C. Womser-Hacker, Hrsgg., *Information Retrieval '93 Von der Modellierung zur Anwendung*, No. 12 in Schriften zur Informationswissenschaft, Seiten 154–168, Konstanz, 1993. Gesellschaft für Informatik, UVK - Universitätsverlag Konstanz.
- E. Rosch. Principles of Categorization. In: E. Rosch und B.B. Lloyd, Hrsgg., *Cognition and Categorization*, Seiten 27–48. Lawrence Erlbaum Associates, Hillsdale (NJ), USA, 1978.
- D. Rosenberg. Language in Multi-media. In: *Proceedings of the CHI 2000 Workshop on Natural Language Interfaces*, 4 2000.
- J. Roukens. The Multilingual Information Society. In: *Forum Sprache ohne Grenzen*, Seite 165. DLR, Projektträger Informationstechnik, 1997.
- D. Rousseau, B. Moulin, und G. Lapalme. Interpreting communicative acts and building a conversation model. *Natural Language Engineering*, 2(3):253–276, 9 1996.

- G. Ruge und S. Goeser. Information Retrieval ohne Linguistik? *nfd Information – Wissenschaft und Praxis*, 49(6):361–369, 1998.
- C. Runciman und N. Hammond. User Programs: A Way to Match Computer Systems and Human Cognition. In: M. D. Harrison und A. F. Monk, Hrsgg., *People and Computers: Designing for Usability*, Seiten 464–481, Cambridge, U.K., 1986. Cambridge University Press.
- E. Sacerdoti. Language Access to Distributed Data with Error Recovery. In: Raj Reddy, Hrsg., *Proceedings of the 5th International Joint Conference on Artificial Intelligence (IJCAI-77)*, Seiten 196–202, Cambridge, MA, August 1977. William Kaufmann.
- E. Sachse, M. Liebig, und W. Gödert. *Automatische Indexierung unter Einbeziehung semantischer Relationen: Ergebnisse des Retrievaltests zum MILOS II-Projekt*, Vol. 14: *Kölner Arbeitspapiere zur Bibliotheks- und Informationswissenschaft*. Fachhochschule Köln, Köln, 12 1998.
- G. Salton, E.A. Fox, und E.M. Voorhees. A Comparison of Two Methods for Boolean Query Relevance Feedback. *Information Processing and Management*, 20(5/6): 637–651, 1984.
- G. Salton, E.A. Fox, und E.M. Voorhees. Advanced Feedback Methods in Information Retrieval. *Journal of the American Society for Information Science*, 36(3):200–210, 1985.
- T. Saracevic. Evaluation of Evaluation in Information Retrieval. In: *Proceedings of the Eighteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Keynote Address, Seiten 138–146, 1995.
- A. Scheffczyk. Das Frameproblem in Philosophie und Künstlicher Intelligenz. Unveröffentlichter Abschlußbericht des DFG-Projektes 'Strukturen des Lexikons' an der Universität Osnabrück, 1990.
- M. Schiehlen. Disambiguation of Underspecified Discourse Representation Structures under Anaphoric Constraints. Technical Report 188, Universität Stuttgart, 1 1997.
- E. Schmidt, G. Weigang, und U. Schulz. OPAC-Nutzer ernstgenommen – Eine Thinking-Aloud-Studie am ALS-OPAC der Hamburger Öffentlichen Bücherhallen. *Bibliothek*, 21(21):215–226, 1997.
- B. Schmitz und J. Quantz. Dialogue Acts in Automatic Dialogue Interpreting. Technical Report 173, Technische Universität Berlin, 9 1996.
- D. Schoder und M. Müller. Potentiale und Hürden des Electronic Commerce. Eine Momentaufnahme. *Informatik-Spektrum*, 22(4):252–260, 8 1999.

- M. Schröder. Knowledge-Based Processing of Medical Language: A Language Engineering Approach. In: Hans Jürgen Ohlbach, Hrsg., *Proceedings of the 16th German Conference on Artificial Intelligence (GWAI-92): Advances in Artificial Intelligence*, Seiten 221–234, Berlin, August 31–September 3 1992. Springer.
- U. Schulz. Was wir über OPAC-Nutzer wissen: Fehlertolerante Suchprozesse in OPACs. *ABI-Technik*, 14(4):299–310, 1994.
- K. Schweizer, M. Paechter, und B. Weidenmann. Sozial wahrnehmbare Merkmale von Agenten in virtuellen Lernumgebungen aus Rezipientensicht. *Künstliche Intelligenz. Organ des Fachbereichs 1 Künstliche Intelligenz der Gesellschaft für Informatik e.V.*, 14(2):22–27, 4 2000.
- J. Searle. Indirect Speech Acts. In: S. Davis, Hrsg., *Pragmatics: A Reader*, Seiten 265–277. Oxford University Press, Oxford, 1975a.
- J. Searle. What is a Speech Act? In: S. Davis, Hrsg., *Pragmatics: A Reader*, Seiten 254–263. Oxford University Press, Oxford, 1975b. reprint aus: Black, M. (ed.): *Philosophy in America*, Allen and Unwin, New York: 1965.
- V. Sethi. Towards Understanding the Role, Importance, and Impact of Natural Language Interfaces to Databases. In: *Proceedings of the 8th International Conference on Information Systems*, Seiten 176–187, Pittsburgh, PA, 1987.
- P. Seuren. *Discourse Semantics*. Basil Blackwell, Oxford, 1985.
- P. Sgall. Natural Language Understanding and the Perspectives of Question Answering. In: J. Horecký, Hrsg., *COLING 82*, Seiten 357–364, Amsterdam, 1982. Coling, North-Holland.
- P. Sheridan, M. Wechsler, und P. Schäuble. Cross-Language Speech Retrieval: Establishing a Baseline Performance. In: N. J. Belkin, A. D. Narasimhalu, und P. Willett, Hrsgg., *Proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR-97)*, Vol. 31, special issue: *SIGIR Forum*, Seiten 99–108, New York, July 1997. ACM Press.
- S. M. Shieber. *An Introduction to Unification-Based Approaches to Grammar*. No. 4 in CSLI Lecture Notes. CSLI, Stanford CA, 1986.
- B. Shneiderman. Improving the Human Factors Aspect of Database Interactions. *ACM Transactions on Database Systems*, 3(4):417–439, December 1978.
- B. Shneiderman. *Software Psychology. Human Factors in Computer and Information Systems*. Winthrop Publishers, Cambridge, 1980.

- B. Shneiderman. Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces. In: *Proceedings of the 1997 International Conference on Intelligent User Interfaces*, Debate: Direct Manipulation vs. Interface Agents, Seiten 33–39, 1997.
- B. Shneiderman, D. Byrd, und W. Croft. Sorting out Searching: A User-Interface Framework for Text Searches. *ACM Communications*, 41(4):95–98, 1998.
- B. Shneiderman und P. Maes. Direct Manipulation vs Interface Agents. *interactions*, 4(6):42–61, 1997.
- Y. Shoham und D. McDermott. Problems in Formal Temporal Reasoning. *Artificial Intelligence*, (36/1988):49–61, 1988.
- T. Shopen. Ellipsis as Grammatical Indeterminacy. *Foundations of Language*, 10: 65–77, 1973.
- W. Sijtsma und O. Zweekhorst. Comparison and Review of Commercial Natural Language Interfaces. In: F.M.G. de Jong und A. Nijholt, Hrsgg., *TWLT 5: Natural Language Interfaces. From Laboratory to Commercial and User Environment*, Seiten 43–57, Enschede, 1993. Universiteit Twente.
- M.E. Sime und M.J. Coombs. Introduction. In: M.E. Sime und M.J. Coombs, Hrsgg., *Designing for Human-Computer Communication*, Computers and People, Seiten 1–20. Academic Press, New York, NY, 1983.
- B.M. Slator, M.P. Anderson, und W. Conley. Pygmalion at the interface. *Communications of the ACM*, 29(7):599–604, July 1986.
- D.W. Small und L.J. Weldon. An Experimental Comparison of Natural and Structured Query Languages. *Human Factors*, 25(3):253–263, 1983.
- S. Soderland, D. Fisher, J. Aseltine, und W. Lehnert. CRYSTAL: Inducing a Conceptual Dictionary. In: C. S. Mellish, Hrsg., *Proceedings of 14th International Joint Conference on Artificial Intelligence (IJCAI-95)*, Vol. 2, Seiten 1314–1319, San Mateo, CA, 1995. Morgan Kaufmann.
- D. Speelman. A natural language interface that uses generalized quantifiers. In: *Proceedings of the Twente Workshop on Language Technology(TWLT5)*, Seiten 69–74, University of Twente, 1993.
- M. Stairmand. Textual Content Analysis for Information Retrieval. In: *Proceedings of Sigir'97*, Seiten 140–147, Philadelphia, PA USA, 1997. ACM.
- E. Stegentritt. EMIR - Multilinguales Freitextretrieval Projekt. In: A. Zimmermann, H. Luckhardt, und H. Schulz, Hrsgg., *Mensch und Maschine - Informationelle Schnittstellen der Kommunikation: Proceedings des 3. Internationalen Symposiums*

- für Informationswissenschaft*, Seiten 179–186, Konstanz, November 1992. Universitätsverlag Konstanz.
- E. Stegentritt. CANAL/LS: ein Versuch, den Zugriff auf Bibliotheksdaten zu verbessern. *ABI-Technik*, 16(4):379–382, 1996.
- E. Stegentritt. Evaluationsresultate des mehrsprachigen Suchsystems CANAL/LS. *ABI-Technik*, 18(1):38–46, 1998.
- E. Stegentritt. Virtual Library. *ABI-Technik*, 19(2):139–140, 1999.
- A. Stein. Active Help and User Guidance in a Multimodal Information System: A Usability Study. In: U. J. Timm, Hrsg., *Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen (ABIS-98)*, Erlangen; München; Passau, 1998.
- A. Stein, J. Atle Gulla, A. Müller, und U. Thiel. Conversational Interaction for Semantic Access to Multimedia Information. In: M. Maybury, Hrsg., *Intelligent multimedia information retrieval*, Kapitel 20, Seiten 399–421. MIT Press, Cambridge MA, USA, 1997.
- O. Stock und C. Strapparava. NL-based Interaction in a Multimodal Environment. In: *Proceedings of the Twente Workshop on Language Technology (TWLT5)*, Seiten 113–118, University of Twente, 1993.
- O. Stock, C. Strapparava, und M. Zancanaro. Explorations in an Environment for Natural Language Multimodal Information Access. In: M. Maybury, Hrsg., *Intelligent multimedia information retrieval*, Kapitel 19, Seiten 381–398. MIT Press, Cambridge MA, USA, 1997.
- G. Stumpf. Quantitative und qualitative Aspekte der verbalen Sacherschließung. *Bibliotheksdienst*, 30(7):1210–1227, 1996.
- M. Svensson, P. Persson, und K. Höök. Using Narratives, Humor, and Social Navigation: An Inspection of Two Systems. In: *'Attitude, Personality and Emotions in User-Adapted Interaction.'* *Proceedings of the UM'99 Workshop*, 1999.
- S. Talja, R. Heinisuo, E. Kasesniemi, H. Kemppainen, S. Luukainen, K. Pispä, und K. Järvelin. Discourse Analysis of User Requests. *ACM Communications*, 41(4): 93–94, 1998.
- A.G. Taylor. Authority Files in Online Catalogs: An Investigation of Their Value. *Cataloging and Classification Quarterly*, 4(3):1–17, 1984.
- W. Teiken. Entwicklungsumgebungen für NL-Parser. Master's thesis, Universität Osnabrück, Osnabrück, 1998.

- H.R. Tennant. Experience with the Evaluation of Natural Language Question Answerers. In: *Proceedings of the Sixth International Joint Conference on Artificial Intelligence*, Vol. 2, Seiten 874–876, Los Altos, CA, 8 1979. Kaufmann.
- H.R. Tennant, K.M. Ross, R.M. Saenz, C.W. Thompson, und J.R. Miller. Menu-based natural language understanding. In: *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, Seiten 151–158, Cambridge, MA, 6 1983a. ACL.
- H.R. Tennant, K.M. Ross, und C.W. Thompson. Usable Natural Language Interfaces Through Menu-Based Natural Language Understanding. In: *Proceedings of ACM CHI'83 Conference on Human Factors in Computing Systems*, Menu and Query Language Design, Seiten 154–160, 1983b.
- A.L. Thomas. Ellipsis: The Interplay of Sentence Structure and Context. *Lingua*, 47: 43–68, 1979.
- J.C. Thomas. Psychological Issues in the Design of Database Query Languages. In: M.E. Sime und M.J. Coombs, Hrsgg., *Designing for Human-Computer Communication*, Computers and People. Academic Press, New York, NY, 1983.
- P. Thompson und C.C. Dozier. Name Searching and Information Retrieval. In: Claire Cardie und Ralph Weischedel, Hrsgg., *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, Seiten 134–140. Association for Computational Linguistics, Somerset, New Jersey, 1997.
- R.M. Tong und D.H. Holtzman. Knowledge-Based Access to Heterogeneous Information Sources. In: *Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries*, 1994.
- S. Towns, C. Callaway, J. Voerman, und J. Lester. Coherent Gestures, Locomotion, and Speech in Life-Like Pedagogical Agents. In: *Proceedings of IUI 98*, Seiten 13–20, San Francisco, CA USA, 1998. ACM.
- A. Treloar. Libraries' New Role in Electronic Scholarly Publishing. *ACM Communications*, 41(4):88–89, 1998.
- J.A. Turner, M. Jarke, E.A. Stohr, Y. Vassiliou, und N. White. Using Restricted Natural Language for Data Retrieval: A Plan for Field Evaluation. Technical Report 38, New York University, Center for Research on Information Systems, 1984.
- H. Turtle. Natural Language vs. Boolean Query Evaluation: A Comparison of Retrieval Performance. In: *Proceedings of the Seventeenth Annual ACM SIGIR Conference on Research and Development in Information Retrieval*, Seiten 212–220, London, 1994. ACM, Springer.

- A. Tversky und D. Kahneman. Judgement under uncertainty: heuristics and biases. In: D. Kahneman, P. Slovic, und A. Tversky, Hrsgg., *Judgement under uncertainty: heuristics and biases*, Kapitel 1, Seiten 3–20. Cambridge University Press, Cambridge, 1982.
- M. Twidale und D. Nichols. Designing Interfaces to Support Collaboration in Information Retrieval. Technical Report CSEG/3/97, University of Lancaster, UK, Cooperative Systems Engineering Group, 1997.
- M. Twidale und D. Nichols. A Survey of Applications of CSCW for Digital Libraries. Technical Report CSEG/4/98, University of Lancaster, UK, Cooperative Systems Engineering Group, 1998.
- R. Van der Sandt. *Context and Presupposition*. Croom Helm, New York, 1988.
- R. Van der Sandt. Denial. In: R.M.Rodriguez L.M.Dobrin, L.Nichols, Hrsg., *CLS 27 - Papers from 27th Regional Meeting of the Chicago Linguistic Society Part Two: The Parasession On Negation*, Seiten 331–344, Chicago, 1991. Chicago Linguistic Society, Chicago Linguistic Society.
- R. Van der Sandt. Presupposition Projection as Anaphora Resolution. *Journal of Semantics*, 9(4), 1992.
- R. A. Van der Sandt. Presupposition and discourse structure. Seiten 267–294. 1989.
- R. A. Van der Sandt und B. Geurts. Presupposition, Anaphora, and Lexical Content. Technical Report 185, Institut für Wissensbasierte Systeme, IBM Deutschland GmbH, 1991.
- S. van Mulken, E. André, und J. Müller. The Persona Effect: How Substantial Is It? In: *Proceedings of the HCI'98 Conference on People and Computers XIII, Usability Testing: Methods and Empirical Studies*, Seiten 53–66, 1998.
- J. Véronis und N. Ide. An assessment of semantic information automatically extracted from machine readable dictionaries. In: *Fifth Conference of the European Chapter of the Association for Computational Linguistics*, Seiten 227–232, Morristown, NJ, 1991. Association for Computational Linguistics.
- I. Wachsmuth. On Structuring Domain-Specific Knowledge. Technical Report 12, Institut für Wissensbasierte Systeme, IBM Deutschland GmbH, 1987.
- I. Wachsmuth. Zur intelligenten Organisation von Wissensbeständen in künstlichen Systemen. Technical Report 91, Institut für Wissensbasierte Systeme, IBM Deutschland GmbH, 1989.

- I. Wachsmuth und B. Gängler. Knowledge Packets and Knowledge Packet Structure. In: O. Herzog und C. R. Rollinger, Hrsgg., *Text Understanding in LILOG*, Seiten 380–393. Springer, Berlin, Heidelberg, New York, 1992.
- A. Wærn. Local Plan Recognition in Direct Manipulation Interfaces. In: *Proceedings of IUI 97*, Seiten 7–14, Orlando, Florida, 1997. ACM.
- W. Wahlster. VERBMOBIL–Translation of Face-to-Face Dialogs. In: O. Herzog, T. Christaller, und D. Schütt, Hrsgg., *Grundlagen und Anwendungen der Künstlichen Intelligenz. 17. Fachtagung für Künstliche Intelligenz*, Seiten 393–402, Berlin, Heidelberg, New York, 1993. Springer.
- W. Wahlster, E. André, W. Finkler, H.-J. Profitlich, und T. Rist. Plan-based integration of natural language and graphics generation. *Artificial Intelligence*, Seiten 387–427, 1993.
- W. Wahlster, E. André, W. Graf, und T. Rist. Designing Illustrated Texts: how Language Production is Influenced by Graphics Generation. Technical Report DFKI-RR-91-05, DFKI — German Research Center for Artificial Intelligence, Saarbrücken, 1991.
- W. Wahlster, H. Marburger, A. Jameson, und S. Busemann. Over-Answering Yes-No Questions: Extended Responses in a NL Interface to a Vision System. In: A. Bundy, Hrsg., *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Seiten 643–646, Karlsruhe, August 1983. William Kaufmann.
- J. Walker, L. Sproull, und R. Subramani. Using a Human Face in an Interface. In: Beth Adelson, Susan Dumais, und Judith Olson, Hrsgg., *Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems*, Vol. 1: *Expressive Interfaces*, Seiten 85–91, New York, NY, USA, April 1994. ACM Press.
- M. Walker, J. Cahn, und S. Whittaker. Linguistic Style Improvisation for Lifelike Computer Characters. In: H. Kitano, Hrsg., *Entertainment and AI, A-Life: papers from the 1996 AAAI Workshop*, Menlo Park, C, 1996. American Association for Artificial Intelligence, AAAI.
- M. Walker und S. Whittaker. When Natural Language is Better than Menus: A Field Study. Technical Report HPL-BRC-TR-89-020, HP Laboratories, Bristol, England, 1989.
- M. Walker und S. Whittaker. Mixed Initiative in Dialogue: An Investigation into Discourse Segmentation. In: *Proc. 28th Annual Meeting of the ACL*, Seiten 70–79, 1990.

- M.A. Walker, J.E. Cahn, und S.J. Whittaker. Improvising Linguistic Style: Social and Affective Bases of Agent Personality. In: W. Lewis Johnson und Barbara Hayes-Roth, Hrsgg., *Proceedings of the First International Conference on Autonomous Agents (Agents'97)*, Seiten 96–105, Marina del Rey, CA, USA, 1997. ACM Press.
- M. Wallace. *Communicating with Databases in Natural Language*. Ellis Horwood/Wiley, Chichester/New York, 1984.
- D. L. Waltz. The state of the art in natural-language understanding. In: W. G. Lehnert und M. H. Ringle, Hrsgg., *Strategies for Natural Language Processing*, Seiten 3–32. Erlbaum, Hillsdale, 1982.
- H. Wätjen. Gerhard - Automatisches Sammeln, Klassifizieren und Indexieren von wissenschaftlich relevanten Informationsressourcen im deutschen World Wide Web. *B.I.T. online: Zeitschrift für Bibliothek, Information und Technologie*, 4:279–290, 1998.
- P. Weinstein. Ontology-based metadata: transforming the MARC legacy. In: *DL '98. Proceedings of the third ACM Conference on Digital libraries*, Seiten 254–263. ACM, 1998.
- J. Weizenbaum. Contextual understanding by computers. *Communications of the ACM, CACM*, 10(8):474–480, 1967.
- S. Wermter und M. Löchel. Learning Dialog Act Processing. Technical Report 139, Universität Hamburg, 7 1996.
- J. Whiteside, S. Jones, P.S. Levy, und D. Wixon. User Performance with Command, Menu, and Iconic Interfaces. In: *Proceedings of ACM CHI'85 Conference on Human Factors in Computing Systems*, Interface Tools and Structures, Seiten 185–191, 1985.
- S. Whittaker und P. Stenton. User Studies and the Design of Natural Language Systems. In: *Fourth Conference of the European Chapter of the Association for Computational Linguistics (EACL'89)*, Seiten 116–123. European Chapter of the Association for Computational Linguistics (EACL), Association for Computational Linguistics, 1989.
- O. Wihofszki. Grenzen des Goldrauschs. *Informationweek*, Themenheft E-Commerce: 16–17, 8 1999.
- S. Willie und P. Bruza. Users' Model of the Information Space: The Case for Two Search Models. In: *Proceedings of the Eighteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Search Interfaces, Seiten 205–210, 1995.

- H. Willumeit, G. Gediga, und K.-C. Hamborg. Validation of the IsoMetrics usability inventory. Technical Report 105, Universität Osnabrück – Fachbereich Psychologie, Osnabrück, 9 1995.
- H. Willumeit, G. Gediga, und K.-C. Hamborg. IsoMetrics: Ein Verfahren zur formativen Evaluation von Software nach ISO 9241/10. *Ergonomie und Informatik*, 27: 5–12, 3 1996.
- W. Winiwarter. Bewältigung der Informationsflut – Stand der Computerlinguistik. *Nachrichten für Dokumentation*, 47(3), 1996.
- I. Witten, C. Nevill-Manning, R. McNab, und S. Cunningham. A Public Library Based on Full-text Retrieval. *ACM Communications*, 41(4):71–75, 1998.
- M. Worrying und A. Smeulders. Internet access to scanned paper documents. In: *DL '98. Proceedings of the third ACM Conference on Digital libraries*, Seiten 313–314. ACM, 1998.
- P. Wright. Manual Dexterity: A User-Oriented Approach to Creating Computer Documentation. In: *Proceedings of ACM CHI'83 Conference on Human Factors in Computing Systems*, Plenary Speakers, Seiten 11–18, 1983.
- D. Wunderlich. Probleme der Wortstruktur. *Zeitschrift für Sprachwissenschaft*, 5(2): 209–252, 1986.
- H. Zimmermann. Automatische Indexierung und elektronische Thesauri. In: E. Niggemann und K. Lepsky, Hrsgg., *Zukunft der Sacherschließung im OPAC: Vorträge des 2. Düsseldorfer OPAC-Kolloquiums am 21. Juni 1995*, Vol. 25: *Schriften der Universitäts- und Landesbibliothek Düsseldorf*, Seiten 37–47. Univ.- und Landesbibliothek, 1996.
- M. Zoepritz. Endbenutzersysteme mit 'natürlicher Sprache' und ihre human factors. In: H. Balzert, Hrsgg., *Software-Ergonomie – Tagung I/1983 des German Chapter of the ACM*, Vol. 14: *Berichte des German Chapter of the ACM*, Seiten 397–410, Stuttgart, 4 1983. German Chapter of the ACM, Teubner.
- E. Zoltan, G. Weeks, und W. Ford. Natural-Language Communication with Computers: A Comparison of Voice and Keyboard Inputs. In: G. Johannsen und J. Rijnsdorp, Hrsgg., *Analysis, Design and Evaluation of Man-Machine Systems – Proceedings of the IFAC/IFIP/IFORS/IEA Conference Baden-Baden*, Seiten 255–260, Oxford, New York, Frankfurt, 9 1982. International Federation of Automatic Control, Pergamon Press.
- E. Zoltan-Ford. Reducing variability in natural-language interactions with computers. In: *Proceedings of the Human Factors Society – 28th annual meeting 1986*, Seiten 768–772, Santa Monica, CA, USA, 1984. Human Factors Society.

Erklärung

Ich versichere hiermit an Eides statt, daß ich die Dissertation mit dem Titel

„Dialogschnittstellen an Online-Informationssystemen: Notwendigkeit, Leistungsfähigkeit und Entwicklungsmöglichkeiten am Beispiel des OSIRIS-Systems“

selbständig und ohne unerlaubte Hilfe verfaßt und die benutzten Hilfsmittel vollständig angegeben habe.

Ich habe die Dissertation noch nicht als Prüfungsarbeit für eine andere wissenschaftliche Prüfung eingereicht und mit dieser Abhandlung noch keinen Doktorgrad erworben.

Marc Ronthaler