

**Aufbruchstimmung in der (Meta-)Lexikographie. Bemerkungen zu: Ingrid Lemberg, Bernhard Schröder und Angelika Storrer (Hrsg.): *Chancen und Perspektiven computergestützter Lexikographie*. Lexicographica Series Maior 107. Niemeyer: Tübingen. 2001 (270 Seiten, 144 DM)**

Lexikographie ist eine "eigenständige, kulturelle und wissenschaftliche Praxis", so führt dieser Band über einen aktuellen, aber trotzdem nicht mehr ganz jungen Zweig der Lexikographie, die computergestützte Lexikographie, den Leser in die Thematik ein. Zweifellos ist die Lexikographie, die an vielen (oft an sprachwissenschaftlichen) Instituten betrieben wird und gar in besonderen Fällen durch dedizierte Lehrstühle vertreten ist, eine eigene Wissenschaft. Unverkennbar ist in diesem Band jedoch die starke Wurzel in der Sprachlexikographie, insbesondere der Lexikographie der Allgemeinsprache und der historischen Lexikographie. Einzelne Beiträge z.B. über Flurnamenbücher entpuppen sich bei genauerer Lektüre als ebenso dem historischen Aspekt verbunden wie enthaltene Beiträge über "Goethes Wortschatz" oder "die Retrodigitalisierung mittelhochdeutscher Wörterbücher". Die (synchrone) Fachlexikographie, ein Zweig der Lexikographie, der nicht zuletzt durch Schlagworte wie *Wissensmanagement* und *Wissensgesellschaft* mit engem Bezug zur Wirtschaft einen stetig wachsenden Bereich in der Lexikographie einnimmt, ist in diesem Band nicht vertreten. - Das Interesse des Lesers kann sich damit bei der Beurteilung des Sammelbandes auf den inhaltlichen Aspekt der jeweiligen Artikel beziehen, um diese in ihrer Relevanz zu prüfen. In diesem Fall spannt sich der Bogen inhaltlicher Kompetenzen und Interessen der Artikel eben von "Goethes Wortschatz" bis zu einem deutsch-polnischen Wörterbuch. Diese Sichtweise würde dem Band jedoch aus zwei Gründen nicht gerecht. Zum einen sind fünf Artikel des Bandes auf einer allgemeineren Ebene gehalten, die sich zumindest im Titel inhaltlicher Schwerpunkte einzelner lexikographischer Projekte (also z.B. ein zweisprachiges Wörterbuch Deutsch-Polnisch) freimacht. Zum anderen ist der Titel des Sammelbandes ein deutlicher Hinweis auf die Intention der Autoren, diesen Band auf einer allgemeineren Ebene unter dem Sichtwinkel der computergestützten Lexikographie sehen zu wollen.

Der Sammelband enthält auf 270 Seiten zwei thematische Hauptkomplexe: "Grundlagen" (5 Beiträge) und "Anwendungen" (8 Beiträge). Entstanden ist diese Sammlung von Einzelbeiträgen aus einem "*Symposium Computergestützte Produktion und Publikation von Wörterbüchern*", das u.a. mit Unterstützung der Gesellschaft für linguistische Datenverarbeitung (Deutschland) 1998 in Heidelberg durchgeführt wurde. Zusammengefasst und erstellt wurde der Band im

Jahr 2000. - Obwohl der Band mit seinen Einzelbeiträgen teilweise thematische Bereiche wie "Hypertext" streift, in denen die Halbwertszeit des aktuellen Wissens bei zwei Jahren oder gar weniger liegt, haben die Beiträge auch 2001, drei Jahre nach den zugrundeliegenden Vorträgen, ihre Aktualität in den wesentlichen Eckpunkten behalten.

Der Beitrag von Georg Büchel und Bernhard Schröder "Verfahren und Techniken in der computergestützten Lexikographie" steht am Beginn des Bandes und lässt mit dieser Positionierung im Bereich "Grundlagen" des Bandes auch Grundlegendes vermuten. So ist denn dieser Artikel im wesentlichen eine Übersichtsdarstellung über die eng verwandten Auszeichnungssprachen SGML und XML sowie ein Einblick in die Datenmodellierung mit Datenbanken relationalen und objektorientierten Zuschnitts. Verständlich scheint im Zusammenhang mit den Auszeichnungssprachen die Aussage, SGML und XML seien insgesamt mächtig genug, jede Datenstruktur zu kodieren (Büchel/Schröder 2001: 16) nur dann, wenn z.B. im Vergleich mit relationalen Datenbanken das Konzept der Normalisierung nicht beachtet wird. Natürlich kann XML als Datenbasis prinzipiell jedes Datum aufnehmen, aber grundlegendes Merkmal einer Speicherung von Daten in Auszeichnungssprachen wird stets eine mehr oder weniger ausgeprägte Redundanz bestimmter, oft datenrelatierender Daten sein. Andererseits - und hier muss dem Büchel/Schröder widersprochen werden - beschränkt sich der Zeichensatz einer XML-Datei eben nicht auf den american standard code for information interchange, ASCII (Büchel/Schröder 2001: 17). Der universelle 16-Bit-Zeichensatz Unicode ist als Standardzeichensatz von XML eine wesentliche Erleichterung z.B. bei der mehrsprachigen Lexikographie (Geeb 2001). Alternativ können in XML dann die im Verhältnis zu ASCII wesentlich erweiterten ISO-Zeichensätze eingesetzt werden. ASCII wird damit für XML keine Rolle mehr spielen. - Im Rahmen der Datenmodellierung in relationalen Datenbanken wird ein Datenmodell vorgestellt (Büchel/Schröder 2001: 22), das sich in der Modellierung an wesentlichen Erfordernissen einer Druckausgabe orientiert. Eine kritischere Kommentierung und Erweiterung dieses im Rahmen eines anderen Projektes angewandten Datenmodells wäre hier wünschenswert gewesen. Insgesamt bringt dieser Beitrag eine Zusammenfassung einiger wichtiger Punkte computergestützter Lexikographie, ohne jedoch neue Wege oder Perspektiven aufzugzeigen oder diese zumindest anzudeuten.

Ingrid Schmidt und Carolin Müller stellen in ihrem Beitrag "Entwicklung eines lexikographischen Modells: Ein neuer Ansatz" interessante Überlegungen zu einem Publikationsmodell in der Lexikographie an. Klar zu unterstützen ist die hier formulierte Aussage, die Kompetenzen von Lexikographie und Neuen Medien seien zur Entwicklung guter lexikographischer Software zusammenzubringen. (Schmidt/Müller 2001: 30). Als Ausgangspunkt für das vorgeschlagene

Publikationsmodell wird eine Inhaltsstrukturmodellierung statt einer Datenmodellierung gesehen (Schmidt/Müller 2001: 33). Dabei sollen Daten nach genuinem Zweck in potentielle Informationseinheiten klassifiziert und dann in eine Struktur eingebunden werden. Diese Inhaltsstruktur trägt jedoch offenbar wesentliche Züge einer Mikrostruktur (Schmidt/Müller 46 f.). Genau besehen ist dies jedoch noch immer die Datenmodellierung. Erst durch die Einbindung von Daten in ein System und dessen Relevanzen werden Daten zu Informationen (Willke 1998: 8). Eine Benutzermatrix (Geeb 1998: 39 ff) kann hierbei in diesem lexikographischen Kontext als Set von Relevanzkriterien für ein System der Darstellung von Lexemwissen (Geeb 1998: 8) angesehen werden. - Die Datengrundlage in diesem Beitrag liegt dabei nicht in dem Bereich, den das Lexem *Wörterbuch* im Allgemeinverständnis (auch der anderen Beiträge in diesem Band) umfasst. Mit dem Datenbeispiel zum Lexem "Argentinien" (Schmidt/Müller: 35) zeigen die Autoren, dass Ihr Verständnis dem weitergefassten Terminus des *Lexikographischen Nachschlagewerks* (Geeb 1998: 35) entspricht. Nicht vergessen werden darf bei diesem Beitrag die intensive Auseinandersetzung mit den TEI-Richtlinien der Text Encoding Initiative. Neben zahlreichen wichtigen Kommentaren zu diesen Strukturregeln wäre jedoch die sehr auffällige und wenig hilfreiche Konstruktion eines Mixed Content in den Verdichtungsangaben der TEI-Richtlinien zu kritisieren (Schmidt/Müller 44, 46). - Schmidt/Müller haben in diesem Beitrag wichtige und interessante Punkte angerissen, die weiter vertieft werden können. Eine Auseinandersetzung mit der bestehenden lexikographischen Theoriebildung steht auch hier noch aus.

Angelika Storrer stützt sich in ihrem Beitrag "Digitale Wörterbücher als Hypertexte: Zur Nutzung des Hypertextkonzepts in der Lexikographie" wesentlich auf das Hypertextkonzept, das in den letzten Jahren neben einer erheblichen Verbreitung auch eine wissenschaftliche Grundlage erhalten hat. Ein wichtiger Aspekt dieses Beitrags ist die Aufzählung und Ausformulierung der Unterschieden von Print- und Hypertextwörterbüchern. Hierbei fällt jedoch auf, dass Hypertextsysteme mit Eigenschaften verbunden werden, die nicht unbedingt hypertextspezifisch sind und deren Fehlen andererseits nicht genügt, einem Hypertextsystem dieses Prädikat zu entziehen. So sind "Such- und Recherchewerkzeuge" sowie die Einbindung "medialer Objekte" (Storrer 2001: 55) zwar eine Eigenschaft der Hypertextsysteme, die wir z.Zt. unstrukturiert als "Internet" bezeichnen, sie sind aber nicht klassenspezifische Merkmale von Hypertextsystemen allgemein. Bei einer - im Artikel nicht vorhandenen - möglichen Erweiterung des Beitrages auf Hypermediasysteme wäre der zuletzt genannte Punkt der eingebundenen Medien dagegen relevant. Eine ähnliche terminologische Unschärfe findet sich in der Aussage "Als E-Texte (elektronische Texte) bezeichnet man Texte, die in das World Wide Web eingebunden sind, ..". Alleine die Tatsache

von elektronischen Volltextdarstellungen auf CD-ROMS oder in E-Book-Readern relativieren diese Aussage. Auch die Beschränkung auf eine "lexikalische Datenbank" als Ziel einer lexikographisch motivierten Modellierung von Daten (Storrer 2001:61) ist nicht einsichtig. Zunächst muss das Hauptaugenmerk auf die lexikalische / lexikographische Datenbasis gelegt werden. Ob die dadurch entwickelte Struktur in einer Datenbank (relational, objektrelational, objektorientiert etc.) abgebildet wird, ist in einem weiteren Diskussionsschritt zu entscheiden. Der Terminus "Datenbank" scheint an diesem Punkt daher unmotiviert. Zweifelhaft ist auch, ob die Aussage, Speicherplatz in digitalen Medien sei keine Kostefrage mehr (Storrer 2001, 63), so aufrecht erhalten werden kann. Bei großen Datenmengen in verteilten Systemen und hier vor allem im Internet ist Speicherplatz mittelbar durch Netzkapazität und die damit verbundenen Kosten (und in der gewerblichen Anwendung auch die Wartezeiten) sehr wohl ein Kostenfaktor. Nicht zutreffend sind Aussagen zur Nutzungsprotokollierung im Internet. Die Nutzung von Internetwörterbüchern wird entgegen Storrer (Storrer 2001: 67) nicht aussagekräftig protokolliert. Die Aussagekraft dieser Protokolle endet in über 50% der Fälle bei einer anonymisierten Adresse, die höchstens zu vernachlässigende Rückschlüsse auf den Nutzer zulässt. In den übrigen Fällen ist der Endnutzer zwar namentlich (oder zumindestens rechner-namentlich) identifizierbar, die Benutzerstruktur (Geeb 1998: 46 ff.) jedoch - außer bei vorheriger Anmeldung am System mit gleichzeitiger Benutzerbefragung - nicht bekannt. Auch wenn weitere Ungenauigkeiten im Bereich der Informationstechnologie vorliegen, ist jedoch die Gesamtaussage des Beitrages, die in 7 Thesen formuliert wird, zukunftsweisend und für die weitere Forschung inspirierend.

Ingrid Lemberg argumentiert in ihrem Beitrag "Aspekte der Online-Lexikographie für wissenschaftliche Wörterbücher" für den vielversprechenden und treffenden Fachterminus *Mehrwert*. Damit ist das zweifellos vorhandene "Mehr" an Informationen im Wörterbuch und an Informationsmöglichkeiten für die Wörterbuchnutzerinnen und -benutzer" (Lemberg 2001: 73) umschrieben. Auch hier werden dann die bereits in anderen Beiträgen dieses Bandes formulierten Aspekte des Hypertextes, der erheblichen Speicherkapazitäten und der Dynamik in Nutzung und Erstellung des Wörterbuchs genannt. - Nicht unumstritten ist die Definition von Hypertext als "computerverwaltete, nicht-lineare Texte, die die Mehrachkodierung von Daten in verschiedenen Medialen Formen (Schrift, Graphik, Ton und Video) erlauben" (Lemberg 2001: 74). Richtungsweisende Arbeiten der Usability-Forschung definieren genau diese Informationsstrukturen als "Hypermedia" (Thissen 2000). In ähnlicher Weise kann auch die Definition von Multimedia als "durch Hypertextsysteme ermöglichte Integration von Text, Videoelementen (...) sowie Audioelementen (...)" (Lemberg 2001: 80) diskutiert werden. Multimedia muss nicht zwangsweise mit dem Konzept des Hypertext verbunden

werden, ist aber Bestandteil von Hypermedia (s.o.). Wichtig ist in diesem Beitrag vor allem auch der Verweis auf das Entstehen offener, dynamischer Wörterbücher, “denen fortlaufend alle Arten von Daten (neue Quellen, neue Zitate, neue Wörter usw.) hinzugefügt werden können bzw. müssen” (Lemberg 2001: 86). Auch wenn die dazu genannten Mittel nicht unbedingt überzeugen, so ist die Aussage auf programmatischer Ebene zweifellos relevant.

Annette Klosa rückt in dem Beitrag “Qualitätskriterien der CD-ROM-Publikation von Wörterbüchern” den Benutzer in den Fokus: “Ein gutes Wörterbuch ist eines, das die Erwartungen der Nutzer(innen) erfüllt und ihre Bedürfnisse befriedigt” (Klosa 2001: 94; ähnlich auch Klosa 2001: 97). Allerdings fällt bereits bei der eingangs in diesem Beitrag eingesetzten Wörterbuchdefinition, die der Definition des lexikographischen Nachschlagewerkes (s.o.) sehr nahe kommt, auf, dass eine Beschränkung auf CD-ROM-Wörterbücher in diesem Beitrag unverständlich erscheint. Wenn diese Einschränkung auf ein Medium (CD-ROM) auf der Grundlage eines begrenzten empirischen Materials gewählt wurde, wäre dies zu formulieren. Im anderen Fall ist zu prüfen, ob die als Qualitätskriterien für CD-ROM-Wörterbücher entwickelten Maßstäbe nicht auch für andere rechnergestützte Wörterbücher (Stichwort Hypermedia) gelten können. Diese fehlende oder zumindest nicht formulierte Differenzierung wird noch unterstrichen, wenn von einem guten CD-ROM-Wörterbuch eine “Aktualisierbarkeit des Datenbestandes” (Klosa 2001: 97) erwartet wird. Die gekaufte Wörterbuch-CD-ROM ist wohl kaum aktualisierbar; aktualisierbar ist mittelbar der Datenbestand auf einem beschreibbarem Medium, dessen Datenbestand natürlich mit dem ersteren verknüpft werden kann. Eine Funktionalität, die jedoch im Gegensatz zu verteilten Systemen in Hypermedia, diesen Datenbestand dann an einen Rechner bindet. - Korrekt und wichtig ist die Aussage, dass bei der Erarbeitung von CD-ROM-Wörterbüchern (und allen rechnergestützten Wörterbüchern - auch in verteilten Systemen) lexikographisches Wissen und Wissen aus der Informatik zusammenkommen müssen. Auch die folgende Aussage, in der Lexikographen zu Horizonterweiterung in Richtung der Neuen Medien auffordern ist zu unterstützen (Geeb 2001).

Im Beitrag von Ulrike Hass-Zumkehr “Zur Mikrostruktur im Hypertext-Wörterbuch” ist die durch Artikelaufbau und Inhalt erwartete Verbindung von Mikrostruktur zur unterliegenden Datenmodellierung nicht deutlich erkennbar.

Zudem fällt auf, dass für den Beitrag wichtige Lexeme wie “Information” oder “Hypertext” nicht erläutert werden (Hass-Zumkehr 2001: 103, 104). Ein Verweis auf andere Beiträge zu diesem Thema (Hass-Zumkehr 2001: 103) ist nicht unbedingt befriedigend aber verständlich. Die Hinnahme eines völlig undefinierten Fachterminus wie “Information” mit dem Verweis dieser Terminus werde “absichtlich völlig unterterminologisch verwendet” (Hass-Zumkehr 2001: 104) ist unver-

ständig. Der Beitrag zur Mikrostruktur im Hypertextwörterbuch schöpft zwangweise wesentliche Inhaltsaspekte aus der Abgrenzung von Daten und Information, von Datenmodellierung und Informationsstruktur. - Unverständlich auch die These, dass eine DTD als Grundlage einer Datenbankmodellierung wichtig sei, "weil man zu Beginn noch nicht weiß, welche Datenbanksoftware man braucht" (Hass-Zumkehr 2001: 106). Zu diesem Zweck - zur Modellierung von Daten unabhängig von einem Datenbankmanagementsystem - ist ein reichhaltiger Bestand an übergeordneten Methoden vorhanden. Von Entity-Relationship-Diagrammen über Datenstrukturdiagramme bis zu Relationenschemata und Data Definition Language-Scripts in SQL seien dies nur einige der bekannteren Methoden und Arbeitsschritte für eine plattformunabhängige Modellierung von Daten für Datenbanken. DTDs sind für eine Modellierung von Datenbankinformationen aus zahlreichen Gründen nur bedingt geeignet. - So ist auch nicht zu verstehen, warum Mikrostrukturen in Wörterbüchern in DTDs abgebildet werden sollen (Hass-Zumkehr 2001: 106). DTDs und deren spezifischere Ausformung in XML-Schemata sind Datendefinitionen und lassen sich zum Zeitpunkt der Datenmodellierung mit der Mikrostruktur nur bedingt in Verbindung bringen. Diese Mikrostruktur wird im Fall von XML bei direktem Rendering generiert durch z.B. durch XSLT (Geeb 2001). Damit sind relativ beliebige Mikrostrukturen auf einer gemeinsamen Datengrundlage (als Grammatik der Daten definiert in der DTD) möglich. So sind es denn auch nicht "andere Inhaltsaspekte, die die Datenmodellierung einerseits und die Gestaltung der Nutzeroberfläche andererseits leiten" (Hass-Zumkehr 2001: 107), sondern die Datenmodellierung geht von den vorhandenen oder zu erwartenden oder gewünschten Daten aus und blickt auf die zu erwartenden Informationen. Im Kontext mit der Benutzerstruktur (Geeb 1998: 46 ff.) wird die Information dann zu einer oder mehreren Mikrostrukturen. Die Tatsache, dass die Mikrostruktur nur Daten als Information abbilden kann, die auch in der Datenbasis vorhanden sind, ist selbstredend. Grundlage des Designprozesses ist aber die Datenmodellierung, nicht der Entwurf der Mikrostruktur. Aus dieser unvorteilhaften Vermischung von Datenmodellierung und Mikrostruktur erwächst dann auch die Darstellung der "informationellen Dimensionen" eines Wörterbuchartikels und eng verbunden damit die Beschreibung von "Schaltflächen" sowie deren Inhalt und Benennung in dem konkreten System (Hass-Zumkehr 2001: 109). Auf einer anderen Ebene zeigt sich diese Verquickung von Fragen zur Datenmodellierung und Aspekten einer konkreten Systemnutzung in dem Hinweis, bei Tags in der DTD (Anmerkung des Verf.: in einer DTD gibt es keine Tags oder Elemente sondern nur Deklarationen dieser) seien Abkürzungen ausreichend, "die die AutorInnen verstehen" (Hass-Zumkehr 2001: 110).

Zur Diskussion stehen dann in diesem Beitrag verschiedene Linktypen, die sich jedoch nicht als allgemeine Klassifikation von Verweisen ansetzen lassen, da sie auf der Datenstruktur des konkreten Systems beruhen. Vor allem die Aussage, dass ein "elektronischer Verweis im Prinzip bidirektional" sei ist nur bedingt richtig. Aus technischer Sicht ist diese Zweiseitigkeit noch nicht wirklich implementiert (XLink/XPointer). Eine in einzelnen Projekten softwaremäßig implementierte Zweiseitigkeit kann nicht verallgemeinert werden; die "Back"-Funktionalität eines Browsers ist technisch nicht in dem von der Autorin genannten Bereich anzusiedeln, da sie auf dem individuellen Informationsweg des Benutzers beruht.

Der Weg von der Datenbasis und Datenstruktur über die Datenmodellierung, der Entscheidung für ein Datenbankmanagementsystem oder ein Auszeichnungssystem sowie die Rolle des Hypertextes in der Datenmodellierung sowie in der Mikrostruktur sind in diesem Beitrag nicht klar getrennt worden. Die Forschungsthematik dagegen ist brisant und wichtig.

Thomas Gloning und Rüdigers Welters Beitrag "Wortschatzarchitektur und elektronische Wörterbücher: Goethes Wortschatz und das Goethe-Wörterbuch" führen mit dem Terminus "Wortschatzarchitektur" eine interessante Komponente in die Diskussion ein. Eng verwandt mit diesem Terminus scheint das von den Autoren eingesetzte Lexem "Wortschatzstrukturierung" (Gloning/Welter 2001: 118) zu sein, das offenbar - eine konkrete Definition fehlt - diesen Architekturvorgang ausmacht. Unklar ist, wer der Akteur der Architektur oder Strukturierung ist, wenn wie in diesem Fall von einem historisch und sprachlich abgegrenzten Bereich ("Goethe") ausgegangen werden kann. Klar werden jedoch in diesem Beitrag die verschiedenen, von den Autoren angesetzten "lexikologischen Organisationsprinzipien" - vermutlich das Architekturprinzip - (Gloning/Welter 2001: 119, 128f), die dann als "besondere lexikalische Herausforderung" in eine "elektronische Datenbasis" mit "Markierungen" für die verschiedenen Benutzungsmöglichkeiten umgeformt werden (Gloning/Welter 2001: 119). Dieses treffende und für die Datenmodellierung sowie Informationsstrukturierung ideal neutrale Lexem der "lexikalischen Datenbasis" (s.o.) entspricht offenbar der "lexikologischen Datenbasis" (Gloning/Welter 2001: 123), wobei dann aber die Aussage, dass diese Menge gleich dem elektronischen Wörterbuch sei, verwundert. Neben der Datenbasis besteht das elektronische Wörterbuch noch aus verschiedenen weiteren Teilen; eine Makro und Mikrostruktur lassen sich in jedem Fall in einem noch so rudimentären elektronischen Wörterbuch ausmachen. - In dem von den Autoren beschriebenen Projekt wurde SGML zur Kodierung der Datenbasis herangezogen. Leider vertiefen die Autoren ihre Bewertung dieses Werkzeuges außer mit den Worten "passende Umgebung" (Gloning/Welter 2001: 123) nicht weiter. Die Feststellung "Neben den erweiterten Auswertungs- und Abfrage-

möglichkeiten bietet eine SGML-Datenbasis den Vorteil, dass sich die Auswahl der Informationstypen und die Dokumentationstiefe vom Benutzer variieren lassen.“ (Gloning/Welter 2001: 125) ist hier wenig hilfreich. Denn offenbar werden hier das Werkzeug zur Datenlagerung (SGML) mit der darüber gelegten Benutzerschnittstelle (ist hier nicht namentlich genannt) sowie dem Umfang, der Tiefe und der Qualität der gespeicherten Daten vermischt. SGML als Datenbasis hat keine Funktionalität. SGML verfügt daher über keinerlei “Auswertungs- und Abfragemöglichkeiten”. Erst die Software zu SGML kann diese Möglichkeiten implementieren. Datenbanken an sich haben ebenso keinerlei Funktionalität (als Datenbasis). Erst die Datenbankmanagementsoftware gibt diese Möglichkeiten (z.B. mit SQL). Daraus folgt, dass die zitierte Aussage (Gloning/Welter 2001: 125 f.) der Autoren zur Wahl von SGML nicht als Begründung dieser Wahl angesehen werden kann. So müsste denn die Feststellung, dass eine SGML-Datenbasis gegenüber der gedruckten Fassung einfacher revidiert und aktualisiert werden kann (Gloning/Welter 2001: 127) allgemeiner formuliert werden als “rechnergestützte Datenbasis”. Gleiches gilt für die Formulierung, dass aus einer Datenbasis in SGML Internet-Versionen erstellbar sind (Gloning/Welter 2001: 127). Zweifellos ist dies eine Möglichkeit, die jedoch bei weitem nicht auf SGML beschränkt ist. - Unterstützenwert ist die Forderung der Autoren, die Inhalte eines solchen Projektes im Sinne einer Open Content-Initiative gratis im Internet zur Verfügung zu stellen. - Der Beitrag hat damit wichtige Punkte einer Strukturdiskussion aufgegriffen, jedoch die Struktur der Datenbasis nur teilweise behandelt. Auch hier ist eine teilweise terminologische Unschärfe zu beobachten.

Die TEI-Richtlinien werden von Thomas Burch und Johannes Fournier in dem Beitrag “Zur Anwendung der TEI-Richtlinien bei der Retrodigitalisierung mittelhochdeutscher Wörterbücher” auf ihre Anwendbarkeit untersucht. Die Richtlinien der TEI können dabei nach Aussagen bei der nachträglichen Erfassung (Digitalisierung) von den ausgewählten Wörterbüchern nur bedingt eingesetzt werden. Über diese Kernaussage hinaus, die aufgrund der empirischen Grundlage nicht ohne Einwände verallgemeinert werden kann, wird die TEI-DTD zu diesem Thema genauer untersucht und folgerichtig bewertet. Nicht klar ist die Aussage, dass durch das gewählte Verfahren in dem der Wörterbuchinhalt “nur noch im Markup präsent ist” eine “Wörterbuchdatenbank” effizient aufgebaut werden könne (Burch/Fournier 2001: 150 f.). So lässt sich hier eine nicht sinnvolle Nutzung des Terminus “Datenbank” vermuten, wenn die SGML-Basis als Wörterbuchdatenbank bezeichnet wird. Andererseits könnte - dieser Gedanke wird jedoch nicht ausgesprochen - auch beabsichtigt werden, den Markupcode in SGML als Grundlage einer Datenbankmodellierung (z.B. relational) zu verwenden. Dieser Gedankengang betritt teilweise Neuland und wäre dann genauer

zu formulieren. - Im vorliegenden Fall stehen jedoch beide Lesearten zur Verfügung, und das Verständnis des Artikels wird damit erschwert.

Ebenfalls der historische Perspektive oder dem historischen Material in der Wörterbucharbeit haben sich Ralf Plate und Ute Recker mit dem Beitrag "Elektronische Materialgrundlage und computergestützte Ausarbeitung eines historischen Belegwörterbuchs. Erfahrungen und Perspektiven am Beispiel des neuen Mittelhochdeutschen Wörterbuchs" verschrieben. In diesem Werkstattbericht werden in einer fundamentierten Argumentation Möglichkeiten der rechnergestützten Lexikographie dargestellt. Zur Datenstrukturierung oder dem Informationsretrieval werden keine neuen Sichtweisen präsentiert. Doch ist die deutliche Forderung nach einer Verbindung des Wörterbuchs mit seinen Quelltexten wichtig und folgerichtig (Plate/Recker 2001, 168). Dies wiegt um so schwerer, als die Autoren darauf verweisen, dass das elektronische Wörterbuch die entsprechende Druckfassung in der Zukunft im Regelfall ersetzen wird (Plate/Recker 2001, 170). Auch die mehrfach formulierte Suchmöglichkeit in Feldern (Plate/Recker 2001, 170 f) ist eine wichtige Perspektive zum Aufbau elektronischer Nachschlagewerke. Zumal die viel gelobte Volltextsuche oftmals "zu einem großen Teil unspezifische Daten" liefert (Plate/Recker 2001, 171).

Auf ebenfalls historischen und gleichzeitig sprachgeschichtlichen Gebiet bewegt sich der Beitrag "Das elektronische Flurnamenbuch - Innovation in der Flurnamenforschung durch den Einsatz neuer Medien" von Gerd Richter. Auch hier werden der Übergang vom Printmedium zum elektronischen Medium beschrieben und Vergleiche gezogen. Unverständlich erscheint hier jedoch die Aussage "Vor allem die Abkürzungen, die im Printmedium als potentielle Verweise vorliegen, müssten im Hypertext expliziert werden" (Richter 2001, 192). Zum einen ist dies eine Frage der Benutzung und der Mikrostruktur und müsste daher auch in einem entsprechenden fachterminologischen Zusammenhang gestellt werden. Zum anderen ist die Verlinkung von Abkürzungen zu deren Auflösung - wenn überhaupt gewünscht - keinesfalls eine Anhäufung von Links als redundante Daten. Das Verweiszil ist nicht redundant, die Anhäufung der Links nur über mehrere Artikel o.ä. hinweg erkennbar und daher kein wirkliches Problem der Nutzung. - Wichtig in diesem Beitrag ist die Aussage von Richter, dass durch eine umfangreiche Datendokumentation auch der Bedarf der Strukturierung dieser Daten wichtiger wird um "(a) nicht nutzlose, weil nicht benutzbare Datenmüllhalden zu produzieren und (b) mögliche Rezeptions- und Navigationsprobleme zu vermeiden" (Richter 2001: 194 f). Auch in diesem Beitrag bleibt jedoch eine konkrete, umfassende Diskussion der Datenstrukturierung auf konzeptionellem Niveau aus.

Krzysztof Petelenz Beitrag "Das Informationsdesign auf der Speicherungsebene eines zweisprachigen Online-Wörterbuchs Polnisch-Deutsch" beschäftigt

sich intensiv mit den Anforderungen verschiedener Benutzerstrukturen an ein Wörterbuch (Petelenz 2001, 213 ff.) Allerdings fällt auf, dass der Autor unter dem Terminus "Datenmodellierung" (Petelenz 2001, 215) offenbar die Modellierung der "Informationstypen" (Petelenz 2001, 214) versteht. Die Differenz zwischen Information und Daten und die damit folgenden Unterschiede in der Modellierung der beiden gedanklichen Einheiten ist wichtig für das Verständnis, wird aber nicht formuliert. Ein Datenmodell im klassischen Sinne fehlt. Die Strukturierung oder Modellierung - beides scheint hier möglich - wird scheinbar auf dem Niveau der Informationen durchgeführt. Interessant ist dabei die graphische Darstellung von Abhängigkeiten der Informationstypen (Petelenz 2001, 216 f). Nur am Rande für das Informationsdesign von Interesse sind dagegen Aussagen zu der Benutzeroberfläche. Die Benutzeroberfläche ist auf der Grundlage der Informationstypen zunächst frei gestaltbar und wäre in einem gesonderten Beitrag zu diskutieren. Die Formulierung "Speicherungsebene" lässt dagegen zum einen eine Diskussion der Speicherplattform (Markup, Datenbanken oder andere) und zum anderen eine Diskussion der Speicherstruktur (DTD, Schema, ERD etc.) erwarten. Das objektorientierte Modell wird im vorliegenden Beitrag erwähnt und treffend charakterisiert (Petelenz 2001: 219 f.), die konkrete Ausformung hiervon auf der "Speicherungsebene" bleibt aber offen. - Der Wunsch nach Koordination von Webprojekten durch eine gemeinsame Software ist verständlich (Petelenz 2001, 223), ist aber wiederum kein Beitrag zur Frage der "Speicherungsebene" elektronischer Wörterbücher, zumal die Beschränkung auf Webprojekte nicht zwingend mit dem Informationsdesign verbunden werden kann. Es hätte dann für ein Informationsdesign in verteilten Systemen argumentiert werden müssen. - Nicht zuletzt die Anwendung und Erweiterung von Hupkas Theorien zum lexikographischen Bild (Petelenz 2001, 220 ff.) und die starke Forderung nach der Benutzerfokussierung des elektronischen Wörterbuchs mit den Worten "erst wenn Hypertextwörterbücher den individuellen Bedürfnissen verschiedener Benutzergruppen entsprechen werden, kann ihre Akzeptanz und damit auch praktische Relevanz steigen" (Petelenz 2001: 224) betonen die wichtigen Aussagen dieser Überlegungen.

Mit treffsicherer Formulierung sehen Claudia Kunze und Andreas Wagner in ihrem Beitrag "Anwendungsperspektiven des GermaNet, eines lexikalisch-semantischen Netzes für das Deutsche" bereits zu Beginn des Artikels, dass elektronische Wörterbücher "ein zentrales Desiderat der Informationsgesellschaft" (Kunze/Wagner 2001: 229) sind. Damit geht die hier formulierte Auffassung von Wörterbüchern wesentlich über eine oft angetroffene Definition hinaus. Bereits die empirische Grundlage dieses Beitrages in einem "lexikalisch-semantischen" Netz lassen diese umfassende Sichtweise erwarten. Die Disambiguierung von Lexemwissen (Geeb 1998: 8) wird auf verschiedene Einsatzbereiche angewandt.

Auch die verschiedenen genannten Relationen zwischen Lexemen sind durch die stetige Einordnung in Unter-, Über- und Gleichordnung richtungsweisend. Vor allem auch die Vermeidung des Lexems "Begriff" zur Darstellung der Inhaltsseite des sprachlichen Zeichens (Kunze/Wagner 2001: 231 ff.) in den verschiedenen Beziehungstypen und der Kreuzklassifikation ist wertvoll. Allzuleicht lassen sich durch Rückgriff auf die Vorstellung eines übersprachlichen Begriffs Probleme z.B. zur Äquivalenz der mehrsprachigen Lexikographie scheinbar lösen (Geeb 1998: 16 ff.). Im vorliegenden Beitrag wird diese Problematik jedoch erkannt und ausformuliert (Kunze/Wagner 2001: 237). So verwundert auch nicht, dass die Autoren einen möglichen Einsatz ihres Systems in der automatischen Informationserschließung sehen (Kunze/Wagner 2001: 240).

"Das Internet als Medium für die Wörterbuchbenutzungsforschung" von Lothar Lemnitzer ist ein Beitrag, der die Möglichkeiten der Benutzerforschung durch Datenspeicherung zu den Nutzern von Internetwörterbüchern auslotet. Die Notwendigkeit der Benutzerforschung wird korrekt belegt (Lemnitzer 2001: 247). Es ist jedoch grundlegend zu dieser empirischen Untersuchung anzumerken, dass Protokolle von Webservern in der hier angeführten Form nur geringen Aufschluss über den Benutzer geben. Wenn von Systemen ohne Passwortschutz und/oder namentlicher Benutzeranmeldung ausgegangen wird, sind in der Mehrzahl die Angaben zur Identität des Benutzers wertlos. Diese Angaben werden in der Regel anonymisiert und lassen sich dann nur noch auf Domänen (z.B. ein Internetprovider oder ein großes Unternehmen) zurückführen. Aus dem Benutzerprofil (Geeb 1998: 39 ff.) sind so die gesamte Benutzerstruktur (Geeb 1998: 46 ff.) und damit die Angaben zur den informationellen Bedürfnissen des Benutzers nicht einsehbar. Besonders wesentlich ist dieser Punkt auch, da mehrere verschiedene Benutzer der selben Domäne (z.B. aus dem selben Unternehmen) im genannten Fall nicht unterschieden werden können. Es können damit lediglich Aussagen über die angeforderten Informationen/Daten gegeben werden (Lemnitzer 2001: 251 ff.). Da aber die Bindung dieser Informationen an die Benutzerstruktur fehlt, bleibt diese Information unproduktiv. Sieht man von den lexikographischen Ansprüchen an die Benutzerforschung ab, machen sich aber auch in anderen Bereichen diese Mängel bemerkbar. So ist ein zielgerichtetes Marketing - z.B. von Wörterbüchern - ohne Wissen über den Käufer und seine Merkmale ebenfalls nicht möglich.

Im Vorwort fassen die drei Herausgeber ihre Sicht auf Lexikographie im Computerzeitalter im wesentlichen in fünf Punkten zusammen:

- Datenbank- und Hypertextsysteme, Computernetzwerke und Speicherkapazitäten ermöglichen weitreichende Verknüpfungen lexikographischer Informationen

- Der Prozess der Neuordnung lexikographischer Theorien und Theoriebildung auf dieser Grundlage hat gerade erst begonnen
- Schwerpunkte metalexikographischer Arbeit wie Makro-, Mikro-, Zugriffs- und Verweisstrukturen müssen neu überdacht werden
- Eine Diskussion der Anordnung lexikographischer Information in der Lemmalist ist zweitrangig; der Rechner ermöglicht je nach Strukturierung der lexikographischen Informationen zahlreiche Lemmalisten u.ä.
- Auf der Seite der Produktion lexikographischer Information revolutioniert und vereinfacht der gezielte, überlegte Rechnereinsatz die Arbeit des Lexikographen

(Lemberg, Schröder/Storrer 2001: 1 ff.)

Allen diesen Punkten kann man nur zustimmen. Zweifellos sind Differenzierungen erforderlich, aber in der Grundthese ist klar, dass die Metalexikographie und die praktische Lexikographie durch den Rechnereinsatz in Produktion und Anwendung lexikographischer Information wesentlich verändert wird. Dieser Änderungsprozess ist momentan in Umfang und Dauer nicht abzuschätzen. Sicher ist auch, dass die Lexikographie seit jeher - und heute noch mehr denn je - auf viele andere Bereiche der Informationsstrukturierung, Recherche, des Retrievals und vor allem der Modellierung ausgreifen kann und sollte. Ein Ausgreifen, das stets ein Wechselspiel im Theoriaustausch mit diesen zahlreichen benachbarten Wissenschaften sein wird und muss.

Zusammenfassungen aller Artikel in englischer Sprache sowie ein Register, das sich auf den gesamten Band erstreckt, runden dieses Werk ab. Zweifellos ist die fachliche Ausrichtung, der qualitative Anspruch und die Theorietiefe der verschiedenen Beiträge sehr unterschiedlich. Sicher ist dieser Band jedoch ein gutes Abbild der Chancen computergestützter Lexikographie, wie sie 1998 bestanden und wie sie auch heute noch die Grundlage weiterer Forschung bilden. Sicher ist jedoch auch, dass gerade die Fachlexikographie mit ihrem immensen Nutzungspotential und tatsächlichen Nutzungseinsatz in einsprachiger und mehrsprachiger Ausprägung hier Beachtung verdient hätte. - Ebenso misslich und vielleicht in der mangelnden Beachtung der Fachlexikographie begründet ist der Mangel an Referenzen zur entsprechenden Literatur der Terminologiewissenschaft, der Informationswissenschaft oder der Usability-Forschung oder gar der Informatik. Lexikographie - gerade die computergestützte Lexikographie im Zeitalter des Hypertextes - kann sich nicht nur auf eigenes, linguistisches Gedankengut verlassen. In der klassischen (Meta-)Lexikographie sind erhebliche Prägnanzen der verfügbaren Theorien durch einige herausragende Forscher dieser Disziplin festzustellen. In dem vorliegenden Band wird die Aufbruchsstimmung in der rechnergestützten Lexikographie deutlich. Doch das neue oder das neu formulierte Gedankengut dieser Forschungsdisziplin steht im Gegensatz zu der

klassischen (Meta-)Lexikographie nicht selten noch auf terminologisch unsicherem Grund.

### Literatur

- Büchel, Gregor & Schröder, Bernhard (2001). Verfahren und Techniken in der computergestützten Lexikographie. In *Lemberg, Schröder & Storrer (2001)*. 7-28.
- Burch, Thomas & Fournier, Johannes (2001). Zur Anwendung der TEI-Richtlinien bei der Retrodigitalisierung mittelhochdeutscher Wörterbücher. In *Lemberg, Schröder & Storrer (2001)*. 133-154.
- Geeb, Franziskus (1998). Semantische und enzyklopädische Informationen in Fachwörterbüchern. Eine Untersuchung zu fachinformativen Informationstypen mit besonderer Berücksichtigung wortgebundener Darstellungsformen. Århus: Handelshøjskole Århus (Online-Version unter [http://www.nordiclex.org/geeb/geeb\\_1998.pdf](http://www.nordiclex.org/geeb/geeb_1998.pdf)).
- Geeb, Franziskus (2001). eXtensible Markup Language (XML) og leksikografi. In *LexicoNordica* 8. 167-182.
- Geeb, Franziskus (2002). leXeML - Vorschlag und Diskussion einer metalexikographischen Auszeichnungssprache. In *Sprache und Datenverarbeitung* (im Druck).
- Gloning, Thomas & Welter, Rüdiger (2001). Wortschatzarchitektur und elektronische Wörterbücher: Goethes Wortschatz und das Goethe-Wörterbuch. In *Lemberg, Schröder & Storrer (2001)*. 117-132.
- Hass-Zumkehr, Ulrike (2001). Zur Mikrostruktur im Hypertext-Wörterbuch. In *Lemberg, Schröder & Storrer (2001)*. 103-116.
- Klosa, Annette (2001). Qualitätskriterien der CD-ROM-Publikationen von Wörterbüchern. In *Lemberg, Schröder & Storrer (2001)*. 93-102.
- Kunze, Claudia & Wagner, Andreas (2001). Anwendungsperspektiven des GermaNet, eines lexikalisch-semantischen Netzes für das Deutsche. In *Lemberg, Schröder & Storrer (2001)*. 229-246.
- Lemberg, Ingrid (2001). Aspekte der Online-Lexikographie für wissenschaftliche Wörterbücher. In *Lemberg, Schröder & Storrer (2001)*. 71-92.
- Lemberg, Ingrid, Schröder, Bernhard & Storrer, Angelika (Hrsg.) (2001). Chancen und Perspektiven computergestützter Lexikographie. *Lexicographica. Series Maior* 107. Tübingen: Niemeyer.
- Lemnitzer, Lothar (2001). Das Internet als Medium für die Wörterbuchbenutzerforschung. In *Lemberg, Schröder & Storrer (2001)*. 247-254.
- Petelenz Krzysztof (2001). Das Informationsdesign auf der Speicherungsebene eines zweisprachigen Online-Wörterbuchs Polnisch-Deutsch. In *Lemberg, Schröder & Storrer (2001)*. 199-228.

- Plate, Ralf & Recker, Ute (2001). Elektronische Materialgrundlage und computergestützte Ausarbeitung eines historischen Belegwörterbuchs. Erfahrungen und Perspektiven am Beispiel des neuen Mittelhochdeutschen Wörterbuchs. In *Lemberg, Schröder & Storrer (2001)*. 155-178.
- Richter, Gerd (2001). Das elektronische Flurnamenbuch - Innovationen in der Flurnamenforschung durch den Einsatz neuer Medien. In *Lemberg, Schröder & Storrer (2001)*. 179-198.
- Schmidt, Ingrid & Müller, Carolin (2001). Entwicklung eines lexikographischen Modells: Ein neuer Ansatz. In *Lemberg, Schröder & Storrer (2001)*. 29-52.
- Storrer, Angelika (2001). Digitale Wörterbücher als Hypertexte: Zur Nutzung des Hypertextkonzepts in der Lexikographie. In *Lemberg, Schröder & Storrer (2001)*. 53-70.
- Thissen, Frank (2000). *Screen Design Handbuch*. Berlin, Heidelberg: Springer.
- Willke, Helmut (1998). *Systemisches Wissens Management*. Stuttgart.

*Franziskus Geeb*