

## Das Informationsdesign auf der Speicherungsebene eines zweisprachigen Online-Wörterbuchs Polnisch-Deutsch

1	Vorbemerkung	4.3	Die expandierten Informationstypen für die Wörterbuchinstanzen
2	Das Computerwörterbuch als Hypermediasystem	4.4	Die Behandlung der Phraseme
2.1	Die Tendenz zur Kumulation der lexikographischen Inhalte	5	Zum Problem der Standardisierung
2.2	Drei Generationen von zweisprachigen Hypermediawörterbüchern	6	Die expandierten lexikographischen Datentypen in der Hypertextbasis
2.3	Neue Perspektiven für zweisprachige Computerwörterbücher	6.1	Verschiedene Angabetypen für verschiedene Wörterbuchbenutzungssituationen
3	Zu den Besonderheiten der zweisprachigen Lexikographie	6.2	Die Datenmodellierung
4	Der typenübergreifende Wörterbuchserver: Zur allgemeinen Vorgehensweise	7	Übersichtsknoten
4.1	Das Schema des Wörterbuchsystems	8	Das objektorientierte Datenmodell
4.2	Die Typisierung als Grundlage der wissensbasierten Relationierung von Hypertexteinheiten	9	Die Rolle der Bilder
		10	Ausblick
		11	Literatur
		11.1	Wörterbücher und Lexika
		11.2	Sonstige Literatur

### 1 Vorbemerkung

Seitdem Nachschlagewerke auf CD-ROMs erscheinen, vergrößert sich die praktische und kommerzielle Bedeutung des Mediums Hypertext für die Lexikographie.<sup>1</sup> Das Aufkommen des WWW und der Hypertext-Sprache HTML brachte neue Qualitäten der Informationsverknüpfung mit sich: offene Systeme, aktuelle Inhalte, erleichterte Kommunikation zwischen den Autoren und den Rezipienten und – dank der Interaktivität – bessere Berücksichtigung von Benutzerbedürfnissen. Deshalb stellt sich die Frage, wie diese im WWW entwickelten Möglichkeiten auch für bessere Online-Wörterbücher nutzbar gemacht werden können.

Der folgende Beitrag sieht sich als Versuch, die Grundlagenarbeit für die Präsentation eines zweisprachigen Wörterbuchs in der Form eines hochwertigen Hypertextes zu leisten. Insbesondere möchte ich im folgenden die Strategie der Strukturierung und Speicherung von Wörterbuchdaten erörtern und am Beispiel eines polnisch-deutschen Wörterbuchteils vorstellen. Zuerst werde ich jedoch kurz die Stellung der zweisprachigen Computerwörterbücher innerhalb von Hypermediasystemen diskutieren und auf einige Besonderheiten der

---

<sup>1</sup> Zum Begriff Hypertext siehe J. Conklin (1987), R. E. Horn (1989), R. Kuhlen (1991), S. Freisler (1994); in Bezug auf die Lexikographie A. Storrer in diesem Band.

zweisprachigen Lexikographie im allgemeinen hinweisen, die für die Strukturierung von zweisprachigen Online-Wörterbüchern von entscheidender Bedeutung sind. Besonderes Augenmerk gilt auch der Verwendung und der Kategorisierung von Bildern, denen in dem neuen Medium endlich genügend Platz eingeräumt werden kann. Abschließend will ich noch auf einige „Mehrwerte“ eines Online-Wörterbuchs eingehen- u.a. die Vernetzungsmöglichkeiten mit anderen Web-Projekten sowie auf die Möglichkeit kontinuierlicher Erweiterung der virtuellen Wörterbücher im Internet.

## 2 Das Computerwörterbuch als Hypermediasystem

### 2.1 Die Tendenz zur Kumulation der lexikographischen Inhalte

Die logische Zusammenführung bestehender Informationssysteme – und jedes zweisprachige Wörterbuch kann man als ein Informationssystem verstehen – hat sich in den letzten Jahren im Bereich der Informationswissenschaften zu einem wichtigen Untersuchungsgegenstand entwickelt. Seitdem hypertextbasierte Netzdienste wie das World Wide Web Zugriff auf verschiedene, heterogene Informationsquellen ermöglichen, begannen die Anstrengungen um transparente föderative Oberflächen, die Informationen aus unterschiedlichen Online-Wörterbüchern integrieren. Gleichzeitig bemüht man sich um die Aufbereitung der Informationen aus bereits bestehenden Quellen, so daß bei Neuentwicklungen auf das schon gesammelte lexikalische Wissen zurückgegriffen werden kann.

Im Bereich der lexikographischen Online-Angebote sind Aktivitäten beider Art anzutreffen.<sup>2</sup> Es wurde zum Beispiel ein WWW-Projekt angekündigt, in dem eine wörterbuchübergreifende Präsentationsform für Inhalte verschiedener zweisprachiger Wörterbücher mit den Sprachen Englisch, Spanisch und Baskisch gefunden werden soll (vgl. Patric; Zubigulla 1994). Atkins (1996) schlägt eine Verbindung von zwei einsprachigen Wörterbüchern mit einem polyfunktionalen, bidirektionalen zweisprachigen Wörterbuch für das Sprachenpaar Englisch-Französisch vor – die Daten dieses virtuellen Nachschlagewerkes sollen in einer Frame-basierten Datenbank gespeichert und online zugreifbar sein.<sup>3</sup> Das einsprachige „Hypertext Webster Gateway“ ermöglicht schließlich unter einer einheitlichen Bedienungsoberfläche den Zugang zu mehreren einsprachigen Englischwörterbüchern; durch Anklicken eines in der Definition auftretenden Wortes wird ein Hyperlink zu demjenigen Backend-Wörterbuch aktiviert, aus dem die jeweilige Definition stammt, und ein neuer Nachschlagevorgang für dieses Wort ausgelöst.<sup>4</sup>

<sup>2</sup> Um den Aufbau wiederverwendbarer, in mehreren Kontexten nutzbarer lexikalischer Online-Ressourcen bemühen sich z.B. die Projekte *Deutscher Wortschatz* (30.08.1999, <http://wortschatz.uni-leipzig.de/inhalt/>) und *LEKSIS* (30.08.1999, <http://www.ids-mannheim.de/wiw/>).

<sup>3</sup> Frames sind komplexe Datenstrukturen, die stereotypische Situationen als Module des syntaktisch-semanticen Wissens über bestimmte Sprachwirklichkeitsausschnitte repräsentieren (siehe Wegner 1985, Lowe, Baker, Fillmore 1997). Atkins (1996:532ff.) schlägt vor, das Frame-Konzept – als ein linguistisches Prinzip der Modellierung eines Datenbankschemas – zur Überwindung der kognitiven Unzulänglichkeit (vgl. Kap. 3.1) des zweisprachigen Wörterbuchs zu benutzen.

<sup>4</sup> Vgl. (30.08.1999) <[http://work.ucsd.edu:5141/cgi-bin/http\\_webster](http://work.ucsd.edu:5141/cgi-bin/http_webster)>.

Während derartige Modelle im Internet noch am Anfang ihrer Entwicklung stehen, lassen sich unter lexikographischen Produkten, die für die Offline-Nutzung bestimmt sind, etliche kumulative Ansätze finden (vgl. dazu auch Petelenz 1998). In den neunziger Jahren sind mehrere Wörterbücher – meist Deutsch-Englisch/Englisch-Deutsch sowie deutsche Fremd- und Synonymwörterbücher – auf CD-ROM in Verbindung mit anderen Nachschlagewerken wie Lexikon, Enzyklopädie, Landkartenatlas u.ä. erschienen. Da sie in sich Texte, Bilder und Ton vereinigen, lassen sie sich als Hypermediasysteme bezeichnen.<sup>5</sup> Untereinander verbundene lexikographische Kompendien wie DUDENs LexiRom, Bertelsmanns INFOROM oder Infopedia von Tewi gestatten dem Benutzer den Zugriff auf Informationen aller Werke gleichzeitig, da die Daten zentral in einer Datenbank abgelegt wurden. In diesen lexikographischen Megawerken der ersten Stunde geht man oftmals davon aus, daß das Übermaß an Information nicht schaden kann. Schlägt man in der Infopedia 2.0 im deutsch-englischen Wörterbuch das Lemma *Auge* nach, weil man z.B. die Übersetzung für die deutsche Redewendung *unter vier Augen* sucht, so bekommt man ungefragt – zusammen mit dem Eintrag des Oxford-Taschenwörterbuchs Dt.-Engl. – eine Bildanimation zu sehen und einen kurzen Vortrag (auf Deutsch) über die Funktion und Aufbau dieses menschlichen Sinnesorgans zu hören. Die englische Sprachausgabe oder IPA-Transkription für *eye* oder *in privacy* wird dem Benutzer aber nicht angeboten. Infopedia 2.0. erlaubt es nicht, die Suche im Datenbestand auf bestimmte Nachschlagewerke einzuschränken. Das deutet auf die Unterschätzung der Hypertextfunktionalität hin; statt dem Benutzer die Möglichkeit zu geben, gezielt weiterführende Informationen vom System anzufordern, wird er in diesem Fall mit enzyklopädischen Angaben versorgt, die bei dieser Nachschlageoperation vollkommen redundant und von den meisten Benutzern sicherlich auch nicht erwünscht sind. Ein derartiges Überangebot an Informationen läßt sich nach Conklin (1997:40) als Beispiel für eine Ablenkung vom eigentlichen Ziel des Nachschlagens („disorientation“) und für die kognitive Überlast („cognitive overhead“) bezeichnen, die allerdings in diesem Fall nicht aus der verwirrenden Menge der einzuholenden Auskünfte resultiert, sondern aus einer Unterschätzung der Fähigkeiten des Benutzers, sich die im gegebenen Kontext interessierenden Informationen selbst aussuchen zu können.

Das Konzept der PC-Bibliothek von DUDEN in Deutschland und des LEKSYKONIA-Systems in Polen setzt ebenfalls auf die Versorgung der Benutzer mit Informationen aus verschiedenen Nachschlagewerken auf einmal, d.h. bei einem Nachschlagevorgang werden – falls der Benutzer es wünscht – mehrere Datenbestände durchsucht. Während Infopedia nicht immer die Informationen aus unterschiedlichen Werken auf eine gelungene benutzerfreundliche Art miteinander kombiniert, funktioniert die PC-Bibliothek in dieser Hinsicht wesentlich besser (vgl. dazu auch Neth; Swanson 1999:106). Auch das reichlich bebilderte dt.-poln./poln.-dt. LEKSYKONIA-Wörterbuch läßt sich um ein externes dt.-poln./poln.-dt. Fachwörterbuch der Wirtschaftssprache oder ein einsprachiges polnisches illustriertes Lexikon der technischen Termini erweitern. So besteht die Möglichkeit, den Bestand des allgemeinen Wörterbuchs gleichsam um Lemmata und zusätzliche Angaben weiterer Werke zu vergrößern. Auf diese Weise kann man je nach Bedarf in einer, in zwei oder in mehreren „Informationsbasen“, wie sie von den Autoren des LEKSYKONIA-Systems benannt wurden, gleichzeitig nachschlagen.

Die Tendenz zur Angabe heterogener Informationstypen in einem zweisprachigen Wörterbuch scheint sich auch auf die Printlexikographie auszuweiten: das Power-Englischwörterbuch von Langenscheidt (mit dem Untertitel „Zum Nachschlagen und Lernen“) ver-

<sup>5</sup> Zur Definition der Hypermedia in Bezug auf Lexikographie siehe Storrer (1998:107).

einigt in sich – außer den „reinen“ Angaben eines zweisprachigen Wörterbuchs – bisher selten verwendete Elemente. So sind da etwa farbige Bilder zu finden, die jeweils entweder einen einzelnen Begriff oder einen größeren Ausschnitt der Wirklichkeit illustrieren und das entsprechende, themengebundene Vokabular samt Beispielsätzen dazu liefern. Kleine Kästchen, die immer wieder die herkömmliche Makrostruktur unterbrechen, enthalten zusätzliche „Sprachglossen“ – wie sie im Wörterbuchvorwort genannt werden – mit Informationen zum richtigen Sprachgebrauch, zu grammatischen, pragmatischen und phonetischen Schwierigkeiten, „falschen Freunden“ und kulturbezogenen Hinweisen zur Landeskunde. Der Benutzer kann folglich zwischen den Wörterbuchartikeln und den ergänzenden Informationstafeln „hin- und herspringen“, ähnlich wie zwischen den Fenstern einer Computeranwendung.

## 2.2 Drei Generationen von zweisprachigen Hypermediawörterbüchern

Artents und Bogaerts (1991) unterscheiden zwischen drei Generationen von Hypermedia-systemen. Die erste Generation arbeitet mit festen Standardattributen, die mit Knoten, Ankern und Links verbunden sind.<sup>6</sup> Die Knoten beinhalten hauptsächlich Text, eventuell integrierte Medieninhalte dienen als ergänzende Illustration, die Benutzer können weder Knoten noch Links verändern, die Kommunikation mit anderen PC-Anwendungen, wie z.B. Textverarbeitungsprogrammen, ist nicht vorgesehen. Zu dieser Gruppe zählen viele der heute verfügbaren zweisprachigen PC-Wörterbücher auf CD-ROM aus der niedrigeren Preisklasse (vgl. Heth, Swanson 1999).

Die zweite Generation eröffnet die Möglichkeit, dem System benutzerdefinierbare Knoten und Links, Attribute und Schlüsselwörter hinzuzufügen. Benutzer können auf diese Weise – wie das beispielsweise in Langenscheidts Handwörterbuch Englisch aus der PC-Bibliothek 2.0. der Fall ist – neue Einträge, Querverweise und Schlagwörter definieren, Artikeltexte der einzelnen Knoten mit Anmerkungen versehen und mit „Leuchstiften“ kennzeichnen. Über eine DDE-Schnittstelle lassen sich Wörter aus einer Textverarbeitung heraus nachschlagen und die Knoteninhalte über die Zwischenablage ganz oder teilweise in andere Anwendungen übernehmen.<sup>7</sup> Gleichzeitig ist es möglich, ohne das System zu verlassen, auf weitere gleich strukturierte Datenbankinhalte – d.h. andere Wörterbücher und Lexika desselben Herausgebers – zuzugreifen. Das LEKSYKONIA-Wörterbuch, das „Umlaut“-Wörterbuch und der „Grundwortschatz Polnisch“ im GlobeDisc-System können z.T. mit einer Funktionalität aufwarten, die es erlaubt, diese Produkte zu der zweiten Generation zu rechnen.<sup>8</sup>

<sup>6</sup> Knoten sind die im Hypertext mit Verknüpfungen (Links) verwobenen Inhalte (multimediale Informationseinheiten: Text, Bild, Ton), Anker stellen Ausgangspunkte der Links dar, mit den Links wird aus einer Informationseinheit auf Informationen in anderen Knoten verwiesen. Die Abbildung der Verweisbeziehungen zwischen den Informationseinheiten (lexikographischen Textsegmenten) eines Printwörterbuchs mithilfe von Links beschreibt ausführlich – am Beispiel der einsprachigen Lexikographie – Kammerer (1998). Zu den Verweisbeziehungen in Wörterbüchern allgemein vgl. Wiegand 1996a.

<sup>7</sup> DDE steht für *Dynamic Data Exchange*.

<sup>8</sup> Augenfällig ist jedoch die Diskrepanz zwischen dem Einsatz fortgeschrittener Softwaretechnik beim Erstellen des Wörterbuchbrowsers und der extrem schlechten lexikographischen Qualität der Daten im Fall des GlobeDisc-Wörterbuchs sowie der ziemlich konzeptionslosen lexikographischen Ausrichtung des LEKSYKONIA-Werkes.

Die dritte Generation von Hypertextsystemen sieht eine komplexe Systemunterstützung für die Handhabung der Knoteninhalte und für die benutzergerechte Navigation vor; sie setzt zu diesem Zweck sog. „link abstract structures“ an. Diese Linkstrukturen stellen die Voraussetzungen für die Entwicklung einer wissensbasierten Systemarchitektur dar, die die Kluft zwischen dem „reader's mental model“ und dem „system's internal model“ beseitigen würde (vgl. Arents, Bogaerts 1991:133). Das Konzept beruht auf der Integration eines semantischen und pragmatischen Rahmens für Inhalt, Kontext und Struktur der Informationen auf den Ebenen der Links, der Knoten und des gesamten Systems. Die Idee der abstrakten Linkstrukturen läuft auf die Behandlung von Knoten und Links als gleichrangige Objekte hinaus – die Links seien keine einfachen Attribute der Knoten, sondern völlig eigenständige Objekte, die je nach Verwendungskontext zur Restrukturierung des Hypertextes dienen können. Diese Restrukturierung setzt eine saubere funktionale Trennung zwischen verschiedenen Typen der Semantik, eine statische Struktur der Informationen sowie eine klar definierte Navigationssemantik voraus. Die von Arents und Bogarts vorgeschlagene MMVP-Architektur (Model-Map-View-Praxis), die in einem Prototyp IKON (Intelligent Knowledge Objects Navigator) realisiert wurde, basiert auf zwei Prinzipien:<sup>9</sup>

- „Links do not express meaning by themselves, but express meaning through their navigation“,
- „Link navigation through message passing“ – eine Anfrage des Benutzers an das System wird durch den „message-passing“-Browsingmechanismus bearbeitet (vgl. Arents, Bogaerts 1991:137, 138).

Die Ideen von Arents und Bogaerts können auch als Anregung für die Computerlexikographie dienen. Ein elektronisches Wörterbuch, das den Ansprüchen der dritten Generation der Hypermediasysteme genügen sollte, müßte folglich über semantisch und pragmatisch typisierte Links zur Steuerung und Kontrolle der Systemunterstützung für unterschiedliche Benutzergruppen und Benutzungssituationen verfügen (siehe Kap. 4. 1); die Knoten der Hypertextbasis müssen zugleich mit referentiellen Links verbunden sein, die die gegenseitigen Beziehungen der Knoten untereinander ausdrücken. Im derart konstruierten Wörterbuchsystem soll auf der einen Seite ein Formalismus zur wissensbasierten Strukturierung der Informationsinhalte, der Knoten und der Links implementiert sein, auf der anderen Seite soll eine erfolgreiche Suche und eine freie Navigation (individuelles Browsing) gewährleistet sein.

Darüber hinaus sollte ein innovatives Computerwörterbuch ein offenes, erweiterbares System darstellen und – anders als eine Offline-Anwendung auf einem lokal verfügbaren Speichermedium – eine breite Kommunikation zwischen den Wörterbuchautoren und -benutzern ermöglichen (siehe den Beitrag von I. Lemberg in diesem Band). Dank dieser Kommunikation lassen sich die Knoteninhalte korrigieren, ergänzen und erweitern.

<sup>9</sup> IKON speichert Informationen in Form von „information units“; sie sind klein, exakt definiert und dienen der Dekontextualisierung von Informationfragmenten, so daß die korrespondierenden Informationseinheiten sinnvoll miteinander verbunden und manipuliert (wiederverwendet und rekombiniert) werden können. Die „information units“ werden im semantischen Hyperindex charakterisiert. Zur semantischen Hyperindexing-Technik siehe Arents, Bogaerts (1993).

### 2.3 Neue Perspektiven für zweisprachige Computerwörterbücher

Völlig neu entwickelte elektronische zweisprachige Wörterbücher, sowohl auf CD-ROM als auch im WWW, fehlen noch (vgl. dazu auch Feldweg 1997). Der Grund dafür ist auf der einen Seite die Zurückhaltung der Verlage, wenn es darum geht, viel Geld in die neuen Medien zu investieren, und sie ist auch verständlich, solange die elektronischen Ressourcen auf Speichermedien dem zügellosen Raubkopieren ausgesetzt sind und solange keine Möglichkeiten der Abrechnung für virtuelle WWW-Angebote bestehen. Auf der anderen Seite fehlt wohl einigen Verlagen noch das Know-how, um die Informationen medienneutral aufzubereiten und in einer zentralen Datenbank aufzubewahren. Es ist vor allem nicht möglich, die linear abgefaßten Wörterbuchartikel der Printlexikographie maschinell in eine mehrdimensionale, hierarchisch organisierte Hypertextstruktur zu überführen, da die paradigmatischen Querverbindungen auf der Ebene der Semantik nur aufgrund intellektueller Analyse abbildbar sind. Ebenso lassen sich die Textverdichtungen in den Wörterbuchartikeln nicht völlig automatisch auflösen. Die CD-ROM-Publikationen auf der Basis herkömmlicher Printwörterbücher – auch wenn sie recht gelungen sind, wie z.B. die elektronische Version des OED oder des Petit Robert – orientieren sich immer noch am linearen Textparadigma. Das 1996 von Heidecke skizzierte Redaktionssystem bringt hinsichtlich der Hypertextualisierung von zweisprachigen Wörterbüchern wenige Innovationen, da es hauptsächlich auf Erstellung von Druckwerken ausgerichtet wurde.

Eine endgültige Technologie der Erstellung und Wartung von großen, komplexen elektronischen Dokumenten und zur Erschließung der Papierwelt für Computer hat sich noch nicht etabliert. Man kann die Bemühungen um die innovativen Publikations- und Kommunikationsmethoden mit dem vagen Begriff Wissensmanagement (Knowledge Management) bezeichnen, der in der Informationswissenschaft u.a. als Oberbegriff für Document Engineering und intelligentes Information Retrieval funktioniert. Document Engineering verbindet die Methoden der angewandten Text- und Computerlinguistik und der objektorientierten Softwareentwicklung; alle Arbeitsschritte werden zu einem ingenieurmäßigen, linguistisch fundierten Vorgehen gebündelt (vgl. Kessler; Freisler 1995, Williams 1997). Die eigentlichen Probleme des Document Engineerings liegen nicht nur auf der technischen Ebene der Textproduktion und -präsentation, sondern vorwiegend auf der semantischen Ebene: Die bestehenden Informationen müssen analysiert und in Bezug auf ihre Beschaffenheit, Bedeutung und Funktion charakterisiert werden. Für eine Computeranwendung, die die Informationen handhaben soll (d.h. dynamisch zueinander in Verbindung setzen soll), muß außer der Information selbst auch ihre Semantik repräsentiert werden. Document Engineering entwickelt Techniken, mit denen der semantische Gehalt von Informationen und die Semantik der Relationen zwischen einzelnen Informationen explizit und deklarativ modelliert und mittels geeigneter Schlußfolgerungsverfahren interpretiert werden. Benutzermodelle sollen die Erstellung von „maßgeschneiderten“ Artikeln möglich machen. Im folgenden wird eine mögliche Organisation der Informationen eines zweisprachigen Wörterbuchs in einer Hypertextbasis skizziert, die – um eine Navigationskomponente ergänzt – den Benutzern die lexikographischen Daten auf eine innovative Art und Weise präsentieren könnte. Auf die Aspekte des Retrievals von Wörterbuchdaten wird aus Platzgründen nicht eingegangen.

Zunächst möchte ich jedoch die spezifischen Probleme des zweisprachigen Wörterbuchs überblickartig beleuchten, wohl wissend, daß sie den mit der zweisprachigen Metalexikographie vertrauten Leserinnen und Lesern hinlänglich bekannt sein dürften. Diese Probleme spielen aber eine entscheidende Rolle für das lexikographische Informationsdesign –

und umgekehrt: ein gelungenes Informationsdesign kann viele dieser Probleme lösen –, sie können also im vorliegenden Beitrag nicht unerwähnt bleiben.

### 3 Zu den Besonderheiten der zweisprachigen Lexikographie

Um sich der Anordnung der lexikographischen Daten in einer Hypertextbasis zu widmen, muß man zuvor kurz auf die Grundbausteine der zweisprachigen Printwörterbücher zu sprechen kommen. Die Werke lassen sich nach Skopus, Funktion, Direktion und dem primären linguistischen Ordnungsprinzip des Datenmaterials charakterisieren (vgl. Hausmann, Werner 1991).

Nach dem Skopus sind ein- und zweiteilige Werke zu unterscheiden, z.B. ein dt.-poln. und poln.-dt. Wörterbuch (biskopal) vs. ein poln.-dt. Wörterbuch ohne Pendant für die Gegenrichtung (monoskopal). Ein anderes Kriterium ist das viel diskutierte Aktiv-Passiv-Prinzip, d.h. die Ausrichtung des Werkes auf die „aktive“ Hinübersetzung und Textproduktion in der ZS<sup>10</sup> bzw. „passive“ Herübersetzung und Textrezeption in der AS (folglich jeweils monofunktional) oder für beide diese Funktionen gleichzeitig (bifunktional).<sup>11</sup> Als bidirektional läßt sich ein Wörterbuch bezeichnen, wenn es die Bedürfnisse beider Benutzergruppen, d.h. der Sprecher der AS und der Sprecher der ZS, gleichermaßen gut befriedigt. Monodirektional ist hingegen ein solches, das grundsätzlich nur eine der beiden Benutzergruppen als Zielpublikum ins Auge faßt.

Weitere Kriterien stellen die Paradigmatik und die Syntagmatik dar (vgl. Hausmann 1991). Sie determinieren die Materialanordnung in der Makro- und Mikrostruktur des Wörterbuchs. Während sich die paradigmatischen zweisprachigen Werke (wie Bildwörterbuch, thematischer Thesaurus oder Wörterbuch der kontrastiven Synonymik) der Onomasiologie und der lexikalischen Austauschbarkeit widmen, beschreiben syntagmatische Wörterbücher (Idiomatik-, Valenz-, Kontextwörterbuch etc.) die kontextuellen und grammatischen Abhängigkeiten der Lexik, ihre Gebundenheit in phraseologischen Einheiten. Im Zentrum der Aufmerksamkeit stehen dabei die Translate, d.h. die syntagmatischen Übersetzungseinheiten. Diese müssen nicht identisch sein mit den aus einer innersprachlichen Analyse hervorgegangenen Phrasemen, das betrifft Kollokationen, Routineformeln und „stehende“ Vergleiche. Ein allgemeines zweisprachiges Wörterbuch versucht sowohl die paradigmatische als auch die syntagmatische Betrachtungsweise in sich zu vereinen, mit dem Resultat, daß an vielen Stellen große Abstriche in der lexikographischen Beschreibung auf allen Ebenen gemacht werden müssen. Die mangelhafte Behandlung der Phraseme im zweisprachigen Wörterbuch wird in Kap. 4. 4 näher besprochen.

Pädagogische Überlegungen legen den Gedanken nahe, zwischen den Bedürfnissen der Laienbenutzer (die weder mit der Zielsprache gut vertraut sind noch sich mit den lexikographischen Konventionen auskennen) und denen der kundigen Benutzer (die sowohl die Zielsprache gut beherrschen als auch über die linguistische Vorbildung und die nötige Übung in der Wörterbuchbenutzung verfügen) zu unterscheiden. Diese Differenzierung soll jedoch an dieser Stelle lediglich signalisiert und nicht weiter diskutiert werden.

<sup>10</sup> ZS = Zielsprache, AS = Ausgangssprache.

<sup>11</sup> Die Zusammenfassung der Diskussion um das Aktiv-Passiv-Prinzip ist bei Tarp (1995) zu finden.

## 4 Der typenübergreifende Wörterbuchserver: Zur allgemeinen Vorgehensweise

### 4.1 Das Schema des Wörterbuchsystems

Der erste Schritt auf dem Weg zu einem typenübergreifenden Wörterbuchserver ist die Analyse der lexikographischen Textsegmente verschiedener Typen von zweisprachigen Wörterbüchern für Deutsch-Polnisch.<sup>12</sup> Die dabei isolierten mikrostrukturellen Textsegmente werden anschließend im Hinblick auf die Bifunktionalisierung und/oder Bidirektionalisierung um entsprechende Angabetypen ergänzt. Die Bifunktionalisierung und Bidirektionalisierung führt z.T. zur Vereinheitlichung, Restrukturierung und Expandierung der in den Printwörterbüchern vorgesehenen Informationstypen.

Die eingescannten Artikel verschiedener Wörterbücher für Deutsch-Polnisch bilden ein Korpus mit den lexikographischen Daten. Diese Daten werden (korrigiert und ergänzt) in die Lexikoneinträge der Hypertextbasis übernommen. In die so entstandene Gesamtstruktur fließen noch zusätzliche Informationstypen mit ein, die in der zweisprachigen Printlexikographie entweder ansatzweise (umfangreiche Textbeispiele und Bilder) oder überhaupt nicht (Töne) vertreten sind. Die ursprüngliche Erfassung erfolgt im Textverarbeitungsprogramm (MS Word), den einzelnen Datentypen werden Absatz- und Zeichenformatvorlagen zugeordnet. Anschließend ist die Konvertierung der Daten in eine XML-Datei vorgesehen – die formbezogenen Informationen aus den Word-Stylesheets sollen durch semantische Tags ersetzt werden<sup>13</sup>. Schließlich findet die Übernahme der lexikographischen Daten in die wörterbuchtypen-übergreifende Struktur statt. Diese Struktur, die sich als „Export-Standard“ im Sinne von Bläsi et al. (1994) bezeichnen läßt, muß auf der Speicherungsebene des Wörterbuchservers mithilfe spezieller Software – SchemaText – modelliert werden (vgl. Abbildung 3).<sup>14</sup> Zuvor wird jedoch der Entwurf einer Hypertextbasis in Angriff genommen: Das konzeptuelle, implementierungsunabhängige Schema der Hypertextbasis setzt sich auf der einen Seite aus den typisierten Knoten zusammen, die die ‚kohäsiv geschlossenen informationellen Einheiten‘ hierarchisch ordnen<sup>15</sup>, und auf der anderen Seite aus den typisierten Links, die die Semantik der Relationen zwischen den Knotentypen festlegen. Die typisierten Links, die die einzelnen Knotentypen verbinden, stellen ‚informationelle Funktionen‘ (vgl. Kuhlen 1991:89) dar, sie drücken entweder die hierarchischen Relationen auf der syntagmatischen und paradigmatischen Ebene des Lexikons aus – die hierarchischen Links – oder sie bezeichnen die Beziehungen der Einheiten zueinander im bestimmten situativen Kontext – die pragmatischen Links.

Die höchste Hierarchiestufe der Hypertextbasis unterscheidet zwischen sprachlichen und metasprachlichen Einheiten. Die sprachlichen Einheiten bilden die Lemmata des Wörterbuchsystems, alle an das jeweilige Lemma adressierten Angaben stellen metasprachliche Einheiten dar. Alle Einheiten zusammen sind in eine semantisch und pragmatisch vordefinierte Struktur organisiert, die modular aufgebaut ist, wie die Abbildung 1 zeigt. Die Unterscheidung auf der Ebene der sprachlichen Einheiten führt zu dem Mehrwortlexem-Modul

<sup>12</sup> Analysiert wurden folgende Wörterbücher: DBWDP, GWPDP/DP, HWBPD/DP, ISNP, SWBPD, STJN und TWBP.

<sup>13</sup> XML steht für EXtensible Markup Language.

<sup>14</sup> (30.8.1999) <http://www.schema.de/sitehtml/site-d/schemat2.htm>.

<sup>15</sup> Die zusammenhängenden informationellen Einheiten (‚information units‘) können größere Einheiten (sog. ‚information chunks‘ oder ‚informations blocks‘, vgl. dazu Horn 1989:40ff.) bilden.



(aus Idiomen, Kollokationen und Sprichwörtern) und Einwortlexem-Modul (aus Verben, Nomen, Adjektiven/Adverbien und einer „Restkategorie“).

Die Unterscheidung auf der Ebene der metasprachlichen Einheiten führt zu dem Monosem-Modul und Polysem-Modul, wobei die Templates beider Module (in der Abb. 1 der Hauptknotentyp ‚Bedeutung‘) gleich strukturiert sind; das Schrift- und Lautform-Modul befindet sich auf der selben Hierarchiestufe. Auf der nächstniedrigeren Ebene befinden sich einzelne Beschreibungssegmente der ZS-Äquivalente für die jeweilige Lemmabedeutung, die je nach Wortart unterschiedliche Slots mit Angaben zur Herkunft, Grammatik, Semantik und Pragmatik sowie Beispielen und Bildern aufweisen können.<sup>16</sup> Die Knoteninhalte mehrerer Slots können identisch sein (z.B. eine Paraphrase verschiedener Redewendungen mit gleicher Bedeutung), was bedeutet, daß derselbe Knoten in mehreren Kontexten verwendbar sein kann.

Die herkömmlichen zweisprachigen Printwörterbücher ermöglichen den Zugriff auf die Lexik in der Regel nur „auf der Form-, nicht auf der Bedeutungsebene“ (Martin 1994:17.; vgl. auch Storrer 1998:116). Die hierarchischen Links stellen also Verknüpfungen hauptsächlich nach semantischen Kriterien her und eröffnen dadurch neue Zugriffswege zum lexikalischen Material. Die pragmatischen Links haben hingegen die Aufgabe, die Generierung der benutzergerechten Artikel zu steuern.

Auf der nächstniedrigeren Ebene kommt der Parameter Adressatenkreis richtig ins Spiel – die metasprachlichen Angaben sind nämlich in der Hypertextbasis grundsätzlich doppelt vorhanden: der eine Angabetyp ist für die deutschen, der andere für die polnischen Benutzer bestimmt.

#### 4.2 Die Typisierung als Grundlage der wissensbasierten Relationierung von Hypertexteinheiten

Die Anpassung der präsentierten Informationen an die Bedürfnisse der Benutzer geschieht in einem Interaktionsprozeß. Am Anfang einer Sitzung mit dem Wörterbuchsystem (d.h. auf der ersten HTML-Seite), wird der Benutzer gefragt, welche Muttersprache er spricht und an dieser Stelle somit die Entscheidung gefällt, welcher Teil des Schemas ausmaskiert bleibt und welche Expander für die Generierung der HTML-Seiten aus der Datenbasis zum Einsatz kommen. (Das Layout für die Zielformate wird in SchemaText in den sog. Expandern festgelegt.<sup>17</sup> Expander sind kleine Scheme-Programme, die das Layout für das Textschema, für Knotentypen, für Knoten, für Linktypen und für Links generieren. Sie können rekursiv weitere Expander aufrufen).<sup>18</sup>

Der „SchemaEditor“ dient in SchemaText zur Modellierung von Informationstypen und deren Relationen. Dies erfolgt auf einer graphischen Oberfläche (vgl. Abb. 1). In einem Schema werden Typen von Textobjekten und Verweisbeziehungen bestimmt (Knotentypen: Idiom, Kollokation, Nomen; Äquivalent, Genusangabe usw. und Linktypen ‚Bezieht-sich-auf‘, ‚Ist-Teil-von‘, ‚Synonym-zu‘ etc.) und damit Operationen auf globalen und lokalen Textstrukturen ermöglicht. Komplexe heterogene Vernetzungsstrukturen können – zwecks

<sup>16</sup> Die Knoteninhalte mit Bildern und Tönen enthalten in der Tat lediglich einen Verweis auf die entsprechende Datei im Ressourcenpool mit binären Daten. Dieser Pool stellt ein separates Modul der Hypermediabasis dar (s. Kap. 4.4.2).

<sup>17</sup> Wie das Navigieren auf der Oberfläche des Wörterbuchsystems im WWW-Browser erfolgt, wird im Petelenz (2000:216ff.) skizziert.

<sup>18</sup> Scheme ist ein weitverbreiteter, international standardisierter LISP-Dialekt.

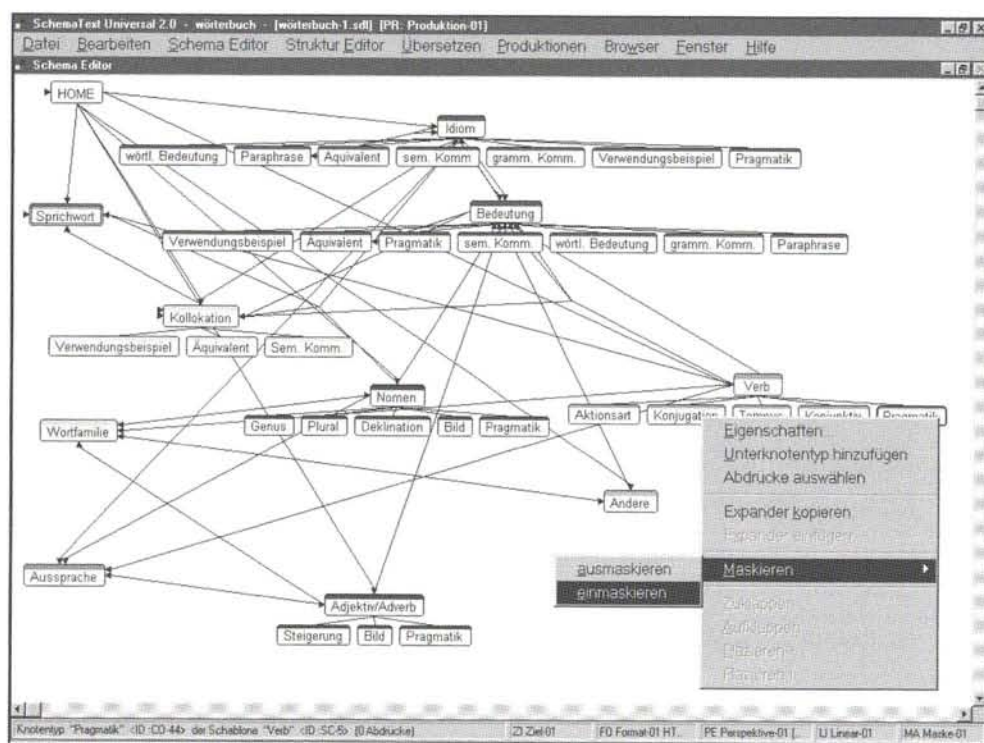


Abb. 1

Mehrfachverwendung und Rekombination der Inhalte – über sog. Multilinks verwaltet werden. Ein Nomen (als Knotentyp) kann z.B. gleichzeitig einen Teil von einer Kollokation, einem Idiom oder einem Sprichwort bilden.

Schema der Hypertextbasis			
Schema 1 WB Poln.-Dt.	Schema 2 WB Dt.-Poln.	Schema 3 Grammatikhilfe	Ressourcen-Pool Bild- und Tondateien

Abb. 2

Die Hypertextbasis setzt sich im SchemaText auf der abstrakten Strukturebene aus Schemata zusammen, diese wiederum aus den hierarchisch organisierten Templates (Schablonen). Jede Schablone hat in der Datenbasis eine physikalische Entsprechung: einen Abdruck, der aus Knoten und Links besteht. Ein Abdruck muß mit dem definierten Template konform sein. (Die Templates lassen sich allerdings auch nachträglich problemlos verändern). Die konkreten Textstrukturen des Wörterbuchs kann man im sog. „Struktur-Editor“ festlegen. Die im Schema gespeicherte Struktur ist von den Instanzen selbst unabhängig, d.h. die Layouteigenschaften der Knoten und Links werden getrennt vom Schema

festgelegt. Die Vererbung bedeutet dann z.B. die Fähigkeit, Änderungen, die am Knotentyp durchgeführt wurden, auf alle Knoten dieses Typs zu übertragen.

Der Skopusbezug wird durch die Teilung der Hypertextbasis in zwei Wörterbuch-schemata bestimmt: ein Schema enthält die deutschen Lemmata mit polnischen Äquivalenten, das andere die polnischen Lemmata mit deutschen Äquivalenten. Die Benutzerprofile (1–4 im Kap. 4.3) sind für jedes Schema einzeln zu implimentieren. Das dritte Schema enthält die Grammatikhilfe (vgl. Abbildung 5), ein für beide Benutzergruppen bestimmtes Handbuch, auf deren Einträge (Kapitel und Unterkapitel) kontextsensitiv aus dem Wörterbuch (Schema 1 und Schema 2 in der Abbildung 1) verwiesen wird. Das Handbuch läßt sich jedoch auch als ein inhärentes Hypertextdokument lesen, ähnlich wie etwa eine Onlinehilfe von üblichen Computerprogrammen. Die einmal erstellte Struktur des Textschemas für die Grammatikhilfe ist in beiden Handbüchern, d.h. in dem für die polnischen und in dem für die deutschen Benutzer, gleich. Der benutzergerechte Zugang wird über die sog. Perspektive gesteuert, also eine Dimension des Schemas, die in diesem Fall zwei Textstrukturkopien der aktuellen Textstruktur enthält, und zwar für beide Sprachen. In diesen zwei Sprachperspektiven sind die Knoteninhalte des Schemas verschieden, die Textstruktur und das verwendete Expanderset jedoch gleich.

#### 4.3 Die expandierten Informationstypen für die Wörterbuchinstanzen

Das System hat Sorge dafür zu tragen, daß der Benutzer nicht mit für ihn redundanten Informationen vom Wörterbuchsystem übersorgt wird. Die Verknüpfung der informationellen Einheiten mit semantisch typisierten Links soll in der Interaktion mit dem Benutzer die dynamische Anpassung des Systems an die jeweilige Wörterbuchbenutzungssituation bewirken. Demnach sind Benutzerprofile zu definieren, sie basieren auf dem in der Künstlichen Intelligenz benutzten Konzept des Benutzermodells. Grundsätzlich sind vier Benutzerprofile denkbar:

1. Benutzerprofil: AS Deutsch ‚aktiv‘ (Hinübersetzung)
2. Benutzerprofil: AS Deutsch ‚passiv‘ (Herübersetzung)
3. Benutzerprofil: AS Polnisch ‚aktiv‘ (Hinübersetzung)
4. Benutzerprofil: AS Polnisch ‚passiv‘ (Herübersetzung)

Kuhlen (1991:330, 331) bezeichnet das Benutzermodell als ein wissensbasiertes Verfahren mit dem Ziel, „Informationssysteme mit dem Wissen über individuell oder stereotypisch definierte Nutzer oder Nutzerklassen zu versorgen, damit Systeme selektiv auf die Bedürfnisse der Nutzer beim Antwortverhalten oder beim Problemlösen eingehen können. [...] Benutzermodelle sind bei Hypertext notwendig, um die mit dem dialogischen Prinzip verknüpften Erwartungen einlösen zu können“. „Dialogisches Prinzip“ sei weiterhin eine „Konzeption von Hypertext, nach der System und Nutzer wechselnde Initiative im Dialog haben sollten, d.h. weder soll die Nutzung über vordefinierte Pfade vom System gesteuert werden, noch soll dem Nutzer, gemäß dem Prinzip der direkten Manipulation, die Zuständigkeit für das Navigieren alleine überlassen werden. Vielmehr soll das System in der Lage sein, dem Benutzer situationsgemäße Nutzungsangebote, z.B. über Pfade, zu machen, denen er folgen kann (oder auch nicht). Voraussetzung für die Realisierung dialogischer Prinzipien sind reiche, wissensbasierte Strukturierung der Hypertexteinheiten und differenzierte Verknüpfungsstrukturen [...]. Das Modellieren der Benutzermodelle kann folglich als

„realistische Zwischenstufe“ zwischen „dummen“ Volltextsystemen und „intelligenten“ Wissensbanken“ angesehen werden (vgl. Kühlen 1991:61).

Die Modellierung von Benutzerprofilen soll mittels des Maskierungsmechanismus geschehen: Mit Masken ist es möglich, durch Eingabe von Regeln oder durch Selektieren auf der graphischen Oberfläche (vgl. das kontextsensitive Menü in der Abb. 1) Varianten eines Dokumentes als Untermenü des Gesamtdokumentes zu definieren. Die Variantenbildung erstreckt sich nicht nur auf die grobe Strukturebene, sondern auch auf die feinkörnige Inhaltsebene.<sup>19</sup> Das Aus- und Einmaskieren von Link- und Knotentypen in bestimmten Verwendungssituationen soll gemäß erstellten Benutzerprofilen erfolgen.

#### 4.4 Die Behandlung der Phraseme

Auf der Ebene der Mikrostruktur in den Printwörterbüchern fällt auf, daß Phraseme – abgesehen von Ausnahmen – als infralemmatische Adressen in verdichteter Form in den Artikeln für Einwortlexeme auftreten. Dieses Verfahren wird den Mehrwortlexemen (MWL) nicht gerecht, da wichtige Angaben (die an MWL adressiert sein müßten) meist im allgemeinen zweisprachigen Wörterbuch fehlen, – zum einen, weil der verfügbare Platz nicht ausreicht, zum anderen, weil das grammatische Verhalten oder die stilistische Markierung des MWLs mit denen des Eingangslemma nicht übereinstimmen und die Wörterbuchautoren die Verwirrung der Benutzer vermeiden wollen. Die Phraseme und andere MWL sind deshalb als eigenständige Lemmata – und nicht als Sublemmata bzw. infralemmatische Adressen der Einwortlexem-Lemmata – in die Datenbasis des neuen Wörterbuchsystems aufzunehmen (vgl. Abbildung 1).<sup>20</sup> Das Phrasem-Modul eröffnet die Möglichkeit, die Mehrwortlexeme als eigenständige Lemmata ausführlich zu bearbeiten (vgl. Abbildung 3).

Das Nachschlagen der Phraseme soll von der Retrieval-Komponente des Wörterbuchsystems erleichtert werden. Hierzu könnte auf Elemente der bereits bestehenden computerlexikographischen Tools, wie etwa des von Xerox patentierten experimentellen LOCOLEX-Nachschlagesystems, zurückgegriffen werden. LOCOLEX leistet nicht nur morphologische Analyse, d.h. die Rückführung auf die Grundform und die Kompositasegmentierung, sondern löst darüber hinaus das Problem der Erkennung von „MWLs ausgehend von einzelnen Textwörtern, die Bestandteil des MWL sind“ sowie der „Erkennung von getrennten Prefixverben und anderen komplexen Ausdrücken“ (vgl. Thielen, Breidt, Feldweg 1998). Dafür wurde die sog. IDAREX-Grammatik (Breidt, Segond, Valetto 1996) benutzt<sup>21</sup>:

<sup>19</sup> Textteile können auch für bestimmte Medien bzw. Zielformate ein- oder ausgeschaltet werden.

<sup>20</sup> „Die in Nischen und Nestern gruppierten Lemmata heißen Sublemmata. Die an sie adressierten Angaben sind sublemmatisch adressiert. Sublemmata gehören zur Makrostruktur. Das erste Lemma einer Gruppe von Sublemmata heißt Eingangslemma (Nischeneingangslemma/Nesteingangslemma). Es ist also streng zwischen Sublemma und infralemmatischen Adressen zu unterscheiden. Während die Sublemmata der Makrostruktur angehören – sie können über einen alphabetischen Suchpfad gefunden werden (der freilich im Falle der Nester recht kurvenreich verlaufen kann) –, kommt infralemmatischen Adressen nur mikrostruktureller Rang zu“ (Hausmann, Werner 1991:2747).

<sup>21</sup> IDAREX steht für *IDioms As Regular Expressions*. Über die Einschränkungen bei der Verwendung der IDAREX-Grammatik für die Beschreibung von MWL in einem englisch-polnischen Wörterbuch berichtet Piotrowski (1999). Diese Einschränkungen resultieren hauptsächlich aus der starken Orientierung an den MWL-Ansatzformen des „Oxford-Hachette English and French Dictionary“ (vgl. Piotrowski 1999:118–119).

„Zweisprachige Wörterbücher enthalten [...] Mehrwortlexeme (MWL, z.B. *sein Teil abbekommen, in Erfahrung bringen*), feste Ausdrücke (z.B. *Eile mit Weile*) und grammatische Kollokationen (z.B. *sich in etw or zu etw verwandeln*), die für die Rezeption durchaus relevant sind und deshalb in ein maschinell verarbeitbares Format überführt werden müssen. [...] Zur Formalisierung von Mehrwortlexemen, festen Ausdrücken und grammatischen Kollokationen werden diese zuerst in ihre kanonische Form umgeformt, wofür wir die in den herkömmlichen Wörterbüchern gebrauchte Notation etwas erweitert haben. Die kanonische Form eines Mehrwortausdrucks enthält alle lexikalisch fixierten Komponenten in der Form, in der sie vorkommen müssen, bzw. bei morphologisch flexiblen Bestandteilen in einer ‚neutralen‘ Form (Verb im Infinitiv Präsens, Nomen im Singular etc.). Verbargumente werden mit Metavariablen (etw, jmd, NPakk) ausgedruckt. Außerdem wird mit ‚°‘ markiert, ob ein Wort morphologisch variiert werden kann. Lexikalische Varianten, die im Wörterbuch durch ‚/‘ getrennt sind, werden mit ‚^‘ geklammert, wenn sie mehr als ein Wort enthalten, um den Skopus deutlich zu machen“. (Thielen, Breidt, Feldweg 1998:187–188; weiterführende Details zu IDAREX kann man auch bei Breidt, Segond, Valetto [1996:21ff.] nachlesen.)

Im Rahmen des EU-Projektes STEEL (Developing Specialized Translation/Foreign Language Understanding Tools for Eastern European Languages) wurden die ersten Erfahrungen mit der Anwendung von IDAREX-Formalismen zur Beschreibung der Phraseme in einem Wörterbuch mit der ZS-Sprache Polnisch gemacht (vgl. Piotrowski 1999). Inwieweit sich diese Technik für die Beschreibung von polnischen MWL eignet, muß in Zukunft noch untersucht werden.

## 5 Zum Problem der Standardisierung

Um lexikalische Daten in standardisierter Form als Hypertext zu organisieren, würde es ideal sein, auf Standards der Printlexikographie zurückgreifen zu können. Für Wörterbücher gibt es jedoch keine typen- und länderübergreifenden Standards, sondern spezifische lexikographische Konventionen, die einerseits aus der Tradition der Wörterbucherstellung in einem gegebenen Land resultieren und andererseits von den spezifischen Erfordernissen eines Wörterbuchtyps abhängen. Die meisten Konventionen betreffen die Textverdichtungsmechanismen, die für das Gebot der Platzersparnis in Printwörterbüchern von entscheidender Bedeutung sind. Sie sollen Wörterbücher handhabbar und benutzerfreundlich machen. Aber tun sie das immer? Mit Sicherheit ja – weil ohne Textverdichtungskonventionen überhaupt kein Wörterbuch möglich ist – aber zugleich nur sehr eingeschränkt. Viele Benutzer haben erhebliche Schwierigkeiten mit den Abkürzungen, der Nischentechnik und der Tildierung zurecht zu kommen. Das Alphabet als primäres Ordnungsprinzip im Printwörterbuch ist ebenfalls problematisch (vgl. auch Kap. 4.2.). Die Hypertextualisierung der zweisprachigen Wörterbücher wurde u.a. durch die Medienpädagogik und die Sprachdidaktik angeregt, aus der Überzeugung heraus, daß paradigmatische Navigationsstrukturen viel besser als die alphabetische Anordnung der Lemmata der Lernfunktion der zweisprachigen Wörterbücher gerecht werden. Die Konvention der alphabetischen Materialanordnung kann ein Beispiel für die psycholinguistische Unzulänglichkeit des zweisprachigen Wörterbuchs sein.<sup>22</sup> Die nach linguistischen Prinzipien gebildeten Nester wichen in

<sup>22</sup> Das alphabetische zweisprachige Wörterbuch ist deswegen psycholinguistisch unzulänglich, da es das Bild von der sprachlichen Wirklichkeit verfälscht: Die zielsprachlichen Äquivalente eines Lemma können vom Benutzer erst dann richtig erfaßt und behalten werden, wenn sie nicht nur im

zweisprachigen Wörterbüchern zunehmend den nach stupider Alphabetordnung gereihten Nischen, da nur diese Methode der Materialanordnung sich in der Printlexikographie als praxistauglich erwiesen hat. Innerhalb von „graphischen Nischen“ (Hausmann, Werner 1991:2747) sind die Lemmata striktalphabetisch sortiert, ohne Rücksicht auf Semantik und Wortbildung. Für die Angabe der Antonyme, Hypo- bzw. Hyperonyme oder der Wortfamilie des Lemma ist in der Regel kein Platz. Ebenfalls gibt es keine Standards für die Behandlung von Phrasemen – unter welchem Stichwort ein Mehrwortlexem lemmatisiert wird, dies wird meist der individuellen Entscheidung des Autors überlassen.

Bei Experimenten mit der Nutzbarmachung von lexikographischen Daten aus traditionellen Wörterbüchern für sprachverarbeitende Systeme stellte man dennoch Überlegungen über abstrakte lexikographische Standards an. Solche Standards sind für die automatische Erkennung von groben syntaktischen Strukturen in Wörterbuchtexten sehr nützlich. Bläsi et al. (1994) unterscheiden zwischen internen Standards und Export-Standards für einzelne Artikel und ganze Wörterbücher. Die internen Standards stellen die abstrakte Artikelhierarchie (die je nach Wortart unterschiedlich ausgeprägt ist) dar, die Export-Standards bestimmen hingegen die Datenbankstruktur, in die die Wörterbuchinhalte überführt werden sollen. Die „Rohdaten“ eines Wörterbuchs müssen folglich anhand von internen Standards validiert und anschließend an die Export-Standards angepaßt werden – dann lassen sie sich in einer Wörterbuchdatenbank<sup>23</sup> ablegen (vgl. Bläsi et al. 1994). In unserem Fall geht es jedoch nicht um das Parsen von einzelnen Wörterbüchern, die man komplett in eine Datenbankanwendung übernehmen möchte. Vielmehr geht es um eine exemplarische Analyse der Informationstypen der einzelnen Werke im Hinblick auf Skopus, Funktion und Direktion. Um eine Vorstellung davon zu bekommen, welche potentiellen Textsegmente ein biskopales, bifunktionales und bidirektionales Wörterbuch haben kann und haben soll, gilt es die Intention der Lexikographen nachzuvollziehen, mit der sie bestimmte funktionale Textsegmente in die Artikel ihrer Wörterbücher aufnahmen. Bei dieser Analyse leistete die metalexikographische Methode der exhaustiven funktional-positionalen Segmentierung von E.H. Wiegand (vgl. Wiegand 1986 und Wiegand 1996b) eine große Hilfe.

Da alle analysierten dt.-poln./poln.-dt. zweisprachigen Wörterbücher für menschliche Benutzer geschrieben wurden, genügen sie nur sehr bedingt einem internen metalexikographischen Standard, da sie in ihre Artikel oft von der schablonenhaften „Idealhierarchie“ mehr oder minder abweichen. Das spielt aber keine große Rolle, da sie, wie schon gesagt – aus unterschiedlichen Gründen – nicht für das Parsen in Frage kommen.<sup>24</sup>

---

Bedeutungsfeld, sondern auch im Bezeichnungsfeld gesehen werden. Die Formebene des Alphabets läßt nur eine – nämlich die semasiologische – Betrachtungsweise des sprachlichen Zeichens zu, sie geht vom Wortkörper aus und gibt Äquivalente für die verschiedenen homonymen Bedeutungen an. Die Onomasiologie geht hingegen nicht von der Form, sondern vom Begriff aus; sie könnte folglich im zweisprachigen Wörterbuch die Verbindungen von einem Begriff zu unterschiedlichen Wortkörpern herstellen und diese innerhalb eines Bedeutungs-, Sach- oder Situationsfeldes je nach Bedarf verdeutlichen. Nur das gleich berichtigte Nebeneinander beider Betrachtungsweisen der Lexik wird den kognitiven und kommunikativen Aspekten des Fremdsprachenerwerbs und -gebrauchs gerecht.

<sup>23</sup> Als Wörterbuchdatenbank ist in diesem Kontext ein geparstes Wörterbuch zu verstehen – vgl. Bläsi, Koch (1991); Hauser, Storrer (1996).

<sup>24</sup> Die wichtigsten Gründe sind: 1. Das Fehlen maschinenlesbarer Vorlagen. 2. Die komplizierte urheberrechtliche Lage. 3. Die Notwendigkeit, die Wörterbuchdaten nicht nur strukturell, sondern vor allem sachlich in Hinblick auf Korrektheit, Aktualität und Vollständigkeit zu überprüfen. Daher würde sich der Aufwand, eine Mikrostrukturgrammatik für die Wörterbücher zu schreiben, nicht lohnen.

## 6 Die expandierten lexikographischen Datentypen in der Hypertextbasis

### 6.1 Verschiedene Angabetypen für verschiedene Wörterbuchbenutzungssituationen

Die Analyse der Wörterbuchtexte erlaubt Rückschlüsse auf die lexikographischen Datentypen, die in der Datenbasis des Online-Wörterbuchs ihren Platz finden sollen, um Bedürfnisse beider Benutzergruppen bei allen potentiellen Nachschlageoperationen zu befriedigen. Im polnisch-deutschen Teil sind somit folgende Angaben zu jedem Lemma vorgesehen:

Für die polnischen Benutzer, sowohl im Wort- als auch im Phrasem-Modul:

1. eine Paraphrase der jeweiligen Bedeutung des AS-Lexems (auf Polnisch)<sup>25</sup>
2. ein oder mehrere ZS-Äquivalente als Vorschläge für die Hinübersetzung
3. ein Beispiel der Verwendung des AS-Lexems im Kontext
4. eine Übersetzung des Beispiels
5. Informationen über Lautform (Rechtschreibung) und Aussprache (als Sprachausgabe und IPA-Angabe) zu jedem ZS-Äquivalent
6. ggf. eine Herkunftsangabe (auf Polnisch)
7. Bei sog. Nulläquivalenz eine ZS-Paraphrase der AS-Bedeutung (deutlich als solche gekennzeichnet) – wenn möglich – als ein einsetzbares Translat
8. ggf. semantisch-pragmatische Kommentare zur konnotativen Verwendung des jeweiligen ZS-Äquivalentes (auf Polnisch)
9. grammatische Konstruktionshinweise bezüglich der Valenz und Rektion des AS-Lemma (auf Polnisch)
10. ggf. ein kontextsensitiver Verweis auf ein entsprechendes Kapitel der polnischen Grammatikhilfe (d.h. zur kleinen Grammatik des Deutschen auf Polnisch)
11. grammatische Konstruktionshinweise bezüglich der Valenz und Rektion zu jedem ZS-Übersetzungsvorschlag (auf Polnisch)
12. paradigmatische Verweise auf sinnverwandte Lexeme, Synonyme, Antonyme, Hypo- und Hyperonyme des jeweiligen ZS-Äquivalenten (zur Abgrenzung von ZS-Äquivalenten deutlich als solche gekennzeichnet)
13. ggf. Verweis(e) auf graphische Illustration(en)
14. ggf. Verweis auf eine verdichtete Artikelübersicht<sup>26</sup>

Zusätzlich im Wort-Modul (d.h. wenn ein Einwortlexem als Äquivalent in Frage kommt) auf der Ebene des jeweiligen Semems:

1. ggf. lexikalische Informationen über die bevorzugten Kontextpartner des jeweiligen ZS-Äquivalentes (Verweise auf Kollokationen)
2. ggf. Verweise auf Redewendungen mit dem jeweiligen ZS-Äquivalent (Idiome, „stehende“ Vergleiche, Routineformeln etc.)
3. ggf. Verweise auf Sprichwörter in denen das ZS-Äquivalent vorkommt
4. ggf. paradigmatische Verweise auf Mitglieder der Wortfamilie zur Veranschaulichung des Wortbildungspotentials des ZS-Äquivalentes

<sup>25</sup> Zur Angabe der jeweiligen Lemmabedeutung soll man anmerken, daß sie in vielen Fällen – besonders bei Verben und Adjektiven – mit der Angabe der Kollokation des Lemma gleichzusetzen ist.

<sup>26</sup> Das impliziert die Notwendigkeit, einige metasprachlichen Angaben zweifach, d.h. einmal in verdichteter und einmal in expliziter Form, in der Datenbasis abzulegen. Die verdichteten Angaben lassen sich auch in einer Printversion des Wörterbuchs einsetzen, falls solche generiert werden soll.

Für die deutschen Benutzer, sowohl im Wort- als auch im Phrasem-Modul:

1. Informationen über Lautform (Rechtschreibung) und Aussprache (als Sprachausgabe und IPA-Angabe) des AS-Lexems
2. ggf. eine Herkunftsangabe (auf Deutsch)
3. ggf. eine Paraphrase der jeweiligen Bedeutung des AS-Lexems (auf Deutsch)
4. ein Beispiel der Verwendung des AS-Lexems im Kontext
5. eine Übersetzung des Beispiels
6. ggf. semantisch-pragmatische Kommentare zur denotativen Rezeption des AS-Lexems (auf Deutsch)
7. ein oder mehrere ZS-Äquivalente als Vorschläge für die Herübersetzung
8. grammatische Konstruktionshinweise bezgl. Valenz und Rektion des AS-Lemma (auf Deutsch)
9. ggf. ein kontextsensitiver Verweis auf ein entsprechendes Kapitel der deutschen Grammatikhilfe (d.h. zur kleinen Grammatik des Polnischen auf Deutsch)
10. grammatische Konstruktionshinweise bezüglich der Valenz und Rektion zu jedem Übersetzungsvorschlag (auf Deutsch)
11. paradigmatische Verweise auf sinnverwandte Lexeme, Synonyme, Antonyme, Hypono- und Hyperonyme des jeweiligen AS-Semems (deutlich als solche gekennzeichnet)
12. bei sog. Nulläquivalenz eine deutsche Paraphrase der Bedeutung (deutlich als solche gekennzeichnet) als ein einsetzbares Translat
13. ggf. Verweis(e) auf graphische Illustration(en)
14. ggf. Verweis auf eine verdichtete Artikelübersicht

Zusätzlich im Wort-Modul (d.h. wenn das ZS-Lemma ein Einwortlexem ist) auf der Ebene des jeweiligen Semems:

1. ggf. lexikalische Informationen über die bevorzugten Kontextpartner des AS-Lexems (Verweise auf Kollokationen)
2. ggf. Verweise auf Redewendungen mit dem AS-Lemma (Idiome, „stehende“ Vergleiche, Routineformeln etc.)
3. ggf. Verweise auf Sprichwörter in denen AS-Lemma vorkommt
4. ggf. paradigmatische Verweise auf Mitglieder der Wortfamilie zur Veranschaulichung des Wortbildungspotentials des AS-Lexems

Es ist nicht möglich, alle diese Informationstypen in einem Zug mit den lexikalischen Daten zu füllen, das kann nur nach und nach geschehen. Man muß also auch hier das Prinzip der Modularität walten lassen. Dennoch ist für den Anfang ein Grundgerüst zu erstellen, in dem alle vorgesehenen Angabetypen tatsächlich vollständig existieren. Auf dieses Fundament kann man dann weiter bauen, d.h. die für weitere Phasen vorgesehenen Instanzen (konkrete Knoteninhalte) hinzufügen. In der Hypertextbasis müssen jedoch für diese Daten entsprechende Informationstypen, die auf dem Grundgerüst aufbauen, im voraus vorgesehen sein.

Die im folgenden Kapitel dargestellten Abbildungen veranschaulichen die Beziehungen zwischen den lexikographischen Datentypen am Beispiel des Lemma *piwo* im polnisch-deutschen Teil des Wörterbuchs sowie der mit ihm assoziierten lexikalischen Daten: Es handelt sich um Lemmata (Ein- und Mehrwortlexeme) sowie die auf sie adressierten Angaben. Die Pfeile (in der Abb. 3 und Abb. 4) drücken lediglich Querverbindungen zwischen den einzelnen Knoten (Kästchen) aus. Diese Verbindungen sollen als Hyperlinks umgesetzt werden, die Pfeile sagen jedoch nichts über die Linktypen aus. Ob die jeweiligen Links ein- oder bidirektional, bzw. als sog. Multilinks realisiert werden sollen, muß im Einzelfall entschieden werden. Die Entscheidung über die Linktypen wird zwar generell auf der Ebene des Schemas durch das im ‚SchemaEditor‘ erstellte abstrakte Modell getroffen. Auf



der Ebene der Wörterbuchinstanzen ist es aber möglich – und wünschenswert – die Default-Attribute der Links gegebenenfalls (im ‚StrukturEditor‘) zu ändern.

## 6.2 Die Datenmodellierung

Die Abbildung 3 zeigt ein abstraktes Schema mit den lexikographischen Daten zum Lemma *piwo* (δτ. ‚Bier‘). Das Einwortlexem (EWL) *piwo* ist auf der einen Seite mit Angaben zur Grammatik (Genus- und Pluralform, Deklinationsmuster) und zur Aussprache (Ton und IPA-transkription) sowie mit morphologisch- und sinnverwandten Lexemen verknüpft, auf der anderen Seite ist das Einwortlexem ‚piwo‘ mit den Lemmata verbunden, deren Bestandteil es bildet. So wird z.B. auf das Mehrwortlexem (MWL) *małe piwo* (dt. Kleinigkeit, Bagatelle) aus dem Knoten ‚piwo‘ (Bier) und aus dem Knoten ‚małe‘ (klein) verwiesen. Sowohl das EWL ‚piwo‘ als auch alle Idiome, Sprichwörter und Kollokationen, die das Wort ‚piwo‘ beinhalten, sind mit den in Frage kommenden deutschen Äquivalenten verlinkt. (In der Abbildung 3 sind der Übersichtlichkeit halber nur Verweise auf ‚Bier‘ und ‚Gerstensaft‘ aus ‚piwo‘ dargestellt). Kollokationen bzw. semantische Zusammensetzungen *piwo butelkowe* (Flaschenbier), *warzyć piwo* (Bier brauen) und *toczyć piwo* (Bier zapfen) sind mit entsprechenden graphischen Illustrationen verbunden. Aus den Bildern sollen auch Verweise auf Lexeme im Sach- und Situationsfeld möglich sein: *kapsel* (Bierdeckel), *pianka* (Schaumkrone), *piwowar* (Brauer) usw.

Die Auskünfte über Genus, Plural und Aussprache zu ‚piwo‘ sind natürlich für die polnischen Benutzer redundant, die in der Abbildung 3 grau schattierten Knoten mit den Angaben zu den deutschen Äquivalenten – ‚Bier‘ und ‚Gerstensaft‘ – (hier nur mit einem Beispiel angedeutet) sind wiederum für die deutschen Benutzer irrelevant.

Wenn der deutsche Benutzer das poln. Lexem ‚piwo‘ nachschlägt, werden die Informationen über Genus, Plural aus den dazugehörigen Knoten inkludiert. Der Benutzer bekommt – indem er einen entsprechenden Button anklickt – die Möglichkeit, sich alle Idiome (vgl. Abbildung 5), Sprichwörter bzw. Kollokationen mit ‚piwo‘ sowie einen kulturell-landeskundlichen Kommentar über die Stellung des Biers in Polen (auf Deutsch) oder die Wortfamilie sowie Wörter im semantischen Umfeld (auch über die Verweise auf Bilder) von ‚piwo‘ einblenden lassen. Er kann sich auf die gleiche Art und Weise über die Aussprache und Deklination von ‚piwo‘ informieren. Wenn der deutsche Benutzer ein MWL als solches identifiziert, muß er nicht unter einem der Bestandteile nachschlagen, sondern kann direkt z.B. *nawarzyć piwa* (jmd. ‚eine schöne Suppe einbrocken‘) als Suchbegriff eingeben. Daraufhin werden die Knoten mit dem grammatischen Kommentar, der Paraphrase, dem Äquivalenten, dem Beispielsatz (oder -sätzen) samt dt. Übersetzung inkludiert und Verweise auf synonyme bzw. sinnverwandte Phraseme angezeigt. Einige Angaben beinhalten weitere Informationen: z.B. die Information über das pragmatische Register bezieht sich auf die Äquivalentangabe zu ‚*nawarzyć (komuś) piwa*‘ (dt. ‚jmd. eine schöne Suppe einbrocken‘) – siehe Abbildung 4. Für diese Benutzungssituation wird in den Äquivalent-Knoten eine auf deutsch verfaßte Entität „umgangssprachlich“ inkludiert.

Manche Unterknoten – etwa mit der Kasus- und Valenzangabe oder mit dem Aktionsart-Formativ des poln. Verbs – sind nur in der AS-Sprache verfaßt, manche sind in der Datenbasis doppelt vorhanden: Die einen beinhalten lexikographische Daten auf Deutsch, die anderen auf Polnisch. Je nach der Benutzungssituation wird ein entsprechender Unterknoten eingeschlossen.



Wenn der polnische Benutzer ‚piwo‘ nachschlägt, so bleiben die Verweise auf die für Deutsche bestimmten Informationen ausmaskiert. Dem polnischen Benutzer werden Angaben zu den dt. Äquivalenten von ‚piwo‘ und Verweise auf die polnischen sinnverwandten Lexeme (*kufel, szkalnka, chmiel, słód, jeczmień, podchmielony, browar, dystrybutor*, usw.) angeboten. Die metasprachlichen Informationen (wie Paraphrasen, explizite Kommentare und Registerangaben), die sich auf das Lemma und das gegebene Äquivalent beziehen, erfolgen auf Polnisch. So bekommt etwa beim Nachschlagen des Lemma *małe piwo* (siehe Abb. 4) der Benutzer, neben dem Hinweis auf die wörtliche Bedeutung ‚kleines Bier‘, zwei idiomatische Bedeutungen zur Auswahl: ‚łatwe do załatwienia‘ (müheles zu Erledigendes) und ‚coś niewiele znaczącego, przeciętnego‘ (etwas Unbedeutendes, Mittelmäßiges). Wählt er im weiteren Verlauf des Nachschlageprozesses die erste Bedeutung der Redewendung aus, so bekommt er den Inhalt des Knoten mit dem dt. Äquivalent, in den ein Unterknoten mit der Registerangabe auf Polnisch inkludiert wurde: ‚Kleinigkeit, Bagatelle; *potocznie* Kinkerlitzchen, Pappenstein‘.

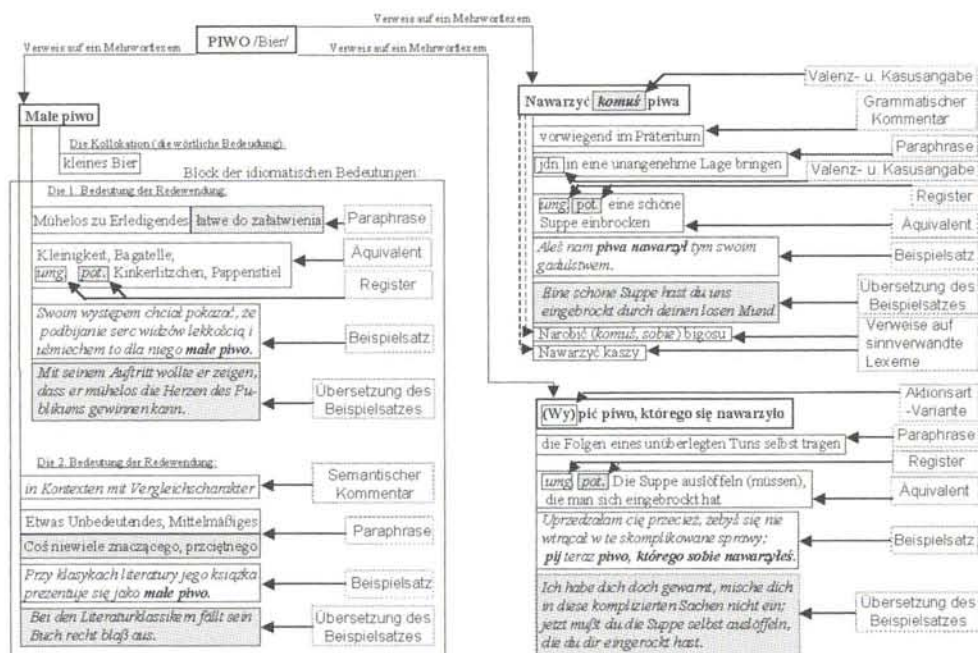


Abb. 4: Mehrwortlemmata

In der Abbildung 4 sind Angaben zu drei poln. MWL dargestellt: ‚małe piwo‘, ‚nawarzyć piwa‘ und ‚(wy)pić piwo, którego się nawarzyło‘. Da das Lemma ‚małe piwo‘ polysem ist, sind die Paraphrasen – als bedeutungsdifferenzierende Glossen – für beide Benutzergruppen nützlich und deshalb in der Datenbasis doppelt vorhanden: auf Polnisch und auf Deutsch. (Im Falle der monosemen Phraseme ist es anders, die Bedeutung, zumindest in den meisten Fällen, kann für den polnischen Benutzer als bekannt vorausgesetzt werden.) Das Lemma *nawarzyć piwa* ist mit zwei sinnverwandten Lemmata verknüpft: *narobić*

*bigosu i nawarzyć kaszy*. In Abbildung 4 stellen die Angaben in grau schattierten Kästchen beispielhaft diejenigen Informationen dar, die im Zuge der Bidirektionalisierung des Materials den lexikalischen Daten aus den analysierten Werken hinzugefügt wurden.

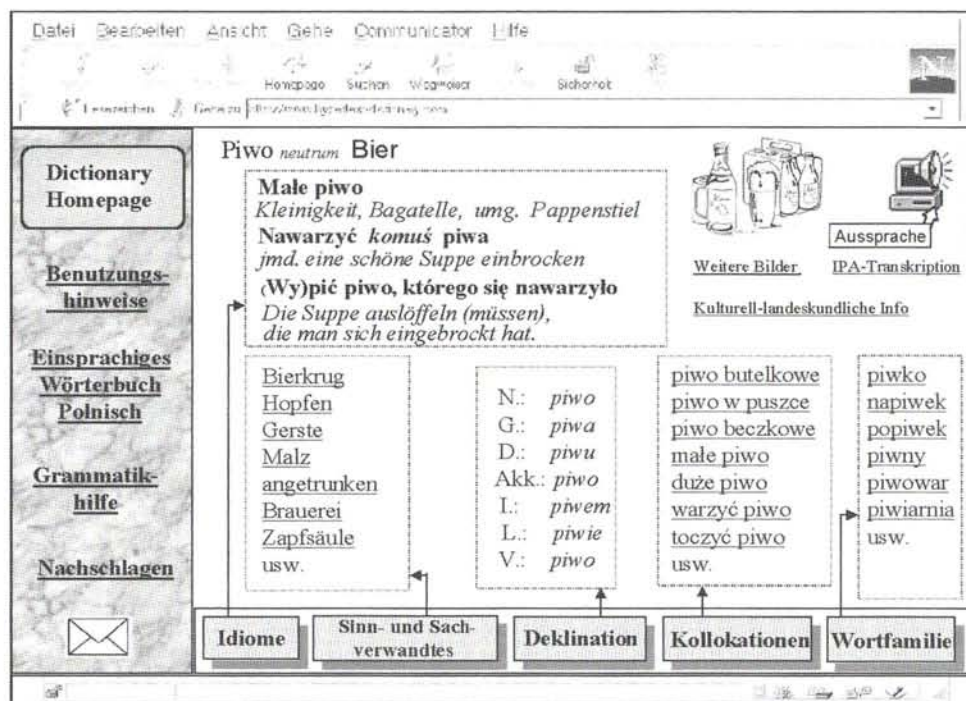


Abb. 5: Benutzeroberfläche

Auf der Benutzeroberfläche (siehe Abb. 5) sollen die deutschen Benutzer des Wörterbuchsystems durch das Anklicken der entsprechenden Buttons z.B. die Mehrwortlexemlemmata abfragen können, die das Einwortlexem 'piwo' als Bestandteil enthalten. Würde der Benutzer hingegen den Button „Deklination“ mit der Maus betätigen, so könnte er das Deklinationsparadigma für das Nomen 'piwo' auf dem Bildschirm sehen. In der Abb. 5 stellen sowohl alle Mehrwortlexeme als auch die unterstrichenen Elemente anklickbare Linkanker dar. Die Pfeile und die gepunkteten Kästchen mit lexikographischen Daten haben auf dem Bild folgende Bedeutung: wenn der Benutzer einen Button drückt, bekommt er die gewünschten (in dem mit entsprechenden Pfeil angedeuteten Kästchen dargestellten) Informationen eingeblendet. Das HTML-Layout wird – wie im Kap. 4. 2 besprochen – über die mit den einzelnen Datenobjekten assoziierten Expander in Verbindung mit dem Maskierungsmechanismus generiert.



## 7 Übersichtsknoten

Viele Autoren plädieren – mit Recht – für die Auflösung der verdichteten Angaben, wie sie in Büchern zum Einsatz kommen, wenn die Wörterbuchdaten in elektronische Nachschlagewerke konvertiert werden oder Neuentwicklungen entstehen. Es gibt jedoch Situationen, in denen die Beibehaltung der Textverdichtung auch in einem elektronischen Wörterbuch sinnvoll zu sein scheint.

Der wesentliche Nachteil des Mediums Hypertext ist bekanntlich die Zersplitterung der Informationen in feinkörnige Einheiten, deren gegenseitige Beziehungen oft nicht klar genug dargestellt werden. Die Benutzer laufen dann Gefahr, aufgrund fehlerhaft konstruierter Navigationsmechanismen den Überblick zu verlieren. Deshalb wären Überblickstafeln mit verdichteten Informationen (die den Artikeln eines Printwörterbuchs gleichen und im Hypertext als ‚information chunks‘ organisiert sind) besonders nützlich: Einerseits leisten sie Orientierungshilfe für alle Benutzer, andererseits bieten sie den kundigen Benutzern komprimierte Informationen. Ein mit den Konventionen der Printlexikographie vertrauter Benutzer kann die Abkürzungen sofort entschlüsseln, den Inhalt des Übersichtsfensters wie im Buch überfliegen und ganz gezielt die Verweise zu den ihn in der konkreten Benutzungssituation interessierenden Informationen verfolgen. Es ist allerdings wichtig, daß die Abkürzungen und Symbole auf einen Mausklick des Benutzers hin angezeigt werden können, sei es, indem sie durchgängig in ganzem Knoten mit voll ausgeschrieben Entitäten ersetzt werden, sei es, indem die Auflösung kontextsensitiv im Pop-up-Fenster erscheint. Denjenigen Benutzern, die im Umgang mit herkömmlichen Wörterbüchern ungeübt sind, helfen hoffentlich die leicht auflösbaren Textverdichtungs-elemente, die Bedeutung der in der Printlexikographie gängigen Abkürzungen für grammatische und pragmatische Kategorien zu lernen. So kann der Hypertext vielleicht auch dazu beitragen, daß die Konventionen der Printlexikographie, die für Nichtlinguisten nicht immer leicht zu verstehen sind, auf dem „Umweg“ über ein elektronisches Medium besser kennengelernt werden.

Natürlich passiert es oft, daß die Benutzer nicht nur die Abkürzungen des Wörterbuchs nicht verstehen, sondern auch mit den aufgelösten Begriffen und explizit formulierten Kommentaren zur grammatischen, semantischen oder pragmatischen Verwendung des Lemmas wenig anfangen können. Deswegen scheint es wichtig, daß sie kontextsensitiv in der „Grammatikhilfe“ nachschlagen können. Die Grammatikhilfe ist als ein separates Schema mit zwei Perspektiven angelegt, d.h. die deutsche Grammatik (auf Polnisch) und die polnische Grammatik (auf Deutsch) sind in Kapiteln mit gleicher Struktur gegliedert und lassen sich deshalb in einem Schema unterbringen.

## 8 Das objektorientierte Datenmodell

Die Implementierung der geeigneten Datenmodelle, die die oben umrissene Funktionalität unterstützen, ist mit dem Programm SchemaText dank komplexer Merkmale der Objektorientiertheit möglich.<sup>27</sup> Zur genauen Veranschaulichung all dieser Merkmale müßte man jeweils ein geeignetes Beispiel anführen, was hier aus Platzgründen nicht möglich ist. Trotzdem sollen diese Prinzipien nicht unerwähnt bleiben, da ihnen in Zukunft bei der Entwicklung von Computerwörterbüchern eine immer größere Rolle zukommen wird:

- Prinzip der Modularisierung (Vgl. Kap. 4. 1, 4. 2 und Kap. 10).
- Prinzip der Maskierung (Vgl. Kap. 4. 2).
- Prinzip der Typisierung von Knoten und Links (Vgl. Kap. 4. 1).
- Prinzip der Vererbung (Es regelt u.a. die Vererbung von Layouteigenschaften, die in den Expandern implementiert wurden. Dadurch, daß alle Informationstypen gleichen Strukturvorgaben genügen, können Eigenschaften von denen auf höherer Ebenen auf solche unterer Ebenen vererbt werden).
- Prinzip der Mehrfachverwendung (Manche Knoten können in mehreren Kontexten verwendet werden, etwa Beispielsätze und ihre Übersetzungen, Bilder oder Paraphasen, die sinngemäß mehreren Lexemen entsprechen).

## 9 Die Rolle der Bilder

Die zweisprachigen Printwörterbücher setzen Bilder entweder gar nicht oder nur sehr sparsam ein (vgl. Hupka 1989:39–43), Illustrationen würden zusätzlichen Platz erfordern und die Kosten in die Höhe treiben.<sup>27</sup> Außerdem stellt sich zunächst die Frage, ob die Bilder im zweisprachigen Wörterbuch überhaupt Sinn machen und, wenn ja, welche Arten von Illustrationen zu welchem Zweck an welcher Stelle am besten geeignet sind. Fest steht, daß die Bilder in einigen PC-Wörterbüchern bis jetzt nur der werbewirksamen „Multimedialität“ wegen eingesetzt werden und durch den unüberlegten und konzeptionslosen Einsatz von Illustrationen kaum zur Verbesserung des Wörterbuchs beitragen. Das LEKSYKONIA-Wörterbuch präsentiert etwa beim Lemma *Egipt* ein Bild mit der Flagge Ägyptens (das Verfahren gilt für alle Länder) und beim Lemma *Glühbirne* ein Photo einer Glühbirne. Dieser konsequente Illustrationsansatz, substantivische Lemmata mit mehr oder weniger aussagekräftigen Photographien zu illustrieren, kann den Benutzern wohl wenige Vorteile bieten. Mehr Sinn machen schon einige unikale Illustrationen, z.B. für verschiedene Gemüsesorten (Lemmata *Kohl*, *Erbse*, *Gurke* usw.), wobei jedes dieser Lemmata einen Verweis auf den hyperonymischen Eintrag *Gemüse* enthält, der wiederum Verweise auf illustrierte Lemmata der einzelnen Gemüsesorten enthält.<sup>29</sup>

Für die Lernfunktion ist es wichtig, verschiedene Bildtypen zu differenzieren und zu überlegen, mit welchen lexikographischen Texten welche Typen von Illustrationen verbunden werden. Hupka (1989, 200–202 und 235–242) unterscheidet zwischen neun Typen von lexikographischen Illustrationen in gedruckten Nachschlagewerken. Alle diese Typen

<sup>27</sup> Das ganze Wörterbuch stellt als ein elektronisches Dokument – in der Sprache der Informatik gesagt – ein Objekt dar. Auf den niedrigeren Ebenen läßt sich dieses Objekt in weitere Objekte „zerlegen“. Jedes dieser Objekte verfügt potentiell über gewisse Eigenschaften, die von einem Laserdrucker, einem WWW-Browser oder einer Belichtungsmaschine anders interpretiert werden. Jedes Objekt trägt gleichsam die „Rezepte“ für seine Verarbeitung mit sich. Das Objekt Wörterbuch zeigt somit ein anders „Gesicht“ gegenüber verschiedenen Ausgabemedien und verschiedenen Benutzern.

<sup>28</sup> Eine Ausnahme stellt nur das zweisprachige Bildwörterbuch dar.

<sup>29</sup> Eine nomenklatorische Illustration, die alle Gemüsesorten auf einmal präsentiert, und somit dem Benutzer die Möglichkeit gibt, die ihm unbekannte ZS-Bezeichnung für eine bestimmte Gemüsesorte (z.B. durch das Anklicken des entsprechenden Bildes) kennenzulernen, gibt es im LEKSYKONIA-Wörterbuch nicht.

können auch für ein elektronisches zweisprachiges Wörterbuch sehr vorteilhaft sein, wenn sie zweckmäßig angewandt werden:

1. Unikale Illustrationen helfen die referenzierten Denotate zu identifizieren. Sie sind in der Herübersetzungssituation bei Lemmata etwa aus dem technischen und naturwissenschaftlichen Bereich einzusetzen, da, wo der Benutzer auch in seiner Muttersprache über keine oder nur vage Vorstellungen von dem mit dem Lemma bezeichneten Begriff verfügt. Eine unikale Illustration kann zugleich sequentiell sein (Typ 3). Unikale Illustrationen dienen der Absicherung und der besseren Memorierung der Information, die durch die Äquivalentangabe vermittelt wird.
2. Aufzählende Illustrationen sind für Lexeme vorgesehen, die eine Klasse von Gegenständen unter einer gemeinsamen Bezeichnung vereinen. Das Aussehen der Referenten kann stark voneinander abweichen. (Hupka nennt als Beispiel das Lemma *Säge*. Zu illustrieren wären dann *Stichsäge*, *Bandsäge*, *Metallsäge*, *Kreissäge*, *Motorsäge* etc.).
3. Sequentielle Illustrationen veranschaulichen Bewegungsabläufe und sind im zweisprachigen Wörterbuch nur in Ausnahmefällen einzusetzen, etwa wenn Verben (*tröpfeln*, *schaukeln*, *rutschen*, *auskippen*, *zwinkern* etc.), Substantive (*Purzelbaum*, *Liegestütz*), Kollokationen (*ein Rad schlagen*) oder kulturspezifische Lemmata (bayer. *Schuhplattler*) illustriert werden sollen. Als technisches Mittel eignen sich hierzu besonders Videosequenzen.
4. Strukturelle Illustrationen verdeutlichen die Teil-Ganzes-Beziehungen und helfen z.B. Teile eines Fahrrads zu benennen.
5. Funktionale Illustrationen zeigen die referierten Objekte im Kontext der angrenzenden Umgebung, die zum Verständnis des Wortes notwendig ist, da sie einen integralen, nicht abtrennbaren Teil einer größeren Einheit darstellen (z.B. das Lemma *Pupille*). Sie sind auf ähnliche Weise wie die unikalen und enzyklopädischen Illustrationen einzusetzen.
6. Nomenklatorische Illustrationen geben Instanzen einer Klasse wieder, z.B. *Gemüse: Gurke, Tomate* usw.
7. Szenische Illustrationen bilden einen Realitätsausschnitt ab: z.B. *Campingplatz*, *Bahnhofshalle*, *Krankenhauszimmer*, *Büro*. Sie eignen sich auch zur Veranschaulichung der Verwendung von Präpositionen.
8. Funktionsschemata könnten z.B. – in Form von Diagrammen – die Frequenz einzelner Äquivalente graphisch präsentieren, vorausgesetzt, es gäbe ein repräsentatives Korpus, an dem die Frequenzmessungen und andere statistischen Auswertungen vornehmbar wären.<sup>30</sup>
9. Enzyklopädische Illustrationen ergänzen Definitionen von landes- und kulturspezifischen Lexemen und bilden typisch deutsche Gegenstände wie etwa *Adventskranz*, *Gartenzwerg* bzw. unikale Eigennamen-Denotate wie *Friedensengel* ab.

Einen für die zweisprachigen Wörterbücher prädestinierten Bildertyp stellt die kontrastierende Illustration dar, die in der Typologie von Hupka (1989) nicht erfaßt wurde.<sup>31</sup> Kontrastierende Illustrationen wurden z.B. im Umlaut-Wörterbuch (monodirektional, d.h. nur für poln. Benutzer) eingesetzt, und verdeutlichen (immer in Form von einer Photographie) die besonderen Merkmale, die bestimmte Denotate in der deutschen Realität charakterisieren: z.B. *Telefonzelle*, *Briefkasten*, *Fahrkarte*, *Aufenthaltserlaubnis*, *Polizist*, *Streifenwagen*, *Imbißbude*, diverse Schilder wie *Naturschutzgebiet*, *Bushaltestelle*, *U-Bahn*

<sup>30</sup> Ein großes, ausbalanciertes, projekteigenes Korpus, in dem Lexikographen und Benutzer recherchieren könnten, wäre natürlich sehr wünschenswert. Dieses Desiderat, ähnlich wie das Morphologieanalyzesystem, muß jedoch als separates Unterfangen betrachtet und ggf. realisiert werden.

<sup>31</sup> Unerwähnt bleibt bei Hupka auch die komplementäre Illustration, z.B. *Steckdose: Stecker* (vgl. LGDaF:942) oder *aufblasen:platzen* und die polysemieaufdeckende Illustration, z.B. *Flügel: Flügel eines Vogels, Konzertflügel, Gebäudeflügel* usw.

usw.<sup>32</sup> Diese Illustrationen vermitteln dem Benutzer zusätzliche landeskundliche Kenntnisse, die ihm einerseits die Orientierung während seiner Deutschlandreise erleichtern, und andererseits zusätzliche Konnotationen für die Memorierung der Begriffe liefern. (Allerdings muß noch der Einfluß einer bildhaften Kontrastierung von mutter- und fremdsprachlichen Referenten auf den Lernprozeß besser untersucht werden).

Alle Typen von Bildern haben primär die Aufgabe, den Lernprozeß zu fördern, indem sie entweder zusätzliche enzyklopädische (onomasiologische und semasiologische) Informationen vermitteln oder assoziative und paradigmatische Beziehungen der Lexik veranschaulichen.<sup>33</sup> Sie sind mit den lexikographischen Textsegmenten gleichrangig, folglich gilt es auch für sie zu entscheiden, in welcher Benutzungssituation sie eingesetzt werden sollen. Ferner gilt es zu bestimmen

- von welchem Knoten auf die Illustration verwiesen wird, d.h. ob der Verweis auf der Ebene des Äquivalentes bzw. der Paraphrase erfolgt oder auf der Ebene des semantischen bzw. pragmatischen Kommentars,
- ob, und wenn ja, wann eine unikale Illustration einen Teil einer aufzählenden, strukturellen, szenischen oder nomenklatorischen Illustration bilden soll. Das würde bedeuten, daß Illustrationen ggf. untereinander verlinkt werden müssen.

## 10 Ausblick

Das Prinzip der Modularisierung ist auf der einen Seite notwendig, um das System inhaltlich zu strukturieren (siehe Kap. 4.2), und auf der anderen Seite, um es von anderen Webprojekten abzugrenzen. Es ist nämlich nicht leicht, der Versuchung zu widerstehen, auch andere als Äquivalenz-Wörterbuchtypen in das System zu integrieren, und dadurch Umfang und Funktionalität immer weiter auszubauen. Besonders im Hinblick auf die Diskussion über erklärende zweisprachige (Übersetzungs-)Wörterbücher (vgl. Duda 1986, Worbs 1997) könnte eine solche zu weite Öffnung bedeuten, daß einsprachige Bedeutungswörterbücher und enzyklopädische Lexika mit eingeschlossen wären, auf die ein Zugriff im Prinzip sehr wünschenswert wäre. Solch ein Ansatz, der für die Benutzer mit Sicherheit einen großen Vorteil bedeuten würde, muß jedoch aus praktischen Gründen modularisiert werden, d.h. es sind getrennte, von zweisprachigen Wörterbüchern unabhängige einsprachige Webprojekte vonnöten, auf deren Lemmata punktuell verwiesen werden kann. Das erlaubt den Autoren nach dem Motto *multum non multa* zu verfahren, birgt jedoch die Gefahr in sich, daß die referentielle Integrität u.U. nicht mehr gewährleistet ist. Hier ist man mit dem grundsätzlichen Problem des WWW konfrontiert:

„Während in einem geschlossenen Hypertextsystem, z.B. einem auf CD-ROM vertriebenen Lernsystem, alle Links, denen der Leser nachgehen kann, vom Autor kontrolliert und mit einer klaren

<sup>32</sup> Das Beispiel ‚Telefonzelle‘ zeigt, daß auch Illustrationen schnell obsolet werden können. Die im Umlaut-Wörterbuch abgebildete gelbe Telefonzelle der Post wurde mittlerweile durch eine rosa-graue Telekom-Zelle beinahe verdrängt. Wenn der Telefonmarkt in Zukunft weiter privatisiert wird, kann es sein, daß es bald kein dominierendes länderspezifisches Design für eine Telefonzelle mehr gibt. Dann soll auch die Illustration zu diesem Lemma wegfallen.

<sup>33</sup> Auskünfte über die Zusammenhänge zwischen dem Text, Bild und Ton in neuen Medien gibt aus der Sicht der kognitiven Psychologie Weidenmann (1997:108ff.).



Intention angelegt und somit integraler Bestandteil des Textes sind, gilt das in einem WWW-Dokument nur für einen Teil der Links, nämlich die internen. Ansonsten ist jedoch das eigene Dokument ein Knoten in einem unkontrollierten, weil unkontrollierbaren, dynamischen, globalen Netzwerk. Bereits die von einem Autor selbst angelegten, von seinem eigenen Dokument nach außen führenden Links unterliegen nur bedingt seiner Kontrolle, da er auf die Kontinuität des referenzierten Materials keinen Einfluß hat. Noch schwieriger ist es mit den externen Verweisen auf den eigenen Text: Hier wird deutlich, dass die pragmatischen Rahmenbedingungen der Textrezeption in einem offenen Hypertextsystem zu einer qualitativen Nebenbestimmung des Textbegriffs führen. [...] Durch dieses (vom Autor selbst nicht kontrollierbare) Wegfallen oder Veralten der Referenzen und die potentiellen Veränderungen des referenzierten Materials verlieren Texte als WWW-Dokumente ihren statischen Charakter, sie verändern sich im Laufe der Zeit“ (Frisch 1998:228–229).

Deswegen wäre es sinnvoll, eine Koordination mit anderen Webprojekten vorzunehmen. Ideale Bedingungen dafür sind geschaffen, wenn man dieselbe Software als Dokumentenverwaltungssystem anwendet.

Sogar ein sehr umfangreiches zweisprachiges Wörterbuch kann nicht alle möglichen Kontexte der Lemmaverwendung berücksichtigen – künftige Wörterbücher sollten also einen Anschluß an Korpusabfragesysteme bieten. Solche Systeme, heute noch selten und nur experimentell betrieben, können bald als ein wichtiges lexikographisches Arbeitsmittel dienen, nicht nur für die Wörterbuchautoren, sondern auch für die gewöhnlichen Benutzer.<sup>34</sup> Weitere Links, etwa auf Seiten mit Online-Sprachkursen, Vokabeltrainern, Informationen zur Landeskunde u.ä. sind denkbar und wünschenswert.

Ebenso nützlich wie ein ausgereifter Bildereinsatz kann die Verwendung der Töne sein. Heute noch, wegen beschränkter Bandbreite der Leitungen, eher die Domäne der Offline-Wörterbücher, bald jedoch wird die Tonübertragung im Internet wahrscheinlich kein technisches Problem mehr darstellen. Dann können ganze Redewendungen vertont werden.

Neue Lemmatypen wie Neologismen, Fachtermini, aber vor allem zahlreiche Varianten eines Lexems, sobald sich in diesen Varianten veränderte Semantik des Lemma manifestiert (sehr wichtig für den polnischen Aspekt und die Diminutiva), lassen sich endlich in das zweisprachige Wörterbuch aufnehmen. Das ist für die stark flektierenden Sprachen wie Polnisch und Deutsch von besonderer Bedeutung. Man findet z.B. bislang in keinem poln.-dt. Wörterbuch alle Verbindungen des Typs Prefix+Verb des poln. Verbs *gotować* (*zgotować, przegotować, wygotować, nagotować, rozzgotować, odgotować, dogotować*). Ein Lemmatisierungsalgorithmus, der alle polnischen Formen *„ić, „ójść, „chodzę, „szedł“* usw. auf die Grundformen *„chodzić, /,ić“* (wo als Äquivalent *„gehen“* in Frage kommt) zurückführen würde, wäre der komplexen Problematik des Aspektes im Polnischen und der Aktionsarten im Deutschen nicht gerecht.

Die vom Autor durchgeführten Experimente, mit dem Ziel, ausgewählte Wörterbuchfragmente in die mit SchemaText erstellte Hypertextstruktur zu überführen, lassen sich noch nicht als völlig erfolgreich bezeichnen – die vollständige Implementierung und Evaluierung eines Prototyps konnte deswegen noch nicht in Angriff genommen werden. Ein inhärentes semantisch-pragmatisches Organisationsprinzip, das eine dynamische Generierung von Online-Wörterbucheinträgen gemäß modellierter Benutzerprofile unterstützt, stößt noch auf technische Hürden auf der Produktionsseite und auf urheberrechtliche Probleme bei der Redaktion der Inhalte. Nach wie vor gilt die Aussage von Drewek (1990:272): „Das Urheberrecht regelt nicht, wem die formale Spezifikation eines Wörterbuchdesigns gehört, obwohl sich hierin die individuelle Handschrift von LexikographInnen

<sup>34</sup> Vgl. Figge (1994), Martin (1995), Krishnamurthy (1996).

schöpferisch manifestiert. Hier öffnet sich eine rechtliche Grauzone, die vom Knowledge Engineering her bekannt ist“.

Auf der technischen Seite ergeben sich außerdem schwerwiegende Probleme mit der Unterstützung der länderspezifischen Buchstaben und der Sonderzeichen. Die proprietären Datenformate können zwar mit diesen Zeichen in der Regel gut umgehen, aber wenn es darum geht, die Daten zwischen den Softwarekomponenten oder Anwendungen auszutauschen, kommt man um trickreiche Methoden nicht herum. Der lautstarken Werbung mancher Hersteller, in ihren Produkten den Unicode-Standard zu unterstützen, darf man nicht immer Glauben schenken.

Der im vorliegenden Beitrag skizzierte Ansatz würde ernste Konsequenzen sowohl für die Lexikographen als auch für die Wörterbuchbenutzer mit sich bringen, und zwar einen Paradigmenwechsel bei der Erstellung von lexikographischen Werken und eine Reorganisation der Wörterbucharbeit durch das Aufheben der strikten Grenze zwischen Makro- und Mikrostrukturen, durch Mehrfachverwendung der Inhalte sowie durch ein dichtes Verknüpfen von lexikographischen Daten mithilfe von Hyperlinks. In das Wörterbuch würden sich nicht nur neue Lemmata und Angaben jederzeit aufnehmen lassen, sondern auch veraltete – wenn nötig – entfernen. Dadurch könnte das Wörterbuch stets aktuell bleiben, es wäre aber nie endgültig fertiggestellt.

Ein ähnlicher Paradigmenwechsel wäre für das Benutzerverhalten notwendig, um das Wörterbuch als ein interaktives Medium erfolgreich zu gebrauchen und es eventuell mit- bzw. weitergestalten. Eine entscheidende Rolle bei weiterer Entwicklung von Online-Nachschlagewerken wird wahrscheinlich die Flexibilität der Systeme spielen: erst wenn Hypertextwörterbücher den individuellen Bedürfnissen verschiedener Benutzergruppen entsprechen werden, kann ihre Akzeptanz und damit auch praktische Relevanz steigen. Hypertexte bringen für die Benutzer bekanntlich nicht nur Vorteile mit sich, sondern bergen etliche Gefahren – Unübersichtlichkeit, Ablenkung und Angebotsfülle können bei der Rezeption der Informationen sehr hinderlich sein. Helfen können gut strukturierte Hypertextbasen und leistungsfähige Navigationsmechanismen. Zudem müssen wir noch mehr über die Beziehungen zwischen dem mentalen Lexikon und der formalen Repräsentation bzw. Rekonstruktion des lexikalischen Wissens für die Konstruktion von Hypertextwörterbüchern in Erfahrung bringen.

## 11 Literatur

### 11.1 Wörterbücher und Lexika

- DBWDP = DUDEN BILDWÖRTERBUCH DEUTSCH UND POLNISCH. Bibliographisches Institut & FA Brockhaus AG & Wiedza Powszechna. Aktualisierte Neuauflage. Warszawa 1998.
- GWPD/DP = GROSSWÖRTERBUCH DEUTSCH-POLNISCH/POLNISCH-DEUTSCH von Juliusz Ippoldt und Jan Piprek. Wiedza Powszechna & Verlag Enzyklopädie Leipzig. Warszawa 1969/1971. 4 Bd.
- GlobeDisc 2.0 GRUNDWORTSCHATZ POLNISCH. (Deutsch-Polnisch/Polnisch-Deutsch). Rossipaul Verlag München & EPP Nürnberg 1996. 3 Disketten.
- HWBPD/DP = HANDWÖRTERBUCH DEUTSCH-POLNISCH/POLNISCH-DEUTSCH von Andrzej Bzdęga, Jan Chodera und Stefan Kubica. Wiedza Powszechna. Warszawa 1971. 2 Bd.
- Infopedia 2.0: COMPTON'S ENZYKLOPÄDIE, OXFORD ENGLISH DICTIONARY ENG.-GERM./GERM.-ENG., WELTLÄNDER-LEXIKON, WELTATLAS, RECHTSCHREIBWÖRTERBUCH, FALKEN'S BUCH DER ZITATE,

- FREMDWÖRTERBUCH, SYNONYMENWÖRTERBUCH, ZWEIFELSFÄLLE DEUTSCH, Leonardo der Erfinder, Jahrgang 1995 der Zeitschrift „Bild der Wissenschaft“. Tewi Verlag. München 1997. 3 CD-ROMs.
- INFOROM '98/99: BERTELSMANN UNIVERSALLEXIKON, WÖRTERBUCH ZUR ALTEN UND ZUR NEUEN RECHTSCHREIBUNG, WÖRTERBUCH ENGLISCH-DEUTSCH/DEUTSCH-ENGLISCH, WÖRTERBUCH FRANZÖSISCH-DEUTSCH/DEUTSCH-FRANZÖSISCH, WELTATLAS, AUTOATLAS DEUTSCHLAND, WAHRIG-FREMDWÖRTERLEXIKON, Rechtschreibprüfung für MS Winword. Bertelsmann. München 1998. 2 CD-ROMs.
- ISNP = ILUSTROWANY SŁOWNIK NIEMIECKO-POLSKI von Aldona Brzeska und Alojzy Brzeski. Wiedza Powszechna. Warszawa 1975.
- LANGENSCHIEDTS POWER DICTIONARY ENGLISCH (ENGLISCH-DEUTSCH/DEUTSCH-ENGLISCH) hrsg. von Vicent Docherty. Langenscheidt. München 1997.
- LANGENSCHIEDTS HANDWÖRTERBUCH ENGLISCH (ENGLISCH-DEUTSCH/DEUTSCH-ENGLISCH) - DIE PROFILINE. In der PC-Bibliothek 2.0. Langenscheidt. München 1997. 1 CD-ROM.
- LE PETIT ROBERT. LA REFERENCE DE LA LANGUE FRANÇAISE. Liris Interactive. Paris 1997. 1 CD-ROM.
- LGDaF = LANGENSCHIEDTS GROßWÖRTERBUCH DEUTSCH ALS FREMDSPRACHE hrsg. von Dieter Götz, Günther Haensch und Hans Wellmann. Langenscheidt. München 1993.
- LEKSYKONIA 4.0 MULTIMEDIALNY SŁOWNIK NIEMIECKO-POLSKI I POLSKO-NIEMIECKI. WNT Warszawa & Lexland Knurów 1997. 1 CD-ROM.
- LexiRom 3.0: MEYERS LEXIKON IN DREI BÄNDEN, DUDEN DEUTSCHE RECHTSCHREIBUNG, DUDEN FREMDWÖRTERBUCH, DUDEN WÖRTERBUCH DER SINN- UND SACHVERWANDTEN WÖRTER, LANGENSCHIEDTS TASCHENWÖRTERBUCH ENGLISCH, WELTATLAS. Bibliographisches Institut & FA Brockhaus AG. Mannheim 1998. 1 CD-ROM.
- OED = OXFORD ENGLISH DICTIONARY ON CD-ROM prepared by John Simpson and Edmund Weiner. Oxford University Press & AND Software Inc. Second Edition. Oxford 1996. 1 CD-ROM.
- OXFORD-HACHETTE ENGLISH AND FRENCH DICTIONARY hrsg. von Marie-Helene Corréard und V. Grundy. Oxford University Press. Oxford 1994.
- PC-Bibliothek 2.0. Langenscheidts HANDWÖRTERBUCH ENGLISCH. München 1997. 1 CD-ROM.
- PWP = PHRASEOLOGISCHES WÖRTERBUCH POLNISCH von Erika Ehegötz, Walter Duda, Maria Frenzel, Maria Gehrmann und Stanisław Skorupka. Verlag Enzyklopädie. Leipzig 1990.
- SPRACHBROCKHAUS. DEUTSCHES BILDWÖRTERBUCH VON A-Z. Neubearbeitung. Wiesbaden 1984.
- STJN = SŁOWNIK TEMATYCZNY JĘZYKA NIEMIECKIEGO von Grażyna Hatałał und Małgorzata Bielicka. Kanion. Zielona Góra 1996.
- SWBPD = SPRICHWÖRTERBUCH DEUTSCH-POLNISCH/POLNISCH-DEUTSCH von Alina Wójcik und Horst Ziebart. Wiedza Powszechna. Warszawa 1997.
- TWBP = LANGENSCHIEDTS TASCHENWÖRTERBUCH POLNISCH von Stanisław Walewski. Langenscheidt. München 1979.
- UMLAUT. MULTIMEDIALNY SŁOWNIK EDUKACYJNY NIEMIECKO-POLSKI I POLSKO-NIEMIECKI. Wiedza Powszechna & Premiere Training Company Ltd. Warszawa 1996. 1 CD-ROM.

## 11.2 Sonstige Literatur

- Arents, Hans C.; Bogaerts, Walter (1991): Towards an Architecture For Third-order Hypermedia Systems. In: Hypermedia 3 (2), 133–152.
- (1993): Concept-based Retrieval of Hypermedia Information: From Term Indexing to Semantic Hyperindexing. In: Information Processing & Management 29 (3), 373–386.
- Atkins, Sue (1996): Bilingual Dictionaries: Past, Present and Future. In: Gellerstam, Martin; Järborg, Jerker; Malmgren, Sven-Göran; Norén, Kerstin; Rogström, Lena; Pappmehl, Catarina (Hg.): Euralex '96 Proceedings. 7<sup>th</sup> International Congress on Lexicography. – Göteborg. Sweden, 515–546.
- Bläsi, Christoph; Koch Heinz-Detlev; Lehr, Andrea; Wiegand, Herbert Ernst (1994): Lexicographic Standards and Reusability – From Metalexicographic Description to a Parsing Procedure. In: Lexicographica 10/94, 221–248.

- Breidt, Elisabeth; Segond, Frédérique; Valetto, Giuseppe (1996): Local Grammars for the Description of Multi-Word Lexemes and their Automatic Recognition in Texts. In: Kiefer, Ferenc; Kiss, Gábor; Pajzs Júlia (Hg.): *Proceedings of the 4<sup>th</sup> International Conference on Computational Lexicography COMPLEX'96*. – Budapest, Hungary, 9–28.
- Conklin, Jeff (1987): Hypertext: An Introduction and Survey. In: *IEEE Computer* 20 (9), 17–41.
- Drewek, Raimund (1990): Skizzen aus der real produzierenden Lexikographie. In: Schaefer, Burkhard; Rieger, Burghard (Hg.): *Lexikon und Lexikographie*. – Hildesheim: Olms (= Sprache und Computer 11, Vorträge aus der Jahrestagung der Gesellschaft für Linguistische Datenverarbeitung e.V., Siegen, 26.–28.3.1990), 265–274.
- Duda, Walter (1986): Ein „aktives“ russisch-deutsches Wörterbuch für deutschsprachige Benutzer? In: *Beiträge zur Lexikographie slawischer Sprachen*. – Berlin (Ost): Akademie der Wissenschaften der DDR (= Linguistische Studien, Reihe A, 147), 9–15.
- Feldweg, Helmut (1997): Wörterbücher und neue Medien: Alter Wein in neuen Schläuchen? In: *Zeitschrift für Literaturwissenschaft und Linguistik* 27/107, 110–123.
- Figge, Udo L. (1994): Lexikographische Datenverarbeitung und maschinelle Lexikographie. In: Figge, Udo L. (Hg.): *Portugiesische und portugiesisch-deutsche Lexikographie*. – Tübingen: Niemeyer (= Lexicographica, Series Maior 56), 209–221.
- Freisler, Stefan (1994): Hypertext. Eine Begriffsbestimmung. In: *Deutsche Sprache* 22/94, 19–50.
- Frisch, Elisabeth (1998): Ausgewählte Aspekte des Publizierens im WWW. In: Storrer, Angelika; Harriehausen, Bettina (Hg.): *Hypermedia für Lexikon und Grammatik*. – Tübingen: Narr (= Studien zur deutschen Sprache 12), 217–232.
- Hauser, Ralf; Storrer, Angelika (1996): Probleme und Lösungen beim Parsen von Wörterbüchern. In: Feldweg, Helmut; Hinrichs, Erhard W. (Hg.): *Lexikon und Text. Wiederverwendbare Methoden und Ressourcen zur linguistischen Erschließung des Deutschen*. – Tübingen: Niemeyer (= Lexicographica, Series Maior 73), 53–68.
- Hausmann, Franz-Josef (1991): Typologie der zweisprachigen Spezialwörterbücher. In: Hausmann, Franz-Josef; Reichmann, Oskar; Wiegand, Herbert Ernst; Zgusta, Ladislav (Hg.): *Wörterbücher. Ein internationales Handbuch der Lexikographie*. 3. Teilband – HSK 5.3 (= HSK 5.1, 5.2, 5.3 Handbücher zur Sprach- und Kommunikationswissenschaft 1989, 1990, 1991). – Berlin: De Gruyter, 2877–2881.
- und Wiegand, Herbert Ernst (1989): Component Parts and Structures of General Monolingual Dictionaries: A Survey. In: Hausmann, Franz-Josef; Reichmann, Oskar; Wiegand, Herbert Ernst; Zgusta, Ladislav (Hg.): *Wörterbücher. Ein internationales Handbuch der Lexikographie*. 1. Teilband – HSK 5.1 (= HSK 5.1, 5.2, 5.3 Handbücher zur Sprach- und Kommunikationswissenschaft 1989, 1990, 1991). – Berlin: De Gruyter, 328–359.
- und Werner, Reinhold Otto (1991) Spezifische Bauteile und Strukturen zweisprachiger Wörterbücher. In: Hausmann, Franz-Josef; Reichmann, Oskar; Wiegand, Herbert Ernst; Zgusta, Ladislav (Hg.): *Wörterbücher. Ein internationales Handbuch der Lexikographie*. 3. Teilband – HSK 5.3 (= HSK 5.1, 5.2, 5.3 Handbücher zur Sprach- und Kommunikationswissenschaft 1989, 1990, 1991). – Berlin: De Gruyter, 2729–2769.
- Heidecke, Stefanie (1996): SGML-Tools in the Dictionary-making Process – Experiences with a German-Polish/Polish-German Dictionary. In: Gellerstam, Martin; Järborg, Jerker; Malmgren, Sven-Göran; Norén, Kerstin; Rogström, Lena; Pappmehl, Catarina (Hg.): *Euralex'96 Proceedings*. 7<sup>th</sup> International Congress on Lexicography. – Göteborg, Sweden, 395–404.
- Horn, Robert E. (1989): Mapping Hypertext. The Analysis, Organisation and Display of Knowledge for the Text Generation of On-line-text and Graphics. – Waltham, MA.: Information Mapping Inc.
- Hupka, Werner (1989): Wort und Bild. Illustrationen in Wörterbüchern und Enzyklopädien. – Tübingen: Niemeyer (= Lexicographica, Series Maior 22).
- Kammerer, Matthias (1998): Hypertextualisierung gedruckter Wörterbuchtexte: Verweisstrukturen und Hyperlinks. In: Storrer, Angelika; Harriehausen, Bettina (Hg.): *Hypermedia für Lexikon und Grammatik*. – Tübingen: Narr (= Studien zur deutschen Sprache 12), 145–171.
- Kessler, Markus; Freisler, Stefan (1995): Document Engineering. In: *Materialien aus der Herbsttagung der Tekom e.V. im November 1995 in Dortmund*. (30.3.1999) <http://www.schema.de/html-deu/schemapu/vortrag/document.htm>.

- Krishnamurthy, Ramesh (1996): The Data is The Dictionary: Corpus at the Cutting Edge of Lexicography. In: Kiefer, Ferenc; Kiss, Gábor; Pajzs Júlia (Hg.): Proceedings of the 4<sup>th</sup> International Conference on Computational Lexicography COMPLEX'96. – Budapest, Hungary, 117–144.
- Kuhlen, Reiner (1991): Hypertext. Ein nicht-lineares Medium zwischen Buch und Wissensbank. – Springer, Edition SEL-Stiftung.
- Lowe, John B.; Baker, Collin F.; Fillmore, Charles J. (1997): A Frame-Semantic Approach to Semantic Annotation. In: Proceedings of the SIGLEX Workshop „Tagging Text with Lexical Semantics: Why, What and How?“ (in conjunction with ANLP-97), 4–5 April 1997. Washington, D.C. HTML-Datei: (24.09.1999). URL: <http://www.icsi.berkeley.edu/~framenet/docs/siglex.html>, PostScript-Datei: <http://www.icsi.berkeley.edu/~framenet/docs/siglex.ps>.
- Martin, Willy (1995): Maschinelle Lexikographie. Ein Blick in die Zukunft. In: Hitzemberger, Ludwig (Hg.): Angewandte Computerlinguistik. – Hildesheim: Olms (= Sprache und Computer 15, Vorträge aus der Jahrestagung der Gesellschaft für Linguistische Datenverarbeitung e.V., Regensburg, 30.–31.3.1995), 1–21.
- Neth, Hansjörg; Swanson, Heidi (1999): Zwei Sprachen, eine Scheibe. Englischwörterbücher auf CD-ROM. In: c't. Magazin für Computertechnik. 2/99, 100–106.
- Patric, Jon; Zhang, Jun; Zubillaga, Artola Xabier (1996): An Architecture for a Federation of Heterogeneous Lexical and Dictionary Databases. In: Proceedings of the 1996 Joint Conference of The Association for Literary and Linguistic Computing and The Association for Computers and the Humanities (ALLC/ACH). – Bergen, Norway. URL: <http://fims-www.massey.ac.nz/~is> (25.3.1999).
- Petelenz, Krzysztof (1998): Lexicon ex machina. Wie können wir elektronische Maschinen für die multimediale Wörterbucharbeit nutzen? In: Orbis Linguarum Vol. 8, 147–158.
- (2000): Zur Hypertextualisierung von zweisprachigen Wörterbüchern. Einige Vorschläge am Beispiel des Sprachenpaares Deutsch-Polnisch. In: Wiegand, Herbert Ernst (Hg.): Wörterbücher in der Diskussion IV. Vorträge aus dem Heidelberger Lexikographischen Kolloquium. – Tübingen: Niemeyer (= Lexicographica, Series Maior 100), 203–231.
- Piotrowski, Tadeusz (1999): Tagging and Conversion of a Bilingual Dictionary for XeLDA, a Xerox Computer Assisted Translation System. In: Kiefer, Ferenc; Kiss, Gábor; Pajzs Júlia (Hg.): Proceedings of the 5<sup>th</sup> International Conference on Computational Lexicography COMPLEX'99. – Budapest, Hungary, 113–120.
- Storror, Angelika (1998): Hypermedia-Wörterbücher: Perspektiven für eine neue Generation elektronischer Wörterbücher. In: Wiegand, Herbert Ernst (Hg.): Wörterbücher in der Diskussion III. Vorträge aus dem Heidelberger Lexikographischen Kolloquium. – Tübingen: Niemeyer (= Lexicographica Series Maior 84), 107–134.
- Tarp, Sven (1995): Wörterbuchfunktionen: Utopische und realistische Vorschläge für die bilinguale Lexikographie. In: Wiegand, Herbert Ernst (Hg.): Studien zur zweisprachigen Lexikographie mit Deutsch II. – Hildesheim: Olms (= Germanistische Linguistik 127–128), 17–62.
- Thielen, Christine; Breidt, Elisabeth; Feldweg, Helmut (1998): COMPASS – Ein intelligentes Wörterbuchsystem für das Lesen fremdsprachiger Texte. In: Storror, Angelika; Harrihausen, Bettina (Hg.): Hypermedia für Lexikon und Grammatik. – Tübingen: Narr (= Studien zur deutschen Sprache 12), 173–194.
- Wegner, Immo (1985): Frame-Theorie in der Lexikographie. Untersuchungen zur theoretischen Fundierung und computergestützter Anwendung kontextueller Rahmenstrukturen für die Lexikographische Repräsentation von Substantiven. – Tübingen: Niemeyer (= Lexicographica, Series Maior 10).
- Weidenmann, Bernd (1997): Multicodierung und Multimodalität im Lernprozeß. In: Issing, Ludwig; Klisma, Paul (Hg.): Information und Lernen mit Multimedia. – Weinheim: Beltz, 65–81.
- Wiegand, Herbert Ernst (1986): Shanghai bei Nacht. Auszüge aus einem metalexikographischen Tagebuch zur Arbeit beim Großen Deutsch-Chinesischen Wörterbuch. In: Wiegand, Herbert Ernst (Hg.): Studien zur neuhochdeutschen Lexikographie VI. – Hildesheim: Olms (= Germanistische Linguistik 87–90), 522–626.

- (1996a): Über die Mediostrukturen bei gedruckten Wörterbüchern. In: Zettersten, Arne; Pedersen, Vigo Hjørnager (Hg.): Proceedings of the Symposium on Lexicography VII, 5.-6. May 1994 at the University of Copenhagen. – Tübingen: Niemeyer (= Lexicographica, Series Maior 76), 11–43.
  - (1996b): Das Konzept der semiintegrierten Mikrostrukturen. Ein Beitrag zur Theorie zweisprachiger Printwörterbücher. In: Wiegand, Herbert Ernst (Hg.): Wörterbücher in der Diskussion II. Vorträge aus dem Heidelberger Lexikographischen Kolloquium. – Tübingen: Niemeyer (= Lexicographica, Series Maior 70), 1–82.
- Williams, Melinda (1997): Document Engineering: A Unique Design and Modeling Discipline. (30.3.1999), <http://www.goinside.com/97/2/doceng.html>.
- Worbs, Erika (1997): Plädoyer für das zweisprachige Wörterbuch als Hilfsmittel des Translators. In: Drescher, Horst W. (Hg.): Transfer. Übersetzen-Dolmetschen-Interkulturalität. – Frankfurt a. M.: Peter Lang, 497–510.

*Krzysztof Petelenz, Stuttgart*