

Erschienen in: Rehm, Georg/Witt, Andreas/Lemnitzer, Lothar (Hrsg.):  
Datenstrukturen für linguistische Ressourcen und ihre Anwendungen.  
Proceedings of the Biennial GLDV Conference 2007. Tübingen: Narr, 2007,  
S. 179-188.

# Das *ellexiko*-Portal: Ein neuer Zugang zu lexikografischen Arbeiten am Institut für Deutsche Sprache\*

Carolin Müller-Spitzer

## Zusammenfassung

Das *ellexiko*-Portal soll verschiedene lexikografische Projekte des IDS in einem Verbund zusammenführen und – soweit das die Inhalte zulassen – gemeinsame Recherchemöglichkeiten über verschiedene lexikografische Produkte hinweg bieten. In diesem Aufsatz geht es v. a. darum zu zeigen, wie die XML-basierte Modellierung für das Portal aufgebaut ist, um zum einen die Basis für diese flexiblen Zugriffsstrukturen zu legen und zum anderen der Verschiedenheit der beteiligten Projekte Rechnung zu tragen. Gleichzeitig werden Perspektiven für eine flexiblere Darstellung der Daten und für die zukünftige Weiterentwicklung von Recherchemöglichkeiten aufgezeigt.

## 1 Das *ellexiko*-Portal

Das Projekt *ellexiko* des Instituts für Deutsche Sprache (IDS) ist zweigeteilt: Auf der einen Seite wird im Projekt ein Wörterbuch zum Gegenwartswortschatz des Deutschen neu und korpusbasiert erarbeitet (im Folgenden *ellexiko*-Wörterbuch genannt), zum anderen soll *ellexiko* auch als Portal für weitere lexikologisch-lexikografische Arbeiten am IDS dienen (im Folgenden *ellexiko*-Portal genannt). Was ist dabei mit der Bezeichnung *ellexiko*-Portal gemeint?

Das zehnbändige „Große Wörterbuch der deutschen Sprache“ des Dudenverlags aus dem Jahr 2000 beschreibt das Wort „Portal“ in seiner allgemeinen Verwendung folgendermaßen:

**Por|tal**, das; -s, -e [spätmhd. portal < mlat. portale = Vorhalle, zu: portalis = zum Tor gehörig, zu lat. porta, Pforte]: **1.** *baulich hervorgehobener, repräsentativ gestalteter größerer Eingang an einem Gebäude*: ein hohes, breites, steinernes P.; das P. einer Kirche; aus dem P., durch das P. treten; Unten, vor dem P. des Polizeihochhauses, sah er die erleuchteten Schilder der U-Bahn und der S-Bahn (Prodöhl, Tod 91). **2.** (Technik) *(fest stehende od. fahrbare) tor-, portalartige Tragkonstruktion (für eine bestimmte Art von Kränen)*. **3.** (EDV) *(meist mit Werbung gestaltete) erste Seite eines Browsers o. Ä.* © 2000 Dudenverlag<sup>1</sup>

\* Erschienen in: *Datenstrukturen für linguistische Ressourcen und ihre Anwendungen*, Georg Rehm, Andreas Witt, Lothar Lemnitzer (Hrsg.), Tübingen: Gunter Narr Verlag, 2007. Seite 179–188.

<sup>1</sup> Duden – Das große Wörterbuch der deutschen Sprache, hg. vom Wissensch. Rat der Dudenredaktion. Die CD-ROM basiert auf der 3., völlig neu bearbeiteten und erweiterten Auflage der Buchausgabe in 10 Bänden (1999).

Da es sich beim *ellexiko*-Portal um ein elektronisches, im Internet verfügbares Produkt handelt, liegt es nahe, die unter dem Bedeutungspunkt 3 angedeutete Einzelbedeutung weiter zu vertiefen. In Wikipedia heißt es im Artikel „Portal“<sup>2</sup> speziell aus dem Bereich der Informatik: „Der Ausdruck Portal (lat. porta, „Pforte“) bezeichnet in der Informatik ein Anwendungssystem, das durch folgende Eigenschaften gekennzeichnet ist:

- Integration von Anwendungen, Prozessen und Diensten
- Bereitstellung von Funktionen zur:
  - Personalisierung, Sicherheit, Navigation und Benutzerverwaltung
  - Suche und Präsentation von Informationen.“

Weiterführend wird ein Portal, vor allem für den Bereich Unternehmenskommunikation, von Kirchhof et al. folgendermaßen in seiner Verwendung präzisiert: „Ein Portal ist definiert als eine Applikation, welche basierend auf Webtechnologien einen zentralen Zugriff auf personalisierte Inhalte sowie bedarfsgerecht auf Prozesse bereitstellt. Charakterisierend für Portale ist die Verknüpfung und der Datenaustausch zwischen heterogenen Anwendungen über eine Portalplattform. Eine manuelle Anmeldung an den in das Portal integrierten Anwendungen ist durch Single-Sign-On nicht mehr notwendig, es gibt einen zentralen Zugriff über eine homogene Benutzungsoberfläche. Portale bieten die Möglichkeit, Prozesse und Zusammenarbeit innerhalb heterogener Gruppen zu unterstützen.“ (Kirchhof et al., 2004, S. 5).

Das *ellexiko*-Portal ist dabei in folgender Hinsicht ein Portal: Es bietet – wie ein Eingang zu einem größeren Gebäude – den Zugriff auf verschiedene lexikografische Inhalte aus unterschiedlichen Projekten. Es soll damit die gemeinsame Suche und Präsentation von Informationen mit einem zentralen Zugriff über eine homogene Benutzungsoberfläche verbinden. Dies sind Eigenschaften des Portals, die von außen, d. h. von potenziellen Benutzern, wahrzunehmen sind. Allerdings spielt der Gedanke des Portals im Fall von *ellexiko* auch eine wichtige Rolle im lexikografischen Prozess, denn die Datenbasen der einzelnen Projekte sollen nach einheitlichen Prinzipien strukturiert werden, damit das Portal mehr als ein Nebeneinander einzelner elektronischer Wörterbücher sein kann. Im Moment sind neben den Inhalten des *ellexiko*-Wörterbuchs die Ergebnisse des Projekts *Neologismenforschung*<sup>3</sup> integriert worden. Resultate weiterer Projekte – wie solche aus *Usuelle Wortverbindungen*<sup>4</sup> oder *Zeitreflexion 1945–55*<sup>5</sup> – sollen im ersten Halbjahr 2007 folgen. Die Vereinigung verschiedener lexikografischer Arbeiten unter einem informationstechnologischen Dach bringt Herausforderungen verschiedenster Art mit sich. Dabei stellen sich Fragen wie:

- Inwiefern können und sollen die Inhalte der einzelnen Projekte aufeinander abgestimmt werden?
- Wie werden die Inhalte modelliert und strukturiert?
- Wie können gemeinsame Zugriffsmöglichkeiten auf die Daten entwickelt werden?

Bei den verschiedenen im *ellexiko*-Portal gezeigten Arbeiten handelt es sich um inhaltlich voneinander unabhängige Projekte. Trotzdem ist es natürlich für den Wert des *ellexiko*-Portals

<sup>2</sup> Siehe [http://de.wikipedia.org/wiki/Portal\\_\(Informatik\)](http://de.wikipedia.org/wiki/Portal_(Informatik)).

<sup>3</sup> Siehe <http://www.ids-mannheim.de/11/Neologie/> sowie Herberg et al. (2004).

<sup>4</sup> Siehe <http://www.ids-mannheim.de/kl/projekte/uwv/>.

<sup>5</sup> Siehe <http://www.ids-mannheim.de/11/Zeitreflexion/>.

für potenzielle Benutzer entscheidend, möglichst viele gemeinsame Zugriffsstrukturen auf die jeweiligen Inhalte zu entwickeln, d. h. mehr als eine reine Zusammenstellung inhaltlich nicht verbundener Wörterbücher bzw. lexikografischer Ressourcen zu bieten. Die miteinander in Einklang zu bringenden Pole sind daher zum einen die Eigenständigkeit der Projekte zu wahren, zum anderen die Integrierbarkeit der Daten zu gewährleisten. Daher wurde auf Ebene der Datenmodellierung, die die Basis für alle weiteren Möglichkeiten im Umgang mit den Daten darstellt – „What you mark is what you get“ (Glöning und Welter, 2001, S. 128) –, viel Zeit darin investiert, die Mikro- bzw. Inhaltsstrukturen der einzelnen Wörterbücher so nach einem einheitlichen Konzept zu modellieren, dass

1. gleiche Phänomene auch gleich strukturiert sind und
2. alle Daten so granular strukturiert sind, dass sehr gezielt auf sie zugegriffen werden kann.

Dafür wurde auf Basis eines neuen Konzepts für die maßgeschneiderte Modellierung lexikografischer Daten (Müller-Spitzer, 2007) ein komplexes XML-DTD-Modulsystem entwickelt, mit dem die oben genannten verschiedenen Pole möglichst gut in Einklang gebracht werden. Die Bausteine dieser DTD-Bibliothek sollen im Folgenden vorgestellt werden. Im Anschluss werden die Präsentation der Daten im elexiko-Portal kurz skizziert und weitere Desiderata die Präsentation betreffend formuliert.

## 2 Aufbau der Modellierung

Unter dem Dach des elexiko-Portals sind verschiedene lexikografische Projekte des IDS vereint. Dabei wurde versucht, die Art der Modellierung auf Ebene der Inhaltsstrukturen, d. h. auf Ebene der einzelnen lexikografischen Angaben, aufeinander abzustimmen, und nicht unabhängig voneinander die einzelnen lexikografischen Prozesse unterschiedlich zu organisieren. Damit stellt das elexiko-Portal langfristig eine andere Art von Portal dar, als zum Beispiel das Wörterbuch-Portal, das vorwiegend eine Einstiegsseite für verschiedene Wörterbücher darstellt.<sup>6</sup> Auch eine gemeinsame Einstiegsseite auf sehr unterschiedliche lexikografische Projekte ist sehr sinnvoll, doch da im elexiko-Portal im Moment nur Projekte des IDS vertreten sind, bestand hier die Chance, die Inhalte von Anfang an tiefergehend zu vernetzen und damit die Basis für sehr flexible Zugriffsmöglichkeiten zu legen. Denn sehr heterogene Datenbasen können zwar über Metastrukturen wie semantische Netze erschlossen werden, ohne die Datenbasis zu verändern, aber je nachdem, wie die einzelnen lexikografischen Inhalte strukturiert und ausgezeichnet sind, sind die dabei entstehenden Schwierigkeiten nicht zu vernachlässigen. Schon verschiedene Stichwortlisten unterschiedlicher Wörterbücher zusammenzubringen, kann – insbesondere bei historischen Wörterbüchern – eine nur mit viel Aufwand zu bewältigende Herausforderung sein, da die Prinzipien zum Lemmaansatz von Wörterbuch zu Wörterbuch differieren. Daneben soll die gleichartige Organisation der einzelnen lexikografischen Prozesse im elexiko-Portal natürlich auch Synergieeffekte hinsichtlich des geforderten Know-hows, der eingesetzten Software etc. erzielen.

<sup>6</sup> Das Wörterbuch-Portal wird betreut von der Berlin-Brandenburgischen Akademie der Wissenschaften und von der Heidelberger Akademie der Wissenschaften, siehe <http://www.woerterbuch-portal.de>.

Die Modellierung erfolgt daher für alle Projekte im Prinzip gleich: Sie ist XML-basiert, um langlebige Datenhaltung zu gewährleisten (Büchel und Schröder, 2001), die Daten sind sehr granular strukturiert, um Flexibilität hinsichtlich der Darstellung und des Zugriffs zu erreichen (Wiegand, 1998, Müller-Spitzer, 2006) und die Modellierung ist so angelegt, dass sie die Lexikografen beim Eingeben der Artikel möglichst gut in der Einhaltung der Artikelstruktur unterstützt (Wiegand, 1998, Müller-Spitzer, 2007).<sup>7</sup> Die im XML-Editor erstellten Artikel werden in einer Oracle-Datenbank verwaltet; die Darstellung erfolgt über XSLT-Stylesheets. Hier soll nun die Organisation der XML-Modellierung im Vordergrund stehen. Denn das *elexiko*-Portal ist im Bereich der wissenschaftlichen Lexikografie ein seltenes Beispiel dafür, dass unterschiedliche Projekte auf Ebene der Datenbasis in der Weise verbunden und gleichartig modelliert werden.

Anfangs wurde gesagt, dass vom Inhalt her gleiche lexikografische Angaben aus den verschiedenen Wörterbüchern möglichst auch gleich strukturiert werden sollen, damit die Basis für einen gemeinsamen Zugriff gelegt wird. Aus diesem Grund wurde eine DTD-Bibliothek für das *elexiko*-Portal erstellt, in der in einzelnen DTDs die Entitäts, Elemente oder Attribute zusammengefasst sind, die in verschiedenen Artikelstrukturen gleich sind. Beispielsweise gibt es zwei DTDs für allgemein verwendete Entitäts bzw. Elemente und Attribute, die in allen im *elexiko*-Portal eingebundenen Projekten verwendet werden (Tabelle 1, S. 184). Dabei handelt es sich um so allgemeine Objekte wie IDs zur Identifikation von Elementen oder die Modellierung von Listen, aber auch um inhaltlich wichtige Angaben, die in allen Strukturen gleich eingesetzt sind, wie die Arten von Kommentaren, verschiedene Arten von Hinweisen oder Belege. Unterhalb dieser ganz allgemeinen Ebene, die in allen (Teil-)Produkten des *elexiko*-Portals verwendet werden, gibt es in der DTD-Bibliothek Objektgruppen, die beispielsweise im *elexiko*- und im Neologismenwörterbuch gleich sind. Auch diese sind wieder in einzelnen DTDs zusammengefasst, wie zum Beispiel in der *ew1-objekte.dtd* alle Objekte für Einwortlemmata, die sowohl im *elexiko*- wie im Neologismenwörterbuch identisch sind oder in der *ew1-grammatik.dtd* alle Objekte zur Einwortlemma-Grammatik. Dort finden sich z. B. alle Angaben aus dem Bereich der Wortbildung oder der Deklination und Valenz, die im *elexiko*- und im Neologismenwörterbuch gleich strukturiert sind. Durch diese Aufteilung ist schon anhand der Modellierung zu sehen, dass diese Angaben damit auch gemeinsam zugreifbar sind. Beispielsweise könnte in beiden Wörterbüchern gleichzeitig nach allen starken Verben gesucht werden, die ein obligatorisches Objekt-Komplement haben, was als Nominalphrase im Akkusativ realisiert ist. Zur Verdeutlichung stellt Listing 1 (S. 185) einen Ausschnitt aus der XML-Instanz des Stichworts *abmahnen* aus dem *elexiko*-Wörterbuch dar, in dem die Angaben zur Valenz zu sehen sind.

Die gesamte DTD-Bibliothek für das *elexiko*-Portal gliedert sich damit in Bausteine für das gesamte Portal, in Bausteine für modulübergreifende Objektgruppen, in Bausteine für einzelne Modulprodukte bzw. -projekte und schließlich Kopf-DTDs für die einzelnen Artikelstrukturen, in denen die einzelnen relevanten Teile der DTD-Bibliothek zusammengezogen werden.<sup>8</sup>

Zur Veranschaulichung wie das Zusammenspiel der DTD-Bibliothek funktioniert, folgend die einzelnen Teile der DTD-Bibliothek, die bei der Struktur für die Einwortlemmata im *elexiko*-Wörterbuch (d. h. in der *elexiko-ew1.dtd*) referenziert werden:

<sup>7</sup> Im Projekt werden nach wie vor XML-DTDs eingesetzt, da sie von allen Projektmitgliedern besser lesbar sind (als Schemata). Für das Einlesen in die Oracle-Datenbank werden sie automatisch in XML-Schemata umgewandelt.

<sup>8</sup> Die einzelnen am *elexiko*-Portal beteiligten Projekte werden auch als Modulprojekte bezeichnet.

- allg-entities.dtd
- allg-elemente.dtd
- elexiko-allgobj.dtd
- ewl\_mwl-objekte.dtd
- ewl-objekte.dtd
- ewl-grammatik.dtd

In Summe besteht die Struktur für zu bearbeitende Einwortlemmata im elexiko-Wörterbuch damit aus über 400 Elementen und dazugehörigen Attributen. In der `elexiko-ewl.dtd` sind dabei nur die Elemente und Attribute enthalten, die in genau der Form nur für diese Objektart verwendet werden. Im Falle der Einwortlemmata (elexiko-Wörterbuch) sind das beispielsweise die Angaben zur Disambiguierung der einzelnen Lesarten oder die Modellierung der typischen Verwendungsmuster. Dagegen wird die Modellierung der Wortbildungsangaben oder die Strukturierung der Kommentare und Hinweise über Bausteine der DTD-Bibliothek geregelt, da sie auch in anderen Artikelstrukturen vorkommen. Durch diesen modularen Aufbau der DTDs ist beispielsweise für die Entwicklung von Suchfunktionen klar erkennbar, welche Angabearten nur für die Einwortlemmata im elexiko-Wörterbuch und welche auch für andere Wörterbücher gelten. Für die Lexikografen spielt es dabei bei der Artikelerarbeitung keine Rolle, in welchem Teil der DTD-Bibliothek ein einzelnes Element definiert ist. Sowohl beim Arbeiten im Editor als auch in den einzelnen DTD-Dokumentationen, die bei Fragen zur Strukturierung zur Hilfe genommen werden können, sind alle einzelnen Artikelstrukturen integrativ zu bearbeiten bzw. zu durchsuchen.<sup>9</sup>

Wenn man gleiche Angaben aus verschiedenen lexikografischen Projekte wie dem elexiko- und dem Neologismen-Wörterbuch auf Ebene der einzelnen lexikografischen Angaben gleich strukturieren will und diese Gleichheit auch auf Ebene der XML-Modellierung abbilden will, führt das zu einem relativ komplexen Modellierungsaufbau. Das heißt natürlich auch, dass ein gewisser Aufwand zur Erstellung, aber auch zur Pflege der DTDs aufgewendet werden muss. Was rechtfertigt diesen Aufwand?

Das Ziel des elexiko-Portals ist es, langfristig wirklich neuartige Zugänge zu lexikografischen Daten aus verschiedenen Wörterbüchern zu bieten, wie das die Wörterbuchforschung schon lange fordert (Schryver, 2003). Dabei ist sowohl an eine ausführliche und kombinierbare Suche nach bestimmten Eigenschaften der bearbeiteten Stichwörter über verschiedene Wörterbücher hinweg gedacht, als auch an benutzeradaptive Zugänge zu den lexikografischen Daten (z. B. unterschieden nach Experten und Laien) und an neuartige Formen der Vernetzung. Wenn man dieses Ziel verwirklichen will, müssen die lexikografischen Daten granular und gezielt zugreifbar sein. Da im elexiko-Portal der besondere Fall vorliegt, dass es sich bei den beteiligten (Teil-)Wörterbüchern um Projekte handelt, die in einem Haus erstellt werden, und dass die Modellierung für alle Projekte von einer Person erstellt wird, lag es nahe, die Strukturierung der Daten von vorneherein aufeinander abzustimmen.<sup>10</sup> Darüber hinaus sprachen folgende Gründe dafür, eine solche Modellierung einzusetzen: Die Lexikografen werden sehr gut durch die Artikelstruktur geführt, sodass die Daten sehr konsistent sind, zumindest so weit man das formal prüfen kann. Dies erleichtert auch wesentlich die Einarbeitung neuer Mitarbeiter. Auch

<sup>9</sup> Die DTD-Dokumentationen sind mithilfe der Software DTDhelp erstellt und ermöglichen eine Navigation durch die XML-Struktur und sind – besonders für das elexiko-Wörterbuch – angereichert mit Informationen zur Modellierung und redaktionellen Richtlinien.

<sup>10</sup> Damit ist die Ausgangssituation eine völlig andere als beispielsweise für das Wörterbuchnetz des „Kompetenzentrums für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften der Universität Trier“, siehe <http://germazope.uni-trier.de/Projects/WBB/woerterbuecher/>.

DTD-Bibliothek für das *lexiko*-Portal

<i>Bausteine für das gesamte lexiko-Portal</i>	<i>allg-entites.dtd</i> DTD für allgemeine Entites	<i>allg-elemente.dtd</i> DTD für allgemeine Elemente und Attribute	
<i>Bausteine für modulier-gesjende Objektsruppen</i>	<i>ewl-objekte.dtd</i> DTD für Einwortlemma-Objekte	<i>ewl_mvj-objekte.dtd</i> DTD für Einwort- und Mehrwortlemma-Objekte	<i>ewl-grammatik.dtd</i> DTD für Grammatik der Einwortlemmata
<i>Bausteine für Modulprojekte bzw. -projekte</i>	<i>lexikoba-allgobj.dtd</i> DTD für allgemeine Objekte für bearbeitete Artikel im <i>lexiko</i> -Wörterbuch	<i>neo-allgobj.dtd</i> DTD für allgemeine Objekte für Artikel im Neologismen-Wörterbuch	
<i>Kopf-DDTs für Modulprojekte bzw. -projekte</i>	<i>lexikoba-ewl.dtd</i> DTD für EWL-Artikel mit lerni-lautomatisch erstellten Angaben im <i>lexiko</i> -Wörterbuch	<i>neo-mwl.dtd</i> DTD für Einwortlemmata im Neologismen-Wörterbuch	<i>mwl.dtd</i> DTD für Artikel zu Mehrwortlemmata aus dem Projekt Usuelle Wortverbindungen
	<i>lexikoba-ewl.dtd</i> DTD für zu bearbeitende EWL-Artikel im <i>lexiko</i> -Wörterbuch	<i>neo-mwl.dtd</i> DTD für Mehrwortlemmata im Neologismen-Wörterbuch	<i>wv.dtd</i> DTD für Wortverbindungen aus dem Projekt Usuelle Wortverbindungen

Tabelle 1: Übersicht über die einzelnen Ebenen des DTD-Modulsystems für das *lexiko*-Portal

```

<vb-valenz>
  <vb-komplemente>
    <subjekt-komp>
      <nom-nominalphrase>
        <phrasenmusterA>jemand mahnt (jemanden/etwas)
          ab</phrasenmusterA>
      </nom-nominalphrase>
    </subjekt-komp>
    <objekt-komp>
      <akk-nominalphrase obligatorisch="ja">
        <phrasenmusterA>jemand mahnt jemanden/etwas ab</phrasenmusterA>
      </akk-nominalphrase>
      <ppraepositionalphrase obligatorisch="nein">
        <pp-phrasenmusterA>jemand mahnt (jemanden/etwas)
          <pp-praeposition>bei</pp-praeposition> etwas ab
        </pp-phrasenmusterA>
        [... Beleg]
      </ppraepositionalphrase>
      <ppraepositionalphrase obligatorisch="nein">
        <pp-phrasenmusterA>jemand mahnt (jemanden/etwas)
          <pp-praeposition>wegen</pp-praeposition> etwas ab
        </pp-phrasenmusterA>
        [... Beleg]
      </ppraepositionalphrase>
    </objekt-komp>
  </vb-komplemente>
</vb-valenz>

```

Listing 1: Angaben zur Valenz zum Stichwort *abmahnen*

die Darstellung per XSLT-Stylesheets ist durch die klare und kleinteilige Auszeichnung der Daten sehr flexibel. Dies ist deshalb wichtig, da benutzeradaptive Zugänge zu den erarbeiteten lexikografischen Daten – verbunden mit einer unterschiedlichen Darstellung – nur über eine solche feine Auszeichnung der Datenbasis selbst erreicht werden können und nicht über die Erschließung durch eine Metastruktur. Der Aufwand für die Modellierung rechnet sich im Moment jedoch vor allem im lexikografischen Prozess. Auf Seite der Präsentation ist dagegen noch viel zu tun, damit das Potenzial der Modellierung auch für die Benutzer erkennbar wird.

### 3 Von der Modellierung zur Präsentation der Daten

Die Möglichkeiten für die Darstellung der Daten und auch für die Entwicklung von Recherchemöglichkeiten, die durch die Modellierung bereitgestellt werden, sind sehr vielfältig. Auch ist der modulare Aufbau des elexiko-Portals im Aufbau der DTD-Bibliothek klar abgebildet. Der jetzige Auftritt des elexiko-Portals im Internet zeigt allerdings bislang nur erste Ansätze sowohl dieser Zugriffsmöglichkeiten als auch des modularen Aufbaus des Portals. So gibt es unter <http://www.elexiko.de> beispielsweise erste Möglichkeiten zur gezielten Recherche nach

bestimmten Angabetypen; allerdings ist es hier noch nicht möglich, die gleichen Suchanfragen z. B. gemeinsam über das *lexiko-* wie das *Neologismen-Wörterbuch* laufen zu lassen, obwohl das von der Aufbereitung der Daten und der Art der Inhalte möglich wäre.

Neben der Entwicklung neuartiger Recherchemöglichkeiten sind auch neue Formen der Darstellung nicht zu vernachlässigen. Denn es gilt als nicht besonders benutzerfreundlich, im elektronischen Wörterbuch zwar flexible Zugriffsmöglichkeiten z. B. über eine Suchfunktion oder eine Stichwortliste auf die Wörterbuchartikel zu bieten, aber als Ergebnis nach wie vor den gesamten Wörterbuchartikel zu zeigen. Denn damit reicht die Nachschlagehandlung – wie beim gedruckten Wörterbuch – nur bis zum Artikelanfang. Daher wird in der Wörterbuchforschung gefordert, abhängig von verschiedenen Benutzungssituationen aus einem Wörterbuchartikel nur die jeweils gesuchten Angaben anzuzeigen zu können. „It is one thing to be able to store ever more data, but another thing entirely to present just the data users want to response to a particular look-up.“ (Schryver, 2003, S. 178). „Die Benutzerschnittstelle von digitalen Wörterbüchern“ sollte daher „an Typen von Benutzungssituationen adaptierbar sein.“ (Storrer, 2001, S. 64, ähnlich auch Engelberg und Lemnitzer, 2001, S. 224). Ein wesentlicher Bestandteil, der in der Metalexikografie oft gefordert wird, ist dabei ein für Laien und Experten unterschiedlicher Zugang zu den lexikografischen Informationen.

Die Voraussetzungen für diese flexiblen Zugriffsmöglichkeiten müssen bei der Modellierung geschaffen werden. Dabei stellt sich jedoch die Frage, wann und wie der Benutzer in die Überlegungen zur Modellierung einzubeziehen ist. Bis zu welchem Grad können Daten unabhängig von unterschiedlichen Benutzergruppen und verschiedenen Benutzungssituationen modelliert und erarbeitet werden? Wo sind die Grenzen dieser benutzerunabhängigen Erarbeitung? Denn Storrer sagt zwar zur benutzeradaptiven Darstellung lexikografischer Daten: „Dass es sich bislang nur um Prototypen handelt, liegt weniger daran, dass nicht bekannt wäre, für welche usuellen Benutzungssituationen typischerweise welche Klassen von Angaben relevant werden. Die Ursache liegt vielmehr darin, dass eine kontextadaptive Präsentation lexikalischer Informationen eine linguistisch motivierte und feinkörnige Modellierung der lexikographischen Daten voraussetzt. Eine derartige Modellierung erfordert, wenn sie auf der Grundlage eines gedruckten Wörterbuchs erfolgt, einen relativ hohen Auf- und Nachbearbeitungsaufwand und lässt sich deshalb am schnellsten realisieren, wenn ein digitales Wörterbuch unabhängig von einer vorhandenen Printvorlage konzipiert werden kann.“ (Storrer, 2001, S. 64). Die Modellierung für die am *lexiko-*Portal beteiligten Projekte, vor allem die für das *lexiko-* und *Neologismen-Wörterbuch*, erfüllt diese Forderungen: Sie wurde wegen der breiten Adressierbarkeit nach dem oben in Ansätzen skizzierten Modellierungskonzept (Müller-Spitzer, 2007) benutzerunabhängig vorgenommen. Jede Angabe ist einzeln zugreifbar, sodass nicht die gesamten Wörterbuchdaten verändert werden müssen, um unterschiedliche Benutzergruppen in verschiedenen Benutzungssituationen anzusprechen, sondern es können unterschiedliche Sichten auf den Datenbestand definiert werden. Doch die Aussagen der Wörterbuchforschung darüber, für welchen Benutzungskontext welche Angaben relevant sind, sind vor allem anhand gedruckter Wörterbücher untersucht worden. Für die elektronische Lexikografie ist hier noch grundlegende Forschungsarbeit zu leisten. Vor allem muss gründlicher untersucht werden, wie Benutzer durch ein digitales Wörterbuch navigieren, wie Suchfunktionen benutzt werden etc.

Trotzdem könnten in einem ersten Schritt wirklich bekannte Erkenntnisse über die Benutzung lexikografischer Daten in unterschiedliche Darstellungen umgesetzt werden. Beispielswei-

se wäre es für einen Nichtmuttersprachler in der Benutzungssituation der Textproduktion sehr hilfreich, wenn er im elexiko-Wörterbuch unmittelbar die Angaben zu den Mitspielern und die typischen Verwendungsmuster sähe, für die Textrezeption wäre dagegen ein besserer Überblick über die einzelnen Bedeutungspunkte hilfreich. Diese verschiedenartigen Darstellungen könnten ohne Veränderung der Datenbasis durch unterschiedliche Stylesheets erstellt werden.

Die zukünftige Arbeit am elexiko-Portal wird neben dem Ausbau der Wörterbücher, der Einbindung weiterer lexikografischer Produkte und der damit verbundenen Modellierungsarbeit deshalb vor allem darin bestehen, auf diesem Weg weiter voranzukommen: Die Wörterbuchbenutzungsforschung auf dem Gebiet der elektronischen Lexikografie zu vertiefen und Ergebnisse dieser Forschungsarbeit in neuartige Formen der Darstellung und des Zugriffs umzusetzen. Die Voraussetzungen dafür sind durch die Modellierung geschaffen. Allerdings stehen diese Punkte nicht im Mittelpunkt der alltäglichen Projektarbeit, sodass es bisher an Zeit und Mitteln fehlte, diese Ziele konsequent zu verfolgen. Trotzdem werden wir versuchen, dem anvisierten Ziel – der neuartigen Darstellung lexikografischer Informationen – stückweise näher zu kommen und so die Differenziertheit der Modellierung auch für die Benutzer sichtbar zu machen.

## Literaturverzeichnis

- Büchel, Gregor und Schröder, Bernhard (2001): "Verfahren und Techniken in der computergestützten Lexikographie". In: *Chancen und Perspektiven computergestützter Lexikographie*, herausgegeben von Lemberg, Ingrid; Schröder, Bernhard und Storrer, Angelika, Tübingen: Niemeyer, Nummer 107 in *Lexicographica*. Series Maior, S. 7–28.
- Engelberg, Stefan und Lemnitzer, Lothar (2001): *Lexikographie und Wörterbuchbenutzung*. Nummer 14 in *Stauffenburg Einführungen*. Tübingen: Stauffenburg.
- Gloning, Thomas und Welter, Rüdiger (2001): "Wortschatzarchitektur und elektronische Wörterbücher: Goethes Wortschatz und das Goethe Wörterbuch". In: *Chancen und Perspektiven computergestützter Lexikographie*, herausgegeben von Lemberg, Ingrid; Schröder, Bernhard und Storrer, Angelika, Tübingen: Niemeyer, Nummer 107 in *Lexicographica*. Series Maior, S. 117–132.
- Haß, Ulrike (Herausgeber) (2005): *Grundfragen der elektronischen Lexikographie. elexiko – das Online-Informationssystem zum deutschen Wortschatz*, Band 12 von *Schriften des Instituts für Deutsche Sprache*. Berlin, New York: de Gruyter.
- Herberg, Dieter; Kinne, Michael und Steffens, Doris (2004): *Neuer Wortschatz. Neologismen der 90er Jahre im Deutschen*. Berlin, New York: de Gruyter.
- Kirchhof, Anja; Gurzki, Thorsten; Hinderer, Henning und Vlachakis, Joannis (2004): "Was ist ein Portal? Definition und Einsatz von Unternehmensportalen". <http://www.gurzki.de/publications/>.
- Müller-Spitzer, Carolin (2006): "Das Konzept der Inhaltsstruktur. Ein Ausschnitt aus einer neuen Konzeption für die Modellierung lexikografischer Daten". *OPAL 2*. <http://www.ids-mannheim.de/pub/laufend/opal/>.
- Müller-Spitzer, Carolin (2007): "Der lexikografische Prozess. Konzeption für die Modellierung der Datenbasis". Erscheint in: *Schriften des Instituts für Deutsche Sprache*.

- Schryver, Gilles-Maurice de (2003): "Lexicographer's Dreams in the Electronic-Dictionary Age". *International Journal of Lexicography* 16 (2): S. 143–199.
- Storrer, Angelika (2001): "Digitale Wörterbücher als Hypertexte: Zur Nutzung des Hypertextkonzepts in der Lexikographie". In: *Chancen und Perspektiven computergestützter Lexikographie*, herausgegeben von Lemberg, Ingrid; Schröder, Bernhard und Storrer, Angelika, Tübingen: Niemeyer, Nummer 107 in *Lexicographica*. Series Maior, S. 53–69.
- Wiegand, Herbert Ernst (1998): *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*, Band 1. Berlin, New York: de Gruyter.