

Roman Schneider

Zur Lexikon-Grammatik-Schnittstelle in einem hypermedialen Informationssystem

1. Hypermedia und Datenbanken: Aufgaben der Lexikon-Grammatik-Schnittstelle

Dieser Beitrag beschreibt Konzeption und Umsetzung der Anbindung von lexikalischen Datenbanken an das grammatische Informationssystem GRAMMIS, das seit Mitte 1993 am Institut für deutsche Sprache (IDS) entwickelt wird. Im Rahmen dieses Projekts wird erforscht, wie grammatisches Wissen mit moderner Computertechnik anschaulich dargestellt und verständlich vermittelt werden kann¹.

Eine besondere Herausforderung an Hypermediasysteme² ist die Nutzung der in externen Datenbanken gespeicherten Informationen. Die Verfügbarkeit solcher Datenbestände innerhalb eines Hypermediasystems erbringt in der Regel einen informationellen Mehrwert. Dieser besteht nicht nur darin, daß sich die absolute Menge des abfragbaren Wissens erhöht. Vielmehr kann der informationelle Mehrwert durch Ausführung folgender drei Schritte erzeugt werden: Wissensrekonstruktion, Informationserarbeitung und Informationsaufbereitung. Der Vorgang der Wissensrekonstruktion läßt sich dadurch charakterisieren, daß Wissensobjekte in von Rechnern verarbeitbare Formen überführt werden. Die Datenbestände werden maschinenlesbar gemacht und derart abgespeichert, daß über ein benutzerfreundliches Datenbankfrontend eine gezielte Suche möglich ist. Die Informationserarbeitung umfaßt den methodischen Zugriff auf die entstandenen Ressourcen über verschiedene Arten von Schnittstellen. Mittels formaler Sprachen, komfortabler Selektionsmasken oder sogar natürlicher Sprache werden die gespeicherten Informationen abgefragt, wobei neue Mehrwerte insbesondere durch die hohe Selektionsleistung maschineller Informationssysteme entstehen. Schließlich müssen die abgefragten Informationen aufbereitet, also an die spezifischen Wünsche des Benutzers angepaßt werden. In diesem Zusammenhang wäre etwa eine graphische oder akustische Präsentation der Informationen in vielen Fällen wünschenswert (Kuhlen 1995, S. 80-90).

¹ Zum GRAMMIS-Projekt vgl. auch Storrer und Strecker (in diesem Band).

² Im folgenden verwende ich den Terminus „Hypermedia“ parallel zur traditionellen Bezeichnung „Hypertext“. Mit beiden Begriffen soll die nicht-lineare Darstellung (multimedial realisierter) Einheiten beschrieben werden; vgl. Kuhlen (1991, S. 12-14). Den Terminus „Hypertext“ benutze ich eher im Zusammenhang mit der Diskussion um Probleme bei der entlinearisierten Wissensdarstellung, den Terminus „Hypermedia“ zur Betonung der multimedialen Aspekte des beschriebenen Systems, vgl. Nielsen (1995, S. 5).

Inhaltliche Grundlage des grammatischen Informationssystems GRAMMIS ist die *Grammatik der deutschen Sprache* (Zifonun/Hoffmann/Strecker 1997). Der engen Verbindung zwischen Grammatik und Lexik, die in deren kategorial-grammatischem Ansatz angelegt ist, wird durch die Modellierung von Schnittstellen zwischen externen Datenbanken und Hypertextbasis Rechnung getragen. Solche Schnittstellen ermöglichen die Verknüpfung textueller Einheiten mit lexikalischen Daten, also beispielsweise von den Stellen im Hypertext aus, in denen verschiedene Typen von Verbkomplementen behandelt werden, mit Verben in der Valenzdatenbank, die über ein Komplement dieses Typs verfügen. Der Vorteil und die besondere Attraktivität dieser Schnittstellen für die verschiedenen Benutzergruppen liegen einerseits in der Erweiterung der verfügbaren Informationen um spezielles Wissen (grammatische Eigenschaften, Beispielsätze u.a.). Andererseits besteht mit ihrer Hilfe die Möglichkeit zur Überprüfung etwa der Vollständigkeit oder der Widerspruchsfreiheit von im Hypertext formulierten grammatischen Regeln und Klassifizierungen.

Die Modellierung der an das Informationssystem GRAMMIS angeschlossenen Datenbanken der deutschen Funktionswörter und der Verbvalenz erfolgte auf der Basis des Entity-Relationship (ER)-Ansatzes, der die Abbildung des in der Datenbank zu modellierenden Wirklichkeitsausschnitts auf ein Datenbankmanagementsystem unterstützt. Dieser Ansatz beruht auf der Unterscheidung zwischen Objekten (Entities) und Beziehungen (Relationships) zwischen Objekten (Chen/Knöll 1991). Entitäten können in verschiedene Entitätstypen klassifiziert werden (z.B. Funktionswort, Verb-Monosem, Passivtyp) und werden in einem Entity-Relationship-Diagramm durch rechteckige Kästen dargestellt. Sie haben Eigenschaften (Attribute), deren Werte (z.B. Passivlabel wie *sein*-Passiv, *werden*-Passiv, *bekommen*-Passiv) in einem Kreis notiert werden. Beziehungen zwischen Entitäten (z.B. Passivierbarkeit eines Verb-Monosems durch einen oder mehrere Passivtypen) werden durch eine Raute dargestellt, die durch Linien mit den entsprechenden Entitätstypen verbunden ist. Es gibt 1:1, 1:n und n:m (viele-zu-viele)-Beziehungen.

In einer ersten Phase wurden alle Objekttypen und Beziehungstypen, die für die zu entwerfenden Datenbanken eine Rolle spielen, identifiziert. Zur datenbankunabhängigen Darstellung in sogenannten konzeptuellen Schemata wurden ER-Diagramme verwendet. Aufbauend auf diesen Diagrammen erfolgte dann in der zweiten Phase die Übersetzung in Datenbankschemata für die benutzten Datenbanksysteme.

Im Falle der Datenbank der Funktionswörter fand die praktische Umsetzung mit Hilfe des Autorensystems *Toolbook* statt, im Falle der Valenzdatenbank wurde das relationale Datenbankmanagementsystem *Paradox* verwendet. Im folgenden sollen zunächst die Inhalte sowie die Struktur der beiden genannten Datenbanken beschrieben und anschließend ihre Anbindung an das Hypertextsystem erläutert werden.

2. Informationen in den GRAMMIS-Datenbanken

2.1 Die Datenbank der Funktionswörter

Zu den Funktionswörtern zählen die sprachlichen Elemente, die primär grammatische Bedeutung tragen und im Gegensatz zu den sogenannten Inhaltswörtern nicht Dinge oder Eigenschaften bezeichnen. Sie erfüllen vielmehr spezielle syntaktisch-strukturelle Funktionen. Zu den Funktionswörtern zählen in GRAMMIS die folgenden Wortarten, die jeweils relativ abgeschlossene Inventare haben und mithin vollständig in der Datenbank der Funktionswörter erfaßt werden konnten: Partikeln, Junktoren, Responsive, Präpositionen, Adverbien und Adkopula.

Das zugrundeliegende Datenmodell (vgl. Abb. 1) besteht aus dem einen Objekttyp *Funktionswort*. Dieser wird durch folgende Attribute charakterisiert:

- *Wortform* (offener Wertebereich)
- *zugehörige Wortart* (geschlossener Wertebereich)
- *spezifische Eigenschaften der Wortform*, z.B. bezüglich Rektion oder Stellung im Satz (offener Wertebereich)
- Verwendungsbeispiele bzw. verdeutlichende Beispielsätze (offener Wertebereich)

Im Zuge der praktischen Umsetzung wurden die genannten Attribute mit Hilfe der von *Toolbook* angebotenen Datensatzfelder (*RecordFields*) abgespeichert, wobei die in den Datensatzfeldern *spezifische Eigenschaften* und *Verwendungsbeispiele* gespeicherten Inhalte nach wie vor komplex sind. Dies steht im Widerspruch zu Regeln, die etwa beim Entwurf relationaler Datenbanken dabei helfen, eine Optimierung des Datenmodells zu erreichen. Das relationale Modell ermöglicht den Prozeß der Normalisierung, der u.a. das Ziel hat, durch die Zerlegung komplexer Attribute in atomare Attribute eine unnötige Redundanz relevanter Informationen zu verhindern. Das Attribut *spezifische Eigenschaften der Wortform* ist nicht atomar, da hier mehrere Inhalte vorkommen können (z.B. die semantische Subklasse, Hinweise zur Betonung und zur Funktion; vgl. Abb. 3). Für die geplanten Datenbankabfragen, die ausschließlich über die (atomaren) Attribute *Wortform* und *Wortart* ablaufen, wurde diese Verfahrensweise mit Nichtschlüsselattributen (vgl. Date 1986) allerdings als ausreichend betrachtet.

Um Suchvorgänge in der Datenbank zu beschleunigen, wurden zwei alphabetische Indizes erstellt. Indiziert wurden die Datensatzfelder *Wortform* und *Wortart*. Der Zugriff auf die einzelnen Datensätze ist über diese Indizes möglich und gegenüber einer sequentiellen Suche ohne Index wesentlich schneller und effizienter.

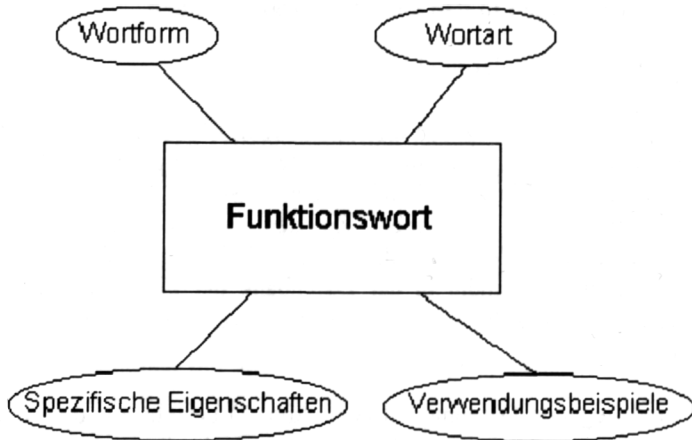


Abb. 1: ER-Diagramm der Datenbank der Funktionswörter

2.2 Die Valenzdatenbank

Im folgenden soll nun das Datenmodell der GRAMMIS-Valenzdatenbank detailliert beschrieben werden, wobei zunächst die Gesamtstruktur und danach die einzelnen Relationen vorzustellen sind. Als Inhalte der Datenbank sollen zunächst die Beschreibungsergebnisse der am Institut für deutsche Sprache in Mannheim angesiedelten Valenzforschungsgruppe – also die Einträge des Valenzwörterbuchs *Verben in Feldern* (Schumacher 1986) – verwendet werden. Das ER-Diagramm der Valenzdatenbank ist in den Abbildungen 2a und 2b dargestellt. Es gibt die folgenden acht grundlegenden Entitätstypen: *Verb-Homonym*, *Verb-Monosem*, *verbspezifische Rolle*, *Konstituententyp*, *semantisches Merkmal*, *Passivtyp*, *Beispielsatz*, *Verbfeld*.

Die Beziehung zwischen *Verb-Homonym* und *Verb-Monosem* ist eine 1:n-Relation, eine Wortform hat also unter Umständen mehrere Bedeutungen. Umgekehrt ist jedes Monosem genau einem Homonym zugeordnet. Der Entitätstyp *Verb-Homonym* besteht aus einer Folge von Buchstaben, nämlich der orthographischen Repräsentation des Verbs. Der Entitätstyp *Verb-Monosem* umfaßt folgende Attribute:

- *Wortbildung* (offener Wertebereich)
- *Phraseologie* (offener Wertebereich)
- *Bedeutungsparaphrase* (offener Wertebereich)
- *Flexion* (offener Wertebereich)
- *Hilfsverb im Perfekt* (fester Wertebereich: *haben* oder *sein*)

- *Anmerkungen* (offener Wertebereich)
- *Quelle* (im Moment fester Wertebereich: *Verben in Feldern*)

Ebenfalls eine 1:n-Beziehung besteht zwischen den Entitätstypen *Verb-Monosem* und *verbspezifische Rolle*: Ein Monosem regiert einen oder mehrere Komplementtypen bzw. Ergänzungsklassen; ein Valenzpartner ist an genau ein Monosem gebunden. Die terminologische Unterscheidung zwischen Komplementtyp (*Grammatik der deutschen Sprache*) bzw. Ergänzungsklasse (*Verben in Feldern*) verdeutlicht ein generelles Problem der Verwaltung von Valenzbeschreibungen unterschiedlicher Herkunft, das in den differierenden Klassifikationen der einzelnen Untersuchungen begründet liegt. Die Einteilung der Verbkomplemente in verschiedene Klassen differiert zwischen der dem Hypertextsystem zugrundeliegenden *Grammatik der deutschen Sprache* und *Verben in Feldern*. Es gibt also keine 1:1-Übereinstimmung in allen Fällen, sondern durchaus Verben, deren Valenzrahmen in den verschiedenen Regelwerken unterschiedlich ist. Um trotzdem die Inhalte sowohl von *Verben in Feldern* als auch anderer Valenzlexika parallel verwalten und nutzen zu können, besitzt der Entitätstyp *verbspezifische Rolle* entsprechend unterschiedliche Attribute. Im einzelnen sind das:

- *Komplementtyp in der Grammatik der deutschen Sprache* (fester Wertebereich)
- *Komplementtyp im Herkunftslexikon* (im Moment nur aus *Verben in Feldern*)
- *Rollenbezeichner* (im Moment: fester Wertebereich entsprechend der *Verben in Feldern*-Variablenbedeutung)
- *Argumentstatus* (fester Wertebereich: *true* oder *false*)
- *Optionalität* (fester Wertebereich: *obligatorisch* oder *fakultativ*)

Eine n:m-Relation besteht zwischen den Entitätstypen *verbspezifische Rolle* und *Konstituententyp*. Valenzpartner können also durch mehrere unterschiedliche Konstituententypen realisiert werden und Konstituententypen können der Realisierung verschiedener Valenzpartner dienen. Der Entitätstyp *Konstituententyp* wird durch die Attribute *Konstituententyp-Label* und *Termklasse* beschrieben, es werden also Informationen über die syntaktische Belegung des Valenzpartners gespeichert.

Eine zweite n:m-Relation findet sich zwischen den Entitätstypen *verbspezifische Rolle* und *semantisches Merkmal*. Valenzpartner können durch verschiedenartige semantische Merkmale belegt werden; semantische Merkmale können der Belegung unterschiedlicher Valenzpartner dienen. Der Entitätstyp *semantisches Merkmal* umfaßt Informationen über die semantische Belegung entsprechend des *Verben in Feldern*-Beschreibungsvokabulars.

Weiterhin angeschlossen an den zentralen Entitätstyp *Verb-Monosem* sind die Entitätstypen *Passivtyp*, *Beispielsatz* und *Verbfeld*. Das ER-Diagramm (Abb. 2a) macht nicht nur die 1:n-Beziehung zwischen einem Verbfeld und einem oder mehreren Verb-Monosemen deutlich, sondern zeigt zusätzlich die in *Verben in Feldern* benutzte hierarchische Struktur von der Art „Feld A ist Subfeld von Feld B“. Auch die Verhältnisse zwischen den Entitätstypen *Passivtyp* mit dem Attribut *Passivlabel* (fester Wertebereich: *werden-*, *sein-* oder *bekommen-Passiv*) und *Beispielsatz* mit den Attributen *Phänomen* und *Text* zum Entitätstyp *Verb-Monosem* sind definiert über 1:n-Beziehungen. Ein Monosemeintrag ist passivierbar durch einen oder mehrere Passivtypen und belegt durch einen oder mehrere Beispielsätze zu einem bestimmten Phänomen.

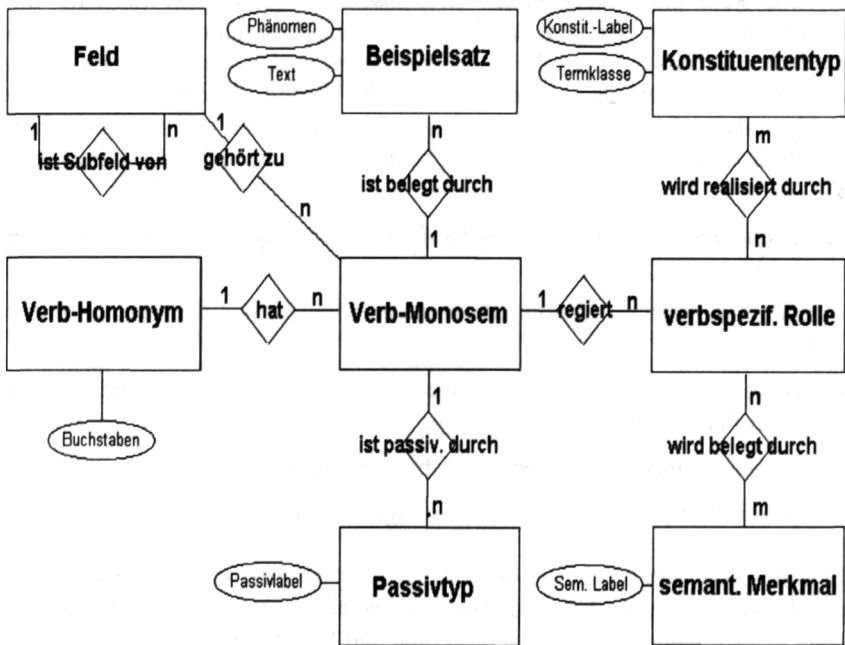


Abb. 2a: ER-Diagramm der Valenzdatenbank

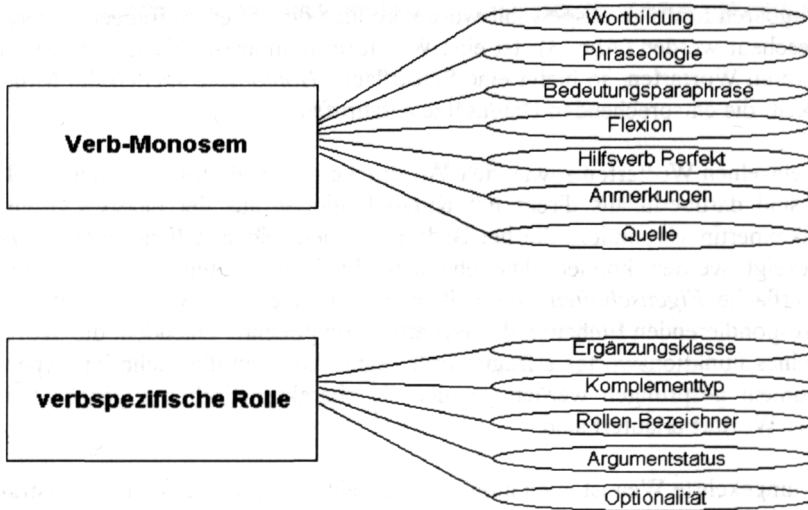


Abb. 2b: Detaillierte Beschreibung der Entitätstypen *Verb-Monosem* und *verbspezifische Rolle*

Zur Verwaltung der im Vergleich mit der Datenbank der Funktionswörter komplexeren Inhalte der Valenzdatenbank bietet sich ein Datenbankmanagementsystem an. Darunter versteht man „den Teil eines (computergestützten) Informationssystems [...], der sich mit der Beschreibung der vorhandenen Daten, ihrer Verwaltung sowie dem Umgang mit und dem Zugriff zu ihnen befaßt“ (Schlageter/Stucky 1983, S. 22). Die Vorteile eines Datenbankmanagementsystems liegen in den durch das System automatisch durchführbaren Konsistenz- und Integritätsprüfungen, hoher Datensicherheit und Redundanzfreiheit, einer integrierten Abfragesprache und der Möglichkeit des Mehrbenutzerzugriffs. Im Falle der GRAMMIS-Valenzdatenbank wird das relationale Datenbanksystem *Paradox* benutzt.

3. Anbindung der GRAMMIS-Datenbanken an den Hypertext

3.1 Anbindung der Datenbank der Funktionswörter

Die Inhalte aus der Datenbank der Funktionswörter können im Informationssystem GRAMMIS von verschiedenen Punkten aus abgerufen werden. Für Benutzer, die sich grammatisches Wissen ausgehend von einer bestimmten Wortform erarbeiten möchten – also die Wortart eines Funktionswortes bestimmen wollen – gibt es die Möglichkeit einer direkten Datenbankabfrage. Nachdem die gesuchte Wortform in ein Suchfeld eingegeben bzw. aus einer Liste der

verfügbaren Einträge ausgewählt wurde, können die ersten Abfrageergebnisse angeschaut werden (Abb. 3). Ist eine Wortform homonym, d.h. gehört sie zu mehreren Wortarten, so bietet eine Schaltfläche *Homonyme suchen* die Möglichkeit, die entsprechenden Datensätze aufzurufen.

Die einzelnen Wortarten sowie viele Ausdrücke innerhalb des Eigenschaftsfeldes sind darüberhinaus direkt mit Kurzdefinitionen aus der GRAMMIS-Grammatikexpertin verbunden, die bei Bedarf in einem eigenen Fenster (*Viewer*) angezeigt werden können. Die ebenfalls für jeden Datensatz vorhandene Schaltfläche *Eigenschaften dieser Wortart* verbindet den Benutzer mit den korrespondierenden Einheiten der Hypertextkomponente. Nachdem die Wortart eines Funktionswortes erfragt wurde, kann direkt an die Stelle im Hyperdokument gesprungen werden, an der die charakteristischen Eigenschaften dieser Wortart erläutert sind.

Der umgekehrte Weg ist auch möglich und stellt die optimale Navigationsstrategie für Benutzer dar, die ausgehend von einem bestimmten thematischen Abschnitt im Hypertext die Datenbank der Funktionswörter nach entsprechenden Einträgen abfragen wollen. Auf jeder Hypertextseite, auf der eine Funktionswortart behandelt wird, befindet sich eine Schaltfläche, die eine vorformulierte Abfrage in der Datenbank durchführt und deren Ergebnisse darstellt. Auf diese Weise lassen sich etwa aus dem Kapitel über die Wortart der Abtönungspartikeln alle in der Datenbank gespeicherten Abtönungspartikeln auflisten und – durch automatische Generierung entsprechender Sprungverbindungen – Verweise in die Datenbankkomponente realisieren. Da das GRAMMIS-Wortartenbuch die Eigenschaften beschreibt, die nach den Definitionen der inhaltlich zugrundeliegenden *Grammatik der deutschen Sprache* für alle Wörter einer Funktionswortart gelten, und in der Datenbank dazu noch Eigenschaften gespeichert sind, die innerhalb der Elemente dieser Wortart variieren, bietet diese Art der Abfrage und Navigation die Möglichkeit, spezifische Zusatzinformationen aus der Lexikonkomponente dem allgemein gehaltenen Regelwerk der Hypertextkomponente anzuschließen.

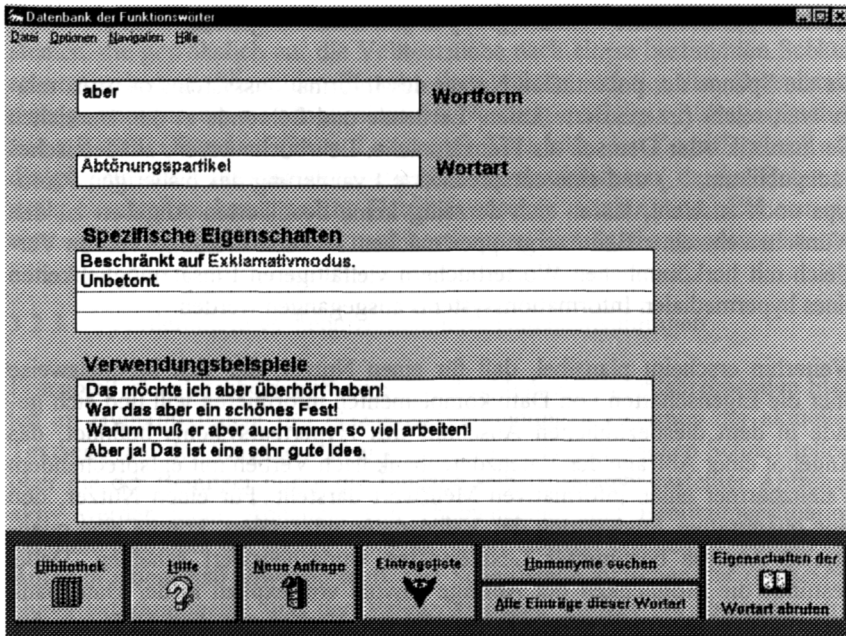


Abb. 3: Benutzeroberfläche der Datenbank der Funktionswörter

3.2 Anbindung der Valenzdatenbank

Die Realisierung einer Verbindung der in der Valenzdatenbank gespeicherten lexikalischen Daten mit dem eigentlichen Hypertext-Informationssystem läßt sich in mehrere zunächst voneinander unabhängige Problemfelder aufteilen. Die zu beantwortenden Fragen sind:

1. Welche Lexikondaten sind für verschiedene Benutzergruppen oder in verschiedenen Benutzungssituationen interessant?
2. Welche Suchtechniken können angewendet werden?
3. An welchen Stellen im Hypertext sollen fest kodierte, vom Autor vordefinierte Abfragen vorgesehen werden bzw. wann eignen sich dynamische, benutzerdefinierte Verbindungen?
4. Wie können die lexikalischen Daten angemessen auf dem Bildschirm dargestellt werden?
5. Welche Teile eines Suchergebnisses lassen sich als Ausgangspunkt für Verweise zurück in den Hypertext nutzen?

3.2.1 Verschiedene Benutzergruppen und Benutzungssituationen

Da die Spanne der potentiellen Nutzer des Informationssystems GRAMMIS relativ weit reicht (interessierte Laien, Lernende und Lehrende in den Bereichen Germanistik oder Deutsch als Fremdsprache, Lehrbuchautoren, akademisches Fachpublikum, ...) und eine entsprechende Evaluierung des bisherigen Prototyps noch aussteht, lassen sich derzeit keine detaillierten Angaben zu den Wünschen einzelner Benutzergruppen machen. Es kann allerdings von im Vergleich mit herkömmlichen Wörterbüchern vielfältigeren Einsatzmöglichkeiten eines hypermedialen Informationssystems ausgegangen werden.

Weiterhin erscheint plausibel, daß für einen Nutzer, der sich beispielsweise über die Eigenschaften von Dativkomplementen informieren will und sich gerade in dem diesbezüglichen Abschnitt des Hypertextsystems befindet, das Angebot einer Abfrage der Valenzdatenbank nach Verben mit entsprechendem Valenzpartner einen informativen Mehrwert darstellt. Für einen Nutzer, den eine quantitative Auswertung der Valenzdaten interessiert, könnte eine quantitative Analyse bestimmter Lexikondaten wünschenswert scheinen. Dabei bestünde die Möglichkeit der Abfrage sowohl von explizit in der Datenbank vorhandenen Informationen (z.B. Anzahl der Verben mit dem Hilfsverb *haben*) als auch von implizit gespeicherten Informationen (z.B. Anzahl der Verben, deren Valenzrahmen zwischen der *Grammatik der deutschen Sprache* und *Verben in Feldern* differiert). Gerade das letztgenannte Beispiel zeigt, daß fortgeschrittene Abfragemöglichkeiten den Wert der Lexikon-Hypertext-Schnittstelle erheblich erhöhen können.

3.2.2 Verwendung unterschiedlicher Suchtechniken

Es gibt verschiedene Suchtechniken in Hypertext- und Datenbanksystemen. Generell lassen sich die Techniken in Inhaltssuche (content search), Struktursuche (structure search) und räumliche Suche (spatial search) unterteilen. Die Inhaltssuche berücksichtigt nicht die Struktur der gespeicherten Informationen, sondern nur deren Inhalt. Die Struktursuche benutzt die Struktur des Informationsnetzwerks, um bestimmte Einträge zu finden. Ein Beispiel dafür wäre die Suche nach Einträgen mit mehr als vier Links. Bei der räumlichen Suche wird die graphische Repräsentation der Einträge untersucht. Als Suchparameter kommt dann die Existenz spezieller Konstituenten und deren Position auf dem Bildschirm zum Einsatz: „Finde alle Einträge mit einem Bild in der linken oberen Fensterecke“ (Kacmar/Leggett/Schnase/Boyle 1988). Daneben existiert die gerade für Hypertextsysteme typische Technik des freien Navigierens und Suchens, die im Gegensatz steht zur üblichen gezielten (rationalen) Recherche (Kuhlen 1991). Interessant für die Suche in der Valenzdatenbank ist in erster Linie die Inhaltssuche, also die Suche nach bestimmten Inhalten von Attributen („Suche alle Einträge, die den Komplement-

typ *Akkusativkomplement* besitzen“). Im Gegensatz zu reinen Hypertextsystemen, die gewöhnlich nur die Volltextsuche nach einem bestimmten Schlüsselwort erlauben, bietet der Einsatz eines echten, relationalen Datenbankmanagementsystems die Möglichkeit der Nutzung mächtiger Abfragesprachen wie z.B. SQL (*Standard Query Language*). Dies hilft, den Zugriffsprozeß auf die gespeicherten Informationen erheblich zu flexibilisieren, da der Benutzer exakt beschreiben kann, welche Daten er ansehen möchte, und nicht nur auf eine wortformenbasierte Suche nach bestimmten Einträgen angewiesen ist.

3.2.3 Vordefinierte vs. benutzerdefinierte Datenbankabfragen

Prinzipiell zu unterscheiden ist zwischen direkten, an den entsprechenden Stellen im Hyperdokument fest plazierten Verbindungen zwischen der aktuellen Textseite und der Valenzdatenbank und zwischen dynamischen Schnittstellen, über die der Benutzer durch die Eingabe selbstzusammengestellter Abfragen aktiv werden kann. Abhängig von Themengebiet und Benutzungssituation kann die eine oder die andere Verbindungsart sinnvoll erscheinen, so daß der Autor in jedem Fall abwägen muß. Die Schnittstelle zwischen der GRAMMIS-Hypertextkomponente und der Valenzdatenbank umfaßt die folgenden drei Verbindungstypen:

3.2.3.1 Verbindung zwischen Themengebiet und Valenzdaten

Von den Stellen des Hyperdokuments aus, an denen verschiedene Typen von Verbkomplementen behandelt werden, können Verben aus der Valenzdatenbank abgerufen werden, die über ein Komplement dieses Typs verfügen. Umgekehrt kann auch von einem in der Datenbank gespeicherten Verb ausgehend zu den Stellen des Hyperdokuments gesprungen werden, an denen die Eigenschaften der Komplemente dieses Verbs näher erläutert sind.

Abb. 4: Verbindung zwischen Themengebiet und Valenzdatenbank

▼ Dativkomplement: Semantische Rollen
<p>Dativkomplemente bezeichnen im allgemeinen einen "indirekt" von einem Ereignis betroffenen Gegenstand, in den meisten Fällen eine Person. Es gibt nur wenige Verben, die auch nicht-personale Belegung des Dativkomplements erlauben. Bei ihnen könnte theoretisch auch satzförmige Realisation sein, belegt sind solche Konstruktionen jedoch äußerst selten.</p>

Eine Verbindung des genannten Typs sieht z.B. folgendermaßen aus: An der Stelle im Hypertext, an der die semantischen Rollen des Dativkomplements behandelt werden, wird beschrieben, daß es nur wenige Verben gibt, die eine nicht-personale Belegung des Dativkomplements erlauben. Über einen typisierten Hyperlink (Abb. 4) ist es dann möglich, alle in der Datenbank gespeicherten Einträge solcher Verben abzurufen. Der Aufruf der Einträge geschieht in diesem Fall – für den Benutzer unsichtbar – über eine vorformulierte Abfrage in SQL, einer nichtprozeduralen Sprache zum Verwalten relationaler Datenbanken. Diese Abfrage wird dem entsprechenden Aktionswort zugeordnet und präsentiert dem Benutzer bei Aktivierung die Abfrageergebnisse auf einer separaten Seite (vgl. Abb. 6).

3.2.3.2 Verbindung zwischen einzelnen Verben und der Datenbank

Oft werden innerhalb eines Hypertextknotens nur ausgewählte Valenzeigenschaften von in Beispielsätzen benutzen Verben erläutert. Eine Anbindung dieser Stellen an die Valenzdatenbank ermöglicht es dem Leser, sämtliche gespeicherten Informationen über das entsprechende Verb abzufragen (Abb. 5) und bei Bedarf durch Verfolgung der im vorigen Punkt erwähnten Navigationspfade an die Stellen im Hyperdokument zu springen, an denen die Eigenschaften weiterer Komplemente des Verbs behandelt werden.

- "Komm, mach **mir** mal die Knöpfe im Rücken auf (...)."
 - Ich kaufe **mir** einen bunten Luftballon.
 - "Ja, Sch Datenbankabfrage als früher, so wie einen Menschen, der **mir** gestorben ist, den ich aber nicht vergessen kann".

Abb. 5: Verbindung zwischen Beispielverb und Valenzdatenbank

Praktisch ist weiterhin folgende Verbindung: Die Hypertextkomponente bietet dem Benutzer die Möglichkeit, die zur Unterscheidung zwischen Komplementen und Supplementen verwendeten Tests praktisch anzuwenden. In diesem Zusammenhang wird dem Benutzer die Möglichkeit geboten, zu Demonstrationzwecken selbstgewählte Verben einzugeben. Falls diese Verben schon in der Datenbankkomponente eingetragen sind, kann das Informationssystem automatisch eine entsprechende Suchanfrage generieren.

3.2.3.3 Benutzerdefinierte Rechercheoptionen in der Datenbank

Die Valenzdatenbank ist wie auch die Datenbank der Funktionswörter aus der GRAMMIS-Bibliothek (zur verwendeten Benutzermetapher vgl. Storrer 1995, S. 297) heraus aufrufbar. Zusätzlich existiert ein Menüeintrag *Abfrage der Va-*

lenzdatenbank, der von jeder Stelle im Hypertext aus aufrufbar ist. Der Anwender gelangt auf diesem Wege auf eine als Frontend für die dahinterliegende Datenbank konzipierte Präsentationsseite (Abb. 6).

Über die Schaltfläche *Abfragegenerator* ist es nun möglich, gezielt nach Belegen für Verben mit bestimmten Eigenschaften zu suchen. Um einerseits vielfältige Retrievaloperationen zu erreichen und andererseits den Benutzer nicht zum Erlernen einer komplexen Abfragesprache zu nötigen, wurde eine möglichst intuitiv verständliche und anwendbare Abfragemaske entworfen (Abb. 7), auf der die gesuchten Attribute einfach per Mausclick ausgewählt und beliebig über *und-* bzw. *oder-*Relationen verknüpft werden können. Ein SQL-Generator übersetzt die ausgefüllte Maske automatisch in die Datenbanksprache SQL, führt die Abfrage durch und läßt den Benutzer anschließend auf der Einstiegsseite durch die gefundenen Einträge blättern. Auf diese Weise können relativ einfach Abfragen durchgeführt werden wie zum Beispiel:

- „Suche alle Verben mit mehr als drei Komplementen.“
- „Suche alle Verben des Besitzens und Besitzwechsels, bei denen das Akkusativkomplement fakultativ ist.“
- „Suche alle Verben, die eine nicht-personale Belegung des Dativkomplements erlauben.“

The screenshot shows a window titled 'Das Valenzbuch' with a navigation pane on the left and a main content area on the right. The navigation pane includes options like 'Ausgang', 'Bibliothek', 'Einführung', 'Abfrage der Valenzdatenbank', 'Abfragegenerator', 'Liste aller Einträge', 'Datenexport in Textdatei', and 'Datenmodell anzeigen'. The main content area is titled 'Abfrage der Valenzdatenbank' and displays the following information for the verb 'abholen':

Verb: abholen
Hilfsverb: haben
Verwendungsbeispiele zu: AdvE
 Da viele Eltern nicht möchten, daß ihre Kinder abends allein nach Hause gehen, holen sie sie selbst von dem Fest ab.

Bedeutung: irgendwohin gehen, um eine Person oder eine Sache in Empfang zu nehmen und wegzubringen
Flexion: holt ab - holte ab - hat abgeholt
Wortbildung: der Abholer/die Abholerin, die Abholung [Vorgang/Handlung]
Quelle: VALBU
Verbleid: unbestimmt
Passivtypen: S-Passiv - W-Passiv

Valenzpartner:

NormE		obl.	+	derjenige, der etwas in Empfang nimmt
AkkE		obl.	+	das, was in Empfang genommen wird
AdvE		fak.	+	Ort, an dem etwas bereitsteht

At the bottom of the window, there are navigation buttons and a page indicator showing '1 von 34'.

Abb. 6: Ergebnispräsentation einer Abfrage der Valenzdatenbank

Mit Hilfe dieser Abfragemaske ist es auch möglich, in der Datenbank nach bestimmten Wortformen zu suchen. Dabei ist es dem Benutzer freigestellt, ob er nach Einträgen für eine einzige Wortform wie z.B. *abholen* suchen will oder durch die Benutzung von Platzhaltern (*wildcards*) die Recherche auf Gruppen von Einträgen erweitern will: Die Suche nach *ab%* findet alle Verben, die mit dem Präfix *ab* beginnen, *ab_* sucht nach Einträgen, die mit *ab* beginnen und insgesamt sieben Buchstaben besitzen.

The screenshot shows the 'Suchkriterien' (Search Criteria) window. It contains the following sections:

- Eintrag:** A text input field for the search term.
- Hilfsverb:** Radio buttons for 'Unspezifiziert' (selected), 'haben', and 'sein'.
- Quelle:** A dropdown menu and a checkbox for 'Auswahl negieren'.
- Feldtyp:** Radio buttons for 'Unspezifiziert' (selected) and 'Spezifiziert'. A text area below shows 'Verben des Helfens und Usatzwechsels' and 'Zustandsverben der allgemeinen Existenz unbestimmt'. A checkbox for 'Auswahl negieren' is also present.
- Passivtyp:** Radio buttons for 'Unspezifiziert' (selected), 'Und - Kombinationen', and 'Oder - Kombinationen'. Checkboxes for 'bekommen - Passiv', 'sein - Passiv', and 'werden - Passiv' are also present.
- Komplementzahl:** Radio buttons for 'Unspezifiziert' (selected) and 'Gensu' (with a value of 1) and 'Valenzpartner'.
- Valenzpartner:** A table with columns for 'Optionalität', 'Arg.-Status', 'Sem. Belegung', and 'Synt. Belegung'. It lists various grammatical categories like AdvE, AkkE, DatE, GenE, NomE, PräpE, and RelI.

Abb. 7: Abfragemaske der Valenzdatenbank

3.2.4 Darstellung der Ergebnisse auf dem Bildschirm

Bei der Darstellung der durch eine Datenbankabfrage gefundenen Informationen sind zwei Problemstellungen zu lösen: die Auswahl der auf einer Seite darzustellenden Informationen sowie die Anordnung der Informationen. Es werden jeweils die Eintragungen zu einem Verb-Monosem dargestellt. Hat eine Abfrage mehrere Verben als Ergebnis, kann der Benutzer mit Hilfe entsprechender Schaltflächen sukzessive alle gefundenen Einträge durchblättern. Die Anordnung der Einträge ist alphabetisch. Generell bestünde aufgrund der zugrundeliegenden Struktur der Valenzdatenbank auch die Möglichkeit – je

denfalls für entsprechend klassifizierte Verben – die Abfolge nach dem onomasiologischen Gesichtspunkt der Wortfeldzugehörigkeit festzulegen, wie es etwa in *Verben in Feldern* gemacht wurde. Da der zur Verfügung stehende Bildschirmplatz durch die Größe des Programmfensters und die zur Navigation benötigten Schaltflächen begrenzt ist und zuviele Angaben auf einen Blick auch der Übersichtlichkeit abträglich sein können, ist die Darstellung nur ausgewählter Angaben zu einem Verb-Monosem sinnvoll.

Zu jedem gefundenen Eintrag werden folgende Informationen aufgeführt: zunächst die Verbform mitsamt dem Hilfsverb, dann die verfügbaren Verwendungsbeispiele und eine Bedeutungsparaphrase, Angaben zu Flexion und Wortbildung sowie Informationen über die Quelle (im Falle von Einträgen aus *Verben in Feldern* komplettiert durch das entsprechende Verbfeld), die Flexion, die möglichen Passivierungstypen und die einzelnen Valenzpartner (siehe auch Abb. 6). Diese innere Anordnung wurde gewählt, um dem Benutzer durch das frühe Aufzeigen der Verwendungsbeispiele eine erste Orientierung zu geben und der Tatsache Rechnung zu tragen, daß gerade der lernende Leser meist induktiv vorgeht (Heringer 1986). Da die in der Datenbank gesammelten Verwendungsbeispiele immer an ein bestimmtes Phänomen (z.B. einzelne Valenzpartner oder Passivierungstypen) gebunden sind und nicht unbedingt für jedes Verb-Monosem zur Verfügung stehen, steht die Bedeutungsparaphrase aus Verständnisgründen ebenfalls vor den übrigen lexikalischen Angaben. Die Liste der Valenzpartner findet sich zwar erst am Schluß eines jeden Eintrags, ist aber durch ihre äußere Form sowie die farblich markierten Aktionswörter hervorgehoben und somit auch direkt identifizierbar in Benutzungssituationen, in denen es dem Anwender um Informationen über die einzelnen Valenzpartner geht. Zu jedem Valenzpartner eines Verbs werden neben dem Komplementtyp entsprechend dem Quellwörterbuch auch evtl. differierende Klassifikationen nach der der Hypertextkomponente zugrundeliegenden *Grammatik der deutschen Sprache* dargestellt. Weiterhin stehen dort Angaben über die Optionalität, den Argumentstatus und den Rollenbezeichner.

3.2.5 Automatisierte Verweise

Die Komplementbezeichner werden bei der Präsentation der Abfrageergebnisse automatisch als Aktionswörter angelegt und gekennzeichnet. Die Abfrage der semantischen und syntaktischen Belegung ist anschließend durch die Aktivierung der entsprechenden typisierten Verknüpfungen möglich. Typisierte Verknüpfungen werden im gesamten GRAMMIS-Informationssystem zur Verbvalenz eingesetzt und bieten die Möglichkeit, von einem Aktionswort aus Verbindungen zu mehreren Hypertexteinheiten zu realisieren. Der Benutzer ist nicht in der unbefriedigenden Situation, vor dem Aktivieren eines Aktionswortes nicht zu wissen, was ihn erwartet (Sprung zu einer anderen Hypertexteinheit, Einblenden eines Beispiels in einem untergeordneten Fenster, Aufruf ei-

ner Animation, ...), sondern kann nach dem Anklicken des Aktionswortes aus einer dann erscheinenden Liste von Verbindungstypen auswählen (siehe dazu auch Abb. 4 und 5).

Diese typisierten Verknüpfungen dienen auch zur Realisierung von Verbindungen zwischen der Frontendseite der Datenbank und anderen Hypertexteinheiten. Dabei gibt es zwei Verbindungstypen: einerseits die Verknüpfung der Komplementtypen mit den Stellen im Hypertext, an denen sie eingehend beschrieben werden, andererseits die Anbindung an die im elektronischen Glossar gespeicherten Kurzdefinitionen, die in einem separaten Fenster eingeblendet werden können.

4. **Schlußbemerkungen**

In diesem Beitrag wurden Modellierung und Anbindung zweier Datenbankkomponenten an das hypermediale Informationssystem GRAMMIS beschrieben: die relativ einfach strukturierte und mit *Toolbook* implementierte Datenbank der Funktionswörter sowie die komplexere, mit Hilfe des Datenbankmanagementsystems *Paradox* umgesetzte Valenzdatenbank. Die beschriebenen Datenstrukturen sind erweiterbar und können bei einer Veränderung der Anforderungen an die Datenbanken – z.B. Bereitstellung neuer Abfragemöglichkeiten – angepaßt werden.

Insbesondere die Anbindung der Valenzdatenbank an den Hypertext über die Verbvalenz hat gezeigt, daß entsprechend den verschiedenen möglichen Benutzersituationen die Realisierung unterschiedlicher Verbindungstypen von Vorteil ist. Nach der im Zuge der Anbindung geleisteten Arbeit entsteht dadurch, daß Grammatik und Lexikon nun einen einheitlichen Informationsraum bilden, ein informationeller Mehrwert.

5. **Literatur**

- Chen, Peter P.-S./Knöll, Heinz-Dieter (1991): Der Entity-Relationship-Ansatz zum logischen Systementwurf: Datenbank- und Programmentwurf. Mannheim/Wien/Zürich.
- Date, Christopher J. (1986): An Introduction to Database Systems. I. Reading (Mass.).
- Heringer, Hans Jürgen (1987): Was lange währt. Gedanken zum Mannheimer Valenzwörterbuch. Rezension über das Valenzwörterbuch Verben in Feldern. In: Zeitschrift für Germanistische Linguistik (ZGL), 15, S. 311-317.
- Kacmar, Charles/Leggett, John/Schnase, John/Boyle, Craig (1988): Data Management Facilities of Existing Hypertext Systems. Technical Report. TAMU 88-018. Hypertext Research Lab. Texas A&M University.

- Kuhlen, Rainer (1991): *Hypertext: Ein nichtlineares Medium zwischen Buch und Wissensbank*. Heidelberg.
- Kuhlen, Rainer (1995): *Informationsmarkt: Chancen und Risiken der Kommerzialisierung von Wissen*. Konstanz.
- Nielsen, Jakob (1995): *Multimedia and Hypertext. The Internet and beyond*. Cambridge.
- Schlageter, Gunter/Stucky, Wolfried (1983): *Datenbanksysteme: Konzepte und Modelle* (2., neubearb. u. erw. Aufl.). Stuttgart.
- Schneider, Roman (1997): *Navigationsangebote und Datenbankintegration in einem hypermedialen Informationssystem zur deutschen Verbvalenz*. Magisterarbeit. Univ. Trier.
- Schumacher, Helmut (Hg.) (1986): *Verben in Feldern. Valenzwörterbuch zur Syntax und Semantik deutscher Verben*. Schriften des Instituts für deutsche Sprache 1. Berlin/New York.
- Storrer, Angelika (1995): *Die Grammatik mit der Maus – Konzeption eines multimedialen Informationssystems zur deutschen Grammatik*. In: Hitzberger, Ludwig (Hg.): *Ange wandte Computerlinguistik*. Hildesheim/Zürich/New York. S. 291-305.
- Zifonun, Gisela/Hoffmann, Ludger/Strecker, Bruno (1997): *Grammatik der deutschen Sprache*. Berlin/New York.