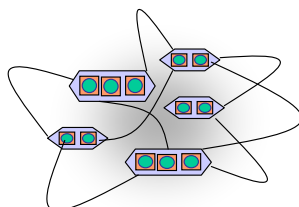


# ***TERMNET*** – ein terminologisches Wortnetz im Stile des ***PRINCETON WORDNET***



Projekt  
**Hypertextualisierung auf  
textgrammatischer Grundlage**  
([www.hytext.info](http://www.hytext.info))

**Technical Report**

Michael Beißwenger  
*beisswenger@hytext.info*

Februar 2008

<b>TERMNET in HyTex: Ein paar Worte vorab zum Projektkontext</b> .....	3
<b>1. Das TERMNET-Modell in Projektphase I (2001-2005):</b> Eine Wortnetzmodellierung im Stile des PRINCETON WORDNET, ergänzt um terminologiespezifische Relationen .....	5
<b>2. Das TERMNET-Modell in Projektphase II (2005-2008):</b> Ausbau der Wortnetz-Modellierungskonstrukte zu Zwecken der differenzierten Modellierung konkurrierender Terminologiesysteme .....	7
2.1 In terminologischen Taxonomien sind Beziehungen der Über-/Unterordnung zwischen Termini i.d.R. als Klasseninklusion konzipiert .....	9
2.2 Innerhalb eines terminologischen Systems entspricht <i>ein</i> Terminus meist genau <i>einem</i> Konzept .....	11
2.3 Einzelne Einheiten verschiedener terminologischer Systeme sind zwar unter Umständen <i>kategorienähnlich</i> – die betreffenden Termini sind jedoch in einem gegebenen Kontext nicht ohne Weiteres gegeneinander austauschbar .....	12
<b>3. Repräsentation von TERMNET II in OWL (Web Ontology Language)</b> .....	17
<b>Erwähnte Literatur</b> .....	18

## TERMNET in HyTex: Ein paar Worte vorab zum Projektkontext

Im Projekt *Hypertextualisierung auf textgrammatischer Grundlage (HyTex)*, siehe <http://www.hytext.info>) entwickeln wir Hypertextualisierungsstrategien, die eine Re-Konnexion von Textmodulen, die durch die Zerlegung ursprünglich sequenziell organisierter Fachtexte gewonnen wurden, nach Kohärenzkriterien ermöglicht. Leitlinie aller im Projekt angestellten Überlegungen ist es, Kohärenzbildungsprozesse beim selektiven Querlesen von Fachtexten besser zu unterstützen als dies in Printmedien möglich ist. Mit Blick auf Rezipienten, für die in Bezug auf eine Fachdomäne nicht Experten-, sondern lediglich Semi-Experten-Status angenommen werden kann<sup>1</sup>, erscheint es uns als eine wichtige Anforderung, die ein Linking nach Kohärenzkriterien speziell zur *Unterstützung der Fachtextrezeption* erfüllen sollte, den Rezipienten Hyperlinks auf Ressourcen bereitzustellen, die zur Gewinnung derjenigen fachlichen und fachsprachlichen (terminologischen) Wissensvoraussetzungen hilfreich sind, die vom Autor eines Fachtextes in Bezug auf seine anvisierten Adressaten getroffen wurden. Die betreffenden Ressourcen können – z.B. im Falle von Definitionen, die in einer bestimmten Textpassage zu einem vom Autor verwendeten Terminus gegeben werden – entweder Teile des Textes selbst sein (deren Kenntnis beim Querlesen aber nicht unweigerlich vorausgesetzt werden kann) oder aber externe Ressourcen (z.B. Textpassagen anderer Texte, auf die der Autor des Textes verweist und deren Kenntnis er für seine eigene Argumentation voraussetzt).

Im *HyTex*-Projekt fokussieren wir in den Arbeitsbereichen „Terminologiesensitives Linking“ und „Wortnetzbasierende Fachdomänenmodellierung“ schwerpunktmäßig

- (a) auf Strategien für die Verlinkung der Verwendungsinstanzen von Fachtermini mit den von den Autoren der betreffenden Texte diesen Verwendungen zugrundegelegten Definitionen und
- (b) auf die korpusgestützte Modellierung eines terminologischen Wortnetzes, das die Fachtextdomäne<sup>2</sup> anhand ihrer zentralen Termini und Konzepte und der zwi-

---

1 Unter einem *Semi-Experten* verstehen wir (in Anlehnung an Arbeiten aus dem Bereich der Wörterbuchforschung von H.E. Wiegand) einen Nutzer einer bestimmten Informationsressource, die Informationen zu einer Domäne bereithält, in der er nicht zu den Experten zählt, aber in Bezug auf die er ein gewisses Vorverständnis mitbringt. Typische Semi-Experten in Bezug auf Wissenschaftsdomänen sind beispielsweise Studierende, Wissenschaftsjournalisten oder auch Wissenschaftler (Experten) anderer Disziplinen im Kontext interdisziplinären Arbeitens.

2 Da Fachdomänen, wenn man sie auf der Grundlage von Textkorpora (und nicht etwa wissenschaftssoziologisch oder in Hinblick auf die mündliche Fachkommunikation) betrachtet, stets nur über die in den jeweiligen Korpora (die grundsätzlich immer nur einen Ausschnitt der für die Fachdomäne relevanten bzw. konstitutiven Texte umfassen) enthaltenen Textdokumente repräsentiert werden, sprechen wir in Bezug auf den Bezugsbereich unserer Modellierung von einer *Fach-*

schen ihnen feststellbaren lexikalischen und konzeptuellen Relationen im Stile des PRINCETON WORDNET (Fellbaum 1998) beschreibt; die dabei entstehende Resource bildet dann die Grundlage für Strategien der Verlinkung der Verwendungsinstanzen von Fachtermini in den Korpus texten mit einer aus dem Wortnetz generierten Glossar-Sicht auf die in der Domäne verwendete Terminologie, die es den Rezipienten erlauben soll, sich – über die zu einzelnen Termini im Korpus belegten Definitionen hinausgehend – über das konzeptuelle Umfeld eines Terminus innerhalb des betreffenden Fachwortschatzes zu informieren und sich so, ausgehend von einzelnen Fachtermini, das relevante Konzeptwissen zur Domäne zu erarbeiten.<sup>3</sup>

Im Fokus der vorliegenden Dokumentation steht das Modell, welches dem Aufbau des terminologischen Wortnetzes (TERMNET) zugrundeliegt und das in den Projektphasen I (2001-2005) und II (2005-2008) in zwei verschiedenen Versionen ausgearbeitet wurde (TERMNET I und TERMNET II).

Das zugrundegelegte Fachtextkorpus wurde für das Projekt aus dem WWW akquiriert und umfasst neben wissenschaftlichen Fachartikeln auch Dokumenttypen mit normativem („Spezifikation“) und didaktischem Anspruch („Tutorial“, „Einführung“, „Überblicksdarstellung“), des Weiteren diskursiv geprägte Textsorten wie FAQs, Mailinglisten- oder Foren-Postings. Ein Ausschnitt aus dem Korpus bildet das *Kernkorpus*, auf dem wir unsere Linking-Strategien in Prototyp-Versionen implementieren und testen. Eine Dokumentation zu Inhalt und Zusammensetzung des Korpus findet sich in Beißwenger & Wellinghoff (2006).<sup>4</sup>

---

*textdomäne*. Dies ist zu lesen als „Die Fachdomäne, so wie sie sich in den zugrunde gelegten Texten zeigt“. Den Ausdruck *Fachdomäne* verwenden wir dann, wenn wir uns ganz allgemein auf das Thema unserer Modellierung, nicht aber auf deren – mit dem Zuschnitt unseres Korpus gegebenen – *Bezugsbereich* beziehen.

3 Zum Stellenwert von TERMNET innerhalb des HyTex-Anwendungsszenarios vgl. ausführlich Storrer (i.Dr.).

4 Das Fachtextkorpus hat einen Umfang von 160 Korpusdokumenten mit insgesamt 2.945.597 laufenden Wortformen. Das Korpus gliedert sich in drei Teilkorpora, die unter [http://www.hytext.info/040\\_werkstatt/010\\_hytextphase\\_eins/030\\_ergebnisse/020\\_korpus](http://www.hytext.info/040_werkstatt/010_hytextphase_eins/030_ergebnisse/020_korpus), z.T. auch in annotierten Versionen, zum Download und zur freien Nutzung bereitgestellt werden.

## 1. Das TERMNET-Modell in Projektphase I (2001-2005): Eine Wortnetzmodellierung im Stile des PRINCETON WORDNET, ergänzt um terminologiespezifische Relationen

In Projektphase I (im Folgenden: *HyTex I*) wurde ein terminologisches Wortnetz zur Fachtextdomäne „Hypertextforschung und Texttechnologie“ aufgebaut, das insgesamt 226 terminologische Einheiten umfasst, die in einem Fachtextkorpus zur Domäne mit jeweils mindestens einer Definition belegt sind (*TermNet I*).<sup>5</sup> Die Ressource wurde komplett im ISO-Standard *XML Topic Maps (XTM)*; vgl. Pepper & Moore 2001) repräsentiert.<sup>6</sup> Ein in *HyTex I* entwickelter Anwendungsprototyp<sup>7</sup>, der die verschiedenen im Projekt entwickelten Strategien für die Hypertextualisierung nach Kohärenzkriterien für eine Auswahl an Korpustexten implementiert und demonstriert, nutzt auch die *TERMNET I*-Ressource und generiert aus ihr eine Glossarsicht, die einerseits als Hilfe bei der selektiven Textlektüre genutzt werden kann (insofern sich aus ihr Informationen zum lexikalischen Umfeld von Termverwendungsinstanzen beziehen lassen) und andererseits eine von den Termini ausgehende Möglichkeit des Einstiegs in die und der Navigation in der Fachtextdomäne darstellt. Die Einträge in

---

5 Modelliert wurden ausschließlich nominale Termini (inklusive Mehrworteinheiten mit nominalem Kopf); Kriterium für die Aufnahme eines im Korpus belegten Terminus in *TERMNET I* war – neben der grammatischen Kategorie – einerseits die Zugehörigkeit zur Fachsprache einer der beiden Domänen „Hypertextforschung“ und „Texttechnologie“, andererseits das Vorhandensein mindestens einer Definition des Terminus im Kernkorpus. Termini, die der Domäne problemlos zugeordnet werden konnten, zu denen im Kernkorpus aber keine Definition belegt war, wurden bei der Modellierung nicht berücksichtigt. Die Koppelung der Entscheidung, ob ein Terminus in *TERMNET I* aufgenommen werden sollte, an das Vorhandensein einer Definition begründet sich aus dem für das Projekt insgesamt definierten Anwendungsszenario, das vorsieht, die Einträge des aus dem terminologischen Wortnetz generierten Glossars in den Korpusdokumenten zu verankern, und zwar an genau denjenigen Textsegmenten, die als Definitionen der jeweiligen Termini identifiziert wurden; umgekehrt werden auch die betreffenden Passagen der Korpustexte wiederum mit den Einträgen im Glossar verlinkt, so dass bei der Textrezeption zu einem an einer bestimmten Textstelle verwendeten Terminus jeweils direkt der korrespondierende Glossareintrag aufgerufen und in seiner Vernetzung mit den übrigen Einträgen des terminologischen Netzes exploriert werden kann. Die Identifizierung und Annotation definitorischer Textsegmente im Textkorpus bildet im Zuge der *HyTex*-Arbeiten einen eigenen Arbeitsbereich (vgl. hierzu Lenz, Beißwenger & Wellinghoff 2006, Wellinghoff 2006 und Beißwenger 2004).

6 Für Aufbau und Pflege der *TERMNET I*-Datenbasis wurde das Werkzeug *K-Infinity* eingesetzt, für dessen Bereitstellung wir der Firma *Intelligent Views* (<http://www.i-views.de/web/>) danken; die Datenbasis wurde anschließend automatisch in ein XTM-Format konvertiert, das die Grundlage für die Durchführung von Inferenzen über der Netzrepräsentation und für die Generierung einer HTML-basierten Anwendungssicht bildete (vgl. hierzu Lenz, Birkenhake & Maas 2003).

7 Der Prototyp kann im WWW unter [http://www.hytex.info/030\\_ergebnisse/010\\_prototyp](http://www.hytex.info/030_ergebnisse/010_prototyp) in Augenschein genommen werden.

der *TERMNET I*-Ressource sind mit den Definitionen der durch sie repräsentierten terminologischen Einheiten im Textkorpus verlinkt. Da Verwendungsinstanzen und Definitionen zu Termini im Korpus annotiert sind, konnte in der Anwendung eine direkte Vernetzung der Glossareinträge mit den Textbelegen und auch eine Vernetzung der Vorkommnisse der Termini in den Texten mit den Einträgen im Glossar umgesetzt werden.

Zentrale Modellierungskonstrukte im *TERMNET I*-Modell sind – ebenso wie im *PRINCETON WORDNET* (Fellbaum 1998) und in seiner deutschsprachigen Entsprechung, dem Tübinger *GERMANET* (Kunze 2001) – zum einen die *lexikalische Einheit* („word“ in der *WORDNET*-Terminologie) und zum anderen das *Synset*, das eine oder mehrere lexikalische Einheiten zusammenfasst, die dasselbe Konzept lexikalisieren. Das *Synset*-Konstrukt beschreibt somit ein Konzept anhand seiner Lexikalisierungen. Bei der Zusammenfassung von lexikalischen Einheiten zu ein- und demselben *Synset* wird hierbei kein striktes, sondern ein weitgefasstes Synonymiekonzept zugrundegelegt. In *TERMNET I* umfasst ein *Synset* jeweils all diejenigen Fachtermini, die *in ungefähr* dasselbe Konzept lexikalisieren. Da alternative Termini zu ähnlichen Konzepten in der Regel keine bloßen Benennungsalternativen darstellen, sondern sich mit alternativen *Terminologisierungen* auch jeweils partiell alternative *Konzeptualisierungen* verbinden, sind die *Synsets* in *TERMNET I* eher Konstrukte zur Beschreibung *thematischer Identität*, indem sie diejenigen terminologischen Einheiten zusammenfassen, die in etwa denselben außersprachlichen Sachverhaltsausschnitt (eine ähnliche Kategorie) bezeichnen, diesen aber u.U. auf unterschiedliche Weise konzipieren. So bezeichnen beispielsweise die Termini ‚Hyperlink‘ und ‚Hypertext‘ ein- und dasselbe Thema (denselben außersprachlichen Sachverhaltsausschnitt), konzipieren es aber schon allein aufgrund der unterschiedlichen Metaphorik unter verschiedenem Aspekt (nämlich einmal als Verknüpfung und einmal als Verweis). Die Termini ‚Hyperlink‘ und ‚Link‘ hingegen sind zwar ebenfalls Alternativen, in ihrem Fall ist die Alternative allerdings wortbildungsmorphologisch zu erklären, insofern ‚Link‘ durch Kürzung aus ‚Hyperlink‘ gebildet wurde. Beide Termini sind somit auch konzeptuell weitestgehend identisch (wenn auch ‚Link‘ als eigenständig idiomatisierte Einheit unabhängig von ‚Hyperlink‘ erworben werden kann, das Wissen um den zugrundeliegenden Kürzungsprozess dabei nicht zwangsläufig miterworben werden muss und ‚Hyperlink‘ durch den Modifikator ‚hyper-‘ einen zusätzlichen Aspekt in das Konzept der bezeichneten Sache einbringt, die in der Bezeichnung ‚Link‘ nicht mittransportiert wird<sup>8</sup>).

---

8 Bei der Besonderheit, dass die Kurzform ‚Link‘ semantisch „ärmer“ ist als die Vollform ‚Hyperlink‘, handelt es sich aber nicht um ein Phänomen der Terminologiebildung, sondern um ein

Das dritte Modellierungskonstrukt in *TERMNET I* ist – neben *lexikalischen Einheiten* und *Synsets* und ebenfalls analog zu *WORDNET* und *GERMANET* – die *Relation*. Die Basisrelation bei der Modellierung semantisch-lexikalischer Wortnetze ist die Synonymierrelation, die zwischen den Mitgliedern ein- und desselben Synsets besteht (also als lexikalische Relation zwischen denjenigen lexikalischen Einheiten, die dasselbe Konzept lexikalisieren). Hinzu kommen weitere Relationen, die – je nach dem, ob sie zwischen lexikalischen Einheiten oder zwischen Synsets (als Repräsentanten von Konzepten) bestehen – zu *lexikalischen* und *konzeptuellen Relationen* zusammengefasst werden.

Das der Modellierung von *TERMNET I* zugrundegelegte Relationeninventar schließt an die in *WORDNET* etablierten Relationstypen an und erweitert diese um genuine *TERMNET*-Relationen, die den spezifischen lexikalischen und konzeptuellen Besonderheiten von Fachwortschätzen Rechnung tragen. Eine Beschreibung des Relationeninventars in *TERMNET I* findet sich in Beißwenger, Storrer & Runte (2003). Eine Liste aller in *TERMNET I* modellierten Termini und Relationen findet sich im Anhang von Runte (2004).

## **2. Das *TERMNET*-Modell in Projektphase II (2005-2008):**

### **Ausbau der Wortnetz-Modellierungskonstrukte zu Zwecken der differenzierten Modellierung konkurrierender Terminologiesysteme**

In Projektphase II (im Folgenden: *HyTex II*) wurde die lexikalische Basis des terminologischen Wortnetzes erweitert und das gesamte *TERMNET* auf Basis eines erweiterten Wortnetzmodells remodelliert (*TERMNET II*). Die Ressource wurde komplett in der Ontologiebeschreibungssprache OWL (Dean & Schreiber 2004) abgebildet und wurde 2007 mit einer OWL-Repräsentation des allgemeinsprachlichen Wortnetzes *GERMANET* verknüpft (Selzam 2008; Kunze, Lemnitzer, Lungen & Storrer 2007).

Mit der Erweiterung der lexikalischen Basis von ursprünglich 226 (definierten) lexikalischen Einheiten auf nunmehr 427 Einträge wurde versucht, *TERMNET* in Hinblick auf die in der Literatur der Fachdomäne dargestellten terminologischen Taxonomien systematisch zu komplettieren. Für die Teildomäne *Hypertextforschung* wurden relevante Termini von Angelika Storrer, für die Teildomäne *Texttechnologie* von Eva Anna Lenz ergänzt.<sup>9</sup>

---

Wortbildungsphänomen; daher erscheint es als legitim, für die Termini ‚Hyperlink‘ und ‚Link‘ konzeptuelle Identität anzunehmen.

<sup>9</sup> Eine Liste aller in *TERMNET II* aufgenommenen Termini ist online unter [http://www.hytext.info/040\\_werkstatt/010\\_hytextphase\\_eins/Liste-aller-TermSets-und-Lexeme.pdf](http://www.hytext.info/040_werkstatt/010_hytextphase_eins/Liste-aller-TermSets-und-Lexeme.pdf) abrufbar. Die Liste umfasst

Der zentrale Unterschied des Modells, nach dem die Datenbasis von *TERMNET I* inklusive der in *HyTex II* hinzugenommenen Erweiterungen in *TERMNET II* remodelliert wurde, besteht darin, dass das Inventar der Modellierungskonstrukte um ein Konstrukt *Termset* erweitert wurde, das es ermöglicht, kategorienähnliche Konzepte aus unterschiedlichen terminologischen Systemen zueinander in Beziehung zu setzen. Das im Projekt verfolgte Teilziel einer Entwicklung von Strategien zur Generierung von Linkangeboten, die die Rekonstruktion terminologiebedingter Wissensvoraussetzungen ermöglichen sollen, wurde in *HyTex II* erweitert um die Perspektive, dem Hypertextrezipienten über die Bereitstellung einer Glossar-Sicht auf die Wortnetzressource nicht nur das schnelle Auffinden von *Bezügen zwischen einzelnen Termini*, sondern darüber hinaus auch von *Bezügen zwischen ganzen (terminologischen) Kategorisierungssystemen* zu ermöglichen. Für viele Bereiche der Wissenschaftskommunikation ist es von großer Bedeutung, nicht nur zu wissen, dass zwei terminologische Einheiten ein ähnliches Konzept bezeichnen, sondern auch, dass sie verschiedenen terminologischen Systemen angehören und daher – trotz der Ähnlichkeit der mit ihnen jeweils verbundenen Konzepte – nicht beliebig füreinander ausgetauscht werden können. Terminologische Systeme konzeptualisieren außersprachliche Wirklichkeit und machen sie begrifflich und sprachlich fassbar. Konkurrierende terminologische Systeme können denselben oder einen ähnlichen Weltausschnitt auf unterschiedliche Weise konzeptualisieren und sprachlich strukturieren. Ähnlichkeiten zwischen Einheiten unterschiedlicher Systeme sind daher eher auf der Ebene der durch die terminologische Strukturierung konstituierten Kategorisierungen als auf der Ebene der Konzepte selbst zu beschreiben. Das Modellierungskonstrukt *Termset* dient dazu, die Ähnlichkeit dessen, was Termini aus unterschiedlichen terminologischen Systemen bezeichnen, als *Kategorienähnlichkeit* zu beschreiben; das Konstrukt des *Synsets* wird beibehalten für Konzepte, die Konzepte können – im Gegensatz zu *TERMNET I* – dann aber als spezifisch für einen jeweils einzelnen Terminus (innerhalb eines bestimmten terminologischen Systems) ausgewiesen werden. Charakteristika fachsprachlicher Kategoriensysteme und Motivation fachsprachlichen Kategorisierens werden dadurch, so ein Anspruch des *TERMNET II*-Modells, differenziert abbildbar, Bezüge zwischen konkurrierenden terminologischen Systemen werden deut-

---

insgesamt 473, die Modellierung hingegen nur 427 Termini, da terminologische Ausdrücke, wenn sie als orthographische Varianten desselben terminologischen Konzepts aufgefasst werden können, in der Netzrepräsentation als Formvarianten desselben Terminus ausgewiesen werden. Beispiele für orthographische Varianz können mit variierender Schreibung von Komposita zusammenhängen (z.B. ‚Hyperlink‘ vs. ‚Hyper-Link‘) oder mit durch die Orthographiereform verursachten Varianten (z.B. ‚referentieller Link‘ vs. ‚referenzieller Link‘). Die in der Liste angegebenen Termini sind nach ihrer Zugehörigkeit zu einzelnen *Termsets* angeordnet; zum Modellierungskonstrukt *Termset* s.u., Abschnitt 2.3.



licher als in einem (für gemeinsprachliche Wortnetze durchaus angemessenen) Modell, das nur lexikalische Einheiten und Synsets kennt.<sup>10</sup> Einige für die Erweiterung des TERMNET-Modells zentrale Überlegungen werden im Folgenden kurz skizziert (Abschnitte 2.1–2.3). Abschnitt 3 gibt einen Überblick über die Modellierung von TERMNET II in der Ontologiebeschreibungssprache OWL (Web Ontology Language).

## 2.1 In terminologischen Taxonomien sind Beziehungen der Über-/Unterordnung zwischen Termini i.d.R. als Klasseninklusion konzipiert

*Hyponymie* besteht zwischen Konzepten, *Inklusion* zwischen Klassen von Dingen in der Welt. Konzepte determinieren Kategorien, die für die Beurteilung der Klassenzugehörigkeit von Dingen/Sachverhalten erforderlich sind; dennoch sind Eigenschaften des unter eine bestimmte Kategorie fallenden außersprachlichen Sachverhaltsausschnitts und seiner Vertreter nicht zwangsweise identisch mit *Bedeutungsmerkmalen*, die Eigenschaften von *Konzepten* sind.

In der Gemeinsprache muss es im Einzelfall nicht notwendigerweise Bestandteil des Bedeutungswissens zu einem mit einem bestimmten Lexem verbundenen Konzept sein zu wissen, zu welchem anderen Konzept es in einer Hyponymiebeziehung steht. So muss zum Lexem ‚Wal‘ nicht notwendigerweise gewusst werden, dass die Extension der dadurch bezeichneten Kategorie eine Subklasse der mit dem Hyperonym ‚Säugetier‘ bezeichneten Kategorie darstellt; Gleiches gilt für das Lexem ‚Pinguin‘ und sein Hyperonym ‚Vogel‘. Dass die Extension von ‚Pinguin‘ von der Extension von ‚Vogel‘ und die Extension von ‚Wal‘ von der Extension von ‚Säugetier‘ inkludiert

---

10 Unterschiedliche, aber kategorienähnliche terminologische Systeme können als alternative Wortschätze zur Beschreibung desselben Sachverhaltsbereichs aufgefasst werden, die aber den Konzeptbereich verschieden strukturieren. Da Termini im Gegensatz zu gemeinsprachlichen Lexemen in aller Regel definiert sind, lassen sich die Grenzen zwischen Konzepten in Fachsprachen deutlich klarer identifizieren als die Grenzen zwischen Wortbedeutungen der Alltagssprache. Termini desselben Urhebers oder derselben Schule, die ontologisch verwandte Sachverhaltsausschnitte bezeichnen, lassen sich somit als ein *System* im Sinne des Strukturalismus Saussurescher Prägung auffassen, in dem die lexikalisierten Bedeutungen sich wechselseitig begrenzen und jede Einheit ihre spezifische Funktion bei der Strukturierung des Sachverhaltsausschnitts aus der Opposition zu den übrigen Einheiten gewinnt. In der Regel ist auch davon auszugehen, dass sich der Urheber einer bestimmten Terminologie in der Art und Weise, wie er von dieser Terminologie Gebrauch macht, an die von ihm per Definition festgelegten Konzeptualisierungen hält; die Regeln des Gebrauchs sind also im Bereich der Terminologie (innerhalb des Paradigmas ein- und desselben terminologischen Systems) nicht in gleicher Weise erweiterbar wie dies in der Alltagssprache möglich ist (die stabile Fixierung von Konzept-Ausdrucks-Verbindungen ist ja gerade ein wesentlicher Zweck des Definierens); terminologische Bedeutung ist in weitaus geringerem Maße kontextabhängig als Wortbedeutungen in alltagssprachlichen Benennungskontexten.

wird, muss von einem Sprecher des Deutschen nicht gewusst werden, um die Ausdrücke ‚Pinguin‘ und ‚Wal‘ so zu verwenden, wie es die übrigen Mitglieder der Sprachgemeinschaft mehrheitlich tun. Die Inklusion spielt hier nur ontologisch eine Rolle, semantisch aber nicht (unbedingt). Hingegen setzt das Verständnis des Konzepts ‚unidirektionaler Link‘ aus der Fachdomäne „Hypertextforschung“ das Wissen voraus, dass die damit bezeichnete Kategorie als eine Subklasse derjenigen Kategorie aufgefasst wird, die mit ‚Link‘ bezeichnet wird. (Die Kompositionsregel zeigt sich bei Mehrworttermini, die Teil einer Taxonomie sind, häufig charakteristischerweise in der Struktur ihrer Benennung.) Das Wissen um die Inklusionsbeziehung zwischen ‚Link‘ und ‚unidirektionaler Link‘ ist somit notwendiger Bestandteil des mit ‚unidirektionaler Link‘ verbundenen Konzepts. Die Beziehung zwischen den beiden Termini kann als *superclass–subclass* interpretiert werden.

Typisch für terminologische Taxonomien ist weiterhin, dass sich die einem generellen Terminus untergeordneten Termini nach dem relevanten klassifikatorischen Merkmal in wechselseitig disjunkte Subklassen gruppieren lassen. Beispielsweise gibt es zum Terminus ‚Verknüpfung‘ in der Fachdomäne „Hypertextforschung“ die spezifischeren Termini ‚monodirektionale Verknüpfung‘, ‚bidirektionale Verknüpfung‘ (Merkmal: *Direktionalität*) und ‚intratextuelle Verknüpfung‘, ‚intertextuelle Verknüpfung‘, ‚extratextuelle Verknüpfung‘ (Merkmal: *Position des Linkziels*). Die Termini, die nach demselben Merkmal klassifiziert sind, denotieren jeweils disjunkte Teilklassen. Eine Modellierung von Termini als Klassen erlaubt eine Beschreibung der Disjunktivität als Beziehung zwischen Extensionen anstatt als ein Phänomen, mit dem untergeordnete Konzepte (Mengen von Kohyponymen zu demselben Hyperonym) gruppiert werden. Dass ‚monodirektionale Verknüpfung‘ und ‚extratextuelle Verknüpfung‘ nicht disjunkt sind und daher in Kombination auf ein Exemplar der Klasse ‚Link‘ zutreffen können, ist nicht ein Merkmal der zugehörigen Konzepte, sondern eine Eigenschaft der durch die Termini bezeichneten Klassen, die sich aus dem Klassifikationskriterium ergibt, das bei ihrer Schöpfung jeweils angewandt wurde. Im *TERMNET I*-Modell, das nur lexikalische Einheiten und Konzepte kannte, wurde über die Angabe des der Genese eines untergeordneten Konzepts zugrundeliegenden Klassifikationsmerkmals die Disjunktivität zwischen mit den Konzepten verbundenen Klassen indirekt dargestellt; im *TERMNET II*-Modell beschreiben wir Termini als Klassen, die hierarchische *superclass-subclass*-Beziehungen aufweisen können und zwischen denen wechselseitige Disjunktivität *als Relation* angegeben werden kann.

## 2.2 Innerhalb eines terminologischen Systems entspricht *ein* Terminus meist genau *einem* Konzept

Innerhalb eines terminologischen Systems gibt es i.d.R. keine lexikalische Ambiguität. *Homonymie* tritt nicht zufällig auf, *Polysemie* ist selten, da sprachlicher Wandel in den terminologischen Ausschnitten von Fachsprachen anderen Regeln unterliegt als der Bedeutungswandel in der Alltagssprache.

*Synonymie* tritt auf, ist aber, wenn sie auftritt, systematisch erklärbar. Typisch für das Auftreten synonyme Termini sind Kürzungsoperationen, mit denen Kurzwörter oder Akronyme gebildet werden (z.B. ‚Hyperlink‘ > ‚Link‘, ‚World Wide Web‘ > ‚WWW‘). Das gleichzeitige und gleichrangige Vorhandensein eines Lehnworts (dt. ‚Link‘ < engl. ‚link‘) und einer Lehnübersetzung (dt. ‚Verknüpfung‘ < engl. ‚link‘) zu einem ursprünglich fremdsprachigen Terminus innerhalb *eines* terminologischen Systems ist eher die Ausnahme; solche terminologieübernahmebedingten Lokalisierungsvarianten kommen eher in unterschiedlichen terminologischen Systemen vor.

Wortbildungsmorphologische Beziehungen zwischen einzelnen Termini werden, wenn ein Terminus offensichtlich durch Kürzung aus einem anderen Terminus gebildet wurde, in Form einer *Abkürzungsrelation* explizit modelliert. Synsets werden im *TERMNET II*-Modell nicht mehr explizit dargestellt, können aber durch Auswertung der zwischen einzelnen Termini dargestellten Abkürzungsrelationen im Bedarfsfall inferiert werden. Generell verliert unter der Maßgabe, dass innerhalb *eines* terminologischen Systems *ein* Terminus genau *einem* Konzept entspricht, das Modellierungskonstrukt *Synset* an Relevanz: Im Idealfall konstituiert *ein* Terminus *ein* Konzept; mehrere Termini werden nur dann noch in einem *Synset* (d.h. zu einer in mehreren sprachlichen Einheiten lexikalisierten Konzeptrepräsentation) zusammengefasst, wenn sie sich wortbildungsmorphologisch als Varianten voneinander ausweisen lassen; *bedeutungsähnliche* Termini aus unterschiedlichen terminologischen Systemen (wie z.B. ‚1:1-Verweis‘, ‚1:1-Link‘ und ‚1:1-Verknüpfung‘ aus drei verschiedenen terminologischen Taxonomien in der Domäne „Hypertextforschung“) werden nicht mehr – wie in *TERMNET I* – in einem *Synset* zusammengeführt bzw. als Lexikalisierungen desselben Konzepts ausgewiesen, sondern zunächst als Klassen innerhalb autonomer terminologischer Systeme dargestellt und erst auf einer weiteren (neuen) Modellierungsebene, der Ebene der *Termsets* (s.u.) – zueinander in Beziehung gesetzt.

### 2.3 Einzelne Einheiten verschiedener terminologischer Systeme sind zwar unter Umständen *kategorienähnlich* – die betreffenden Termini sind jedoch in einem gegebenen Kontext nicht ohne Weiteres gegeneinander austauschbar

In unterschiedlichen terminologischen Systemen kann ein- und derselbe terminologische Ausdruck unterschiedliche Konzepte bezeichnen. So bezeichnet ‚Partikel‘ in der DUDEN-Grammatik eine bestimmte Subklasse der nichtflektierbaren Wortarten, in manchen Schulgrammatiken wird der Ausdruck hingegen als Oberbegriff für alle nichtflektierbaren Wortarten verwendet. Dieser Fall ist für eine Wortnetzmodellierung wie die in *TERMNET I* verfolgte unproblematisch, da die betreffenden Termini als Homonyme behandelt und als zwei *lexikalische Einheiten* modelliert werden können, die unterschiedlichen *Synsets* angehören.

Problematischer sind Fälle, in denen in verschiedenen terminologischen Systemen *ähnliche* Konzepte durch *unterschiedliche* terminologische Etiketten konstituiert werden. Solche Einheiten als *lexikalische Einheiten* zu modellieren, die demselben *Synset* angehören und damit dasselbe Konzept lexikalisieren, wird den Charakteristika terminologischer Systeme und den Motivationen beim fachsprachlichen Klassifizieren nur bedingt gerecht. Unter der Perspektive, dass Terminologisieren kein bloßes Benennen darstellt, sondern eine Operation im Rahmen der Konzeptualisierung und Systematisierung außersprachlicher Wirklichkeit, die Sachverhaltsbereiche kategoriell fassbar macht und einen exklusiven sprachlichen Zugriff auf sie bereitstellt, ist es für die Konzeptualisierung eines Sachverhaltsausschnitts nicht unwesentlich, *unter welcher Motivation* das für ihn reservierte sprachliche Etikett gewählt oder gebildet wird und wie der übrige Sachverhaltsbereich terminologisch-konzeptuell strukturiert ist.

#### Ein Beispiel:

Die Bezeichnung ‚Proterm‘ in der „Grammatik der deutschen Sprache“ (Zifonun, Hoffmann & Strecker 1997) bezeichnet eine Kategorie, die der Kategorie ‚Pronomen‘ aus anderen Grammatiken sehr ähnlich ist – das Konzept, das mit der Benennung ‚Proterm‘ von der Kategorie entworfen wird, ist aber ein anderes als dasjenige, das in manchen anderen Grammatiken mit ‚Pronomen‘ verbunden wird. Die Kategorie wird in der GDS also nicht lediglich anders *benannt*, sondern vielmehr auch anders *konzipiert*; die Spezialität dieser alternativen Konzeptualisierung wird aber genau dann erst ersichtlich, wenn in einem Wortnetzmodell zur Terminologie im Bereich „Grammatik/Wortarten“ die beiden Bezeichner nicht als Lexikalisierungen desselben (oder – undifferenziert – eines „irgendwie ähnlichen“) Konzepts ausgewiesen werden. Für den Semi-Experten ist es hilfreich, zum Terminus ‚Proterm‘ zu wissen,

- (a) dass der Terminus ‚Proterm‘ einem anderen terminologischen System angehört als der Terminus ‚Pronomen‘ (und ggf., wie die unterschiedlichen Systeme den Konzeptbereich sprachlich differenzieren);

- (b) dass dem Terminus ‚Proterm‘ innerhalb des Systems, dem er zugehört, ein Terminus ‚Determinativ‘ nebengeordnet ist, der eine Kategorie bezeichnet, die Überschneidungen zu derjenigen Kategorie aufweist, die in anderen Systemen mit ‚Pronomen‘ bezeichnet wird;
- (c) dass der Terminus ‚Proterm‘ somit nicht „nur“ eine *alternative Konzeption* dessen darstellt, was ansonsten häufig als ‚Pronomen‘ bezeichnet wird, sondern dass das terminologische System, dem ‚Proterm‘ angehört, insgesamt den sprachlich zu fassenden Sachverhaltsausschnitt (= die Eigenschaften der Wörter des Deutschen) anhand spezifischer Kriterien konzipiert, die gegenüber anderen Systemen zu spezifisch unterschiedlichen Kategorisierungen führen, die sich nicht allein auf Wortebene (d.h. durch die bloße – instensionale und extensionale – Gegenüberstellung mit Termini anderer Systeme, die ein ähnliches Konzept bezeichnen) verstehen lassen (zum „Verstehen“ der Kategorienbildung *PROTERM*, die durch die spezifische Konzeptualisierung in der GDS vorgenommen wird, gehört mehr als die Kenntnis aller Mitglieder der Kategorie).

Das Beispiel illustriert: Ein- und derselbe Sachverhaltsbereich kann durch unterschiedliche theoretische Zugriffe unterschiedlich (u.U. in bewusster Abgrenzung voneinander) konzipiert werden. Durch die unterschiedliche Konzeption wird der zugehörige Sachverhaltsbereich unterschiedlich strukturiert, die spezifische Art der konzeptuellen Strukturierung wiederum determiniert spezifische kategoriale Strukturen. Unterschiedliche Wortartenklassifikationen sind hierfür ein gutes Beispiel. Selbst wenn in zwei unterschiedlichen Systemen dasselbe Wort zwei identisch benannten und mit ähnlichen Konzepten verbundenen Wortartenkategorien angehört, so wird durch die Gesamtheit des jeweiligen terminologischen Systems (und der damit verbundenen Konzeptualisierung des gesamten Sachverhaltsbereichs) das betreffende Wort in seiner Zugehörigkeit zur betreffenden Kategorie, zumindestens partiell, unterschiedlich motiviert. Dass „schön“ sowohl nach der Einteilung der DUDEN-Grammatik als auch für die GDS ein Vertreter der Kategorie ‚Adjektiv‘ ist, bedeutet nicht zwangsläufig, dass die Aspekte, unter welchen „schön“ den jeweiligen Kategorien zuzurechnen ist, identisch sind. Die Kategorien ‚Adjektiv‘ in beiden Grammatiken sind konzeptuell nicht identisch – beispielsweise sind auf ausschließlich prädikativen Gebrauch beschränkte, nicht flektierbare Wörter nach DUDEN (noch) Adjektive, während die GDS dafür eine eigene Kategorie (‚Adkopula‘) konzipiert. „schön“ ist zwar kein solcher „Problemfall“ – für die Identifizierung von „schön“ als Vertreter der Kategorie ergeben sich durch diese unterschiedlichen Konzeptionen von ‚Adjektiv‘ (und benachbarte Kategorien) aber unterschiedliche Begründungen. Dass „schön“ der Kategorie ‚Adjektiv‘ zugehört, kann nach der DUDEN-Klassifikation beispielsweise nicht allein dadurch begründet werden, dass „schön“ flektiert und *nicht nur* prädikativ (sondern auch attributiv) verwendbar ist, nach der GDS hingegen schon.

Selbst wenn zwei Termini aus zwei verschiedenen terminologischen Systemen dasselbe oder ein annähernd gleiches Konzept bezeichnen, können sie in der Fachkommunikation nicht beliebig für einander ausgetauscht werden. Charakteristisch für konkurrierende terminologische Systeme ist nicht nur, dass sie unterschiedliche Konzeptualisierungen ähnlicher Sachverhaltsbereiche leisten, sondern auch, dass sie ihre Konzepte mit sprachlichen Etiketten versehen, die gerade als charakteristisch für genau dieses System verstanden werden. Charakteristisch für die terminologische Klassifikation von Sachverhaltsbereichen (bzw. die semantisch-lexikalische Struktu-

rierung fachsprachlicher Wortfelder) ist beispielsweise, dass nominale Benennungen für hyponyme Konzepte die Benennung des hyperonymen Konzepts als Basis verwenden und durch Attribuierung spezifizieren. Die so entstehenden Mehrworttermini sind insofern „sprechend“, als sie für Rezipienten, denen das hyperonyme Konzept sowie ggf. innerhalb der Attribuierung(en) vorkommende adjektivische oder nominale Termini geläufig sind, als kompakte Konzeptualisierungen mit Definiens-Struktur fungieren können: Der Terminus ‚unidirektionaler Link‘ bezeichnet Exemplare des Typs Link (*genus proximum*), die unidirektional sind (*differentia specifica*), der Terminus ‚durchgängig subintegrierte Mikrostruktur‘ bezeichnet Exemplare des Typs Mikrostruktur (*genus proximum*), die das zusätzliche Merkmal „durchgängig subintegriert“ (*differentia specifica*) aufweisen.

Die Termini ‚unidirektionaler Verweis‘ und ‚unidirektionale Verknüpfung‘ bezeichnen kategorienähnliche Konzepte, gehören aber unterschiedlichen terminologischen Systemen an und sind in einem gegebenen Kontext deshalb nicht ohne Weiteres füreinander austauschbar. Die begrenzte Austauschbarkeit hat dabei zum einen ausdrucksseitige Gründe: In der Regel wird ein Autor, der zur Gruppe der Experten in der Fachdomäne gehört, über die er schreibt, nicht willkürlich Termini aus unterschiedlichen terminologischen Systemen mischen, ohne dies entsprechend zu markieren oder zu motivieren (in Texten von Semi-Experten, z.B. Seminararbeiten von Studierenden, kommt dies hingegen durchaus vor – dort dann häufig unbeabsichtigt, da nicht erkannt wird, dass es in der betreffenden Domäne konkurrierende terminologische Systeme gibt). Die begrenzte Austauschbarkeit hat aber auch konzeptuelle Gründe: *Kategorienähnlichkeit* ist – aufgrund des schon erwähnten Zusammenhangs von Benennung und Konzeptualisierung – nicht mit *Kategorienidentität* gleichzusetzen. Konzeptuelle Strukturierung determiniert kategoriale Ordnung; die Gesamtheiten der durch ‚unidirektionaler Verweis‘ und ‚unidirektionale Verknüpfung‘ denotierten Instanzen weisen sicherlich eine beträchtliche Schnittmenge auf, eine vollständige Übereinstimmung dürfte hingegen schwierig nachweisbar sein. Selbst wenn die Definitionen der beiden Termini in den betreffenden terminologischen Systemen identisch wären und die beiden terminologischen Systeme den Sachverhaltsbereich auch ansonsten absolut identisch strukturierten, bliebe letztlich immer noch die unterschiedliche Benennungsmotivik, die den in Rede stehenden Sachverhaltsausschnitt einmal unter dem Aspekt des Verknüpfens (als Lehnübersetzung von engl. ‚link‘) und einmal unter dem Aspekt des Verweisen konzeptualisiert.

Während Termini in *TERMNET II* als Klassen modelliert und, wenn sie demselben terminologischen System zugehören, anhand „harter“ ontologischer, zwischen Klassen definierter Relationen (*superclass-subclass*-Relation, *Disjunktivität*) zueinander in

Beziehung gesetzt werden, bietet sich für die Darstellung von Kategorienähnlichkeiten zwischen Termini aus *unterschiedlichen* terminologischen Systemen eine solch „harte“ Modellierung gerade nicht an. Kategorienähnliche Termini unterschiedlicher Systeme – wie z.B. ‚unidirektionaler Verweis‘ und ‚unidirektionale Verknüpfung‘ oder ‚Proterm‘ und ‚Pronomen‘ – weisen zwar denotativ charakteristische Schnittbereiche auf, konstituieren aber nicht dieselben Klassen von Referenten in der Welt. Für die Modellierung von Kategorienähnlichkeiten erscheint daher gerade der „weiche“ Zuschnitt des Modellierungskonzepts *Synset* aus dem Wortnetz-Ansatz als besonders geeignet – bezogen allerdings nicht, wie in allgemeinsprachlichen Wortnetzen, auf Mengen von lexikalischen Einheiten, die (in einem weiten Sinne synonymisch) dasselbe Konzept lexikalisieren, sondern übertragen auf die in konkurrierenden terminologischen Systemen durch systemspezifische terminologische Strukturierung konstituierte Konzepte, die ähnliche Kategorien (Sachverhaltsausschnitte) konzeptualisieren. Dieses Modellierungskonstrukt nennen wir *Termset*. *Termsets* repräsentieren Kategorien; Beziehungen zwischen *Termsets* werden durch „weiche“ Wortnetz-Relationen ausgedrückt – die dann allerdings nicht, wie in allgemeinsprachlichen Wortnetzen und in *TERMNET I*, Beziehungen zwischen Konzepten, sondern Beziehungen zwischen (durch u.U. konkurrierende Konzepte konstituierte) Kategorien beschreiben. Zentrale Relationen sind hier die *Hyponymie/Hyperonymie*- und die *Meronymie/Holonymie*-Relation. Das Modellierungskonstrukt des *Synsets* findet in *TERMNET II* lediglich implizite Anwendung, kann aber im Bedarfsfall anhand der zwischen einzelnen Termini etablierten lexikalischen Relationen aus der Datenbasis rekonstruiert werden (vgl. oben, 2.2).

Die Unterschiede zwischen den Modellierungskonstrukten *Synset* und *Termset* in *TERMNET II* lassen sich abschließend wie folgt zusammenfassen:

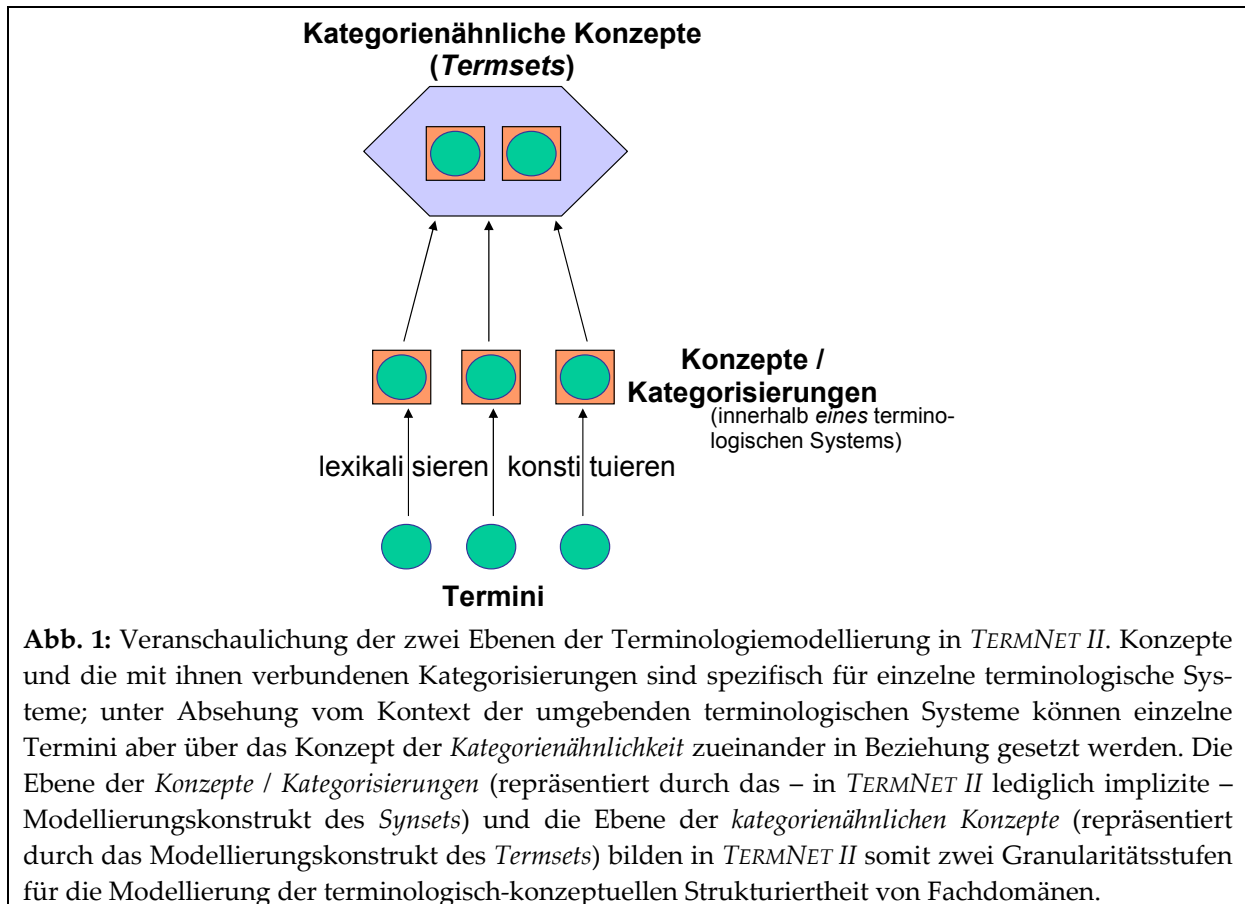
***Synset:***

- (1) Synsets repräsentieren **Konzepte** bzw. **Kategorisierungen**.
- (2) Die enthaltenen Einheiten sind **lexikalische Einheiten** (einzelne **Termini**).
- (3) Die enthaltenen Einheiten sind **synonym**.

***Termset:***

- (1) Termsets repräsentieren **Kategorien**.
- (2) Die enthaltenen Einheiten sind **terminologische Konzepte aus verschiedenen Terminologien, die extensional einen relativ großen Schnittbereich haben**.
- (3) Die enthaltenen Einheiten sind **kategorienähnlich**.

Die Einheiten von *Termsets* sind im Gegensatz zu *Synset*-Einheiten in einem Text nicht gegeneinander austauschbar.

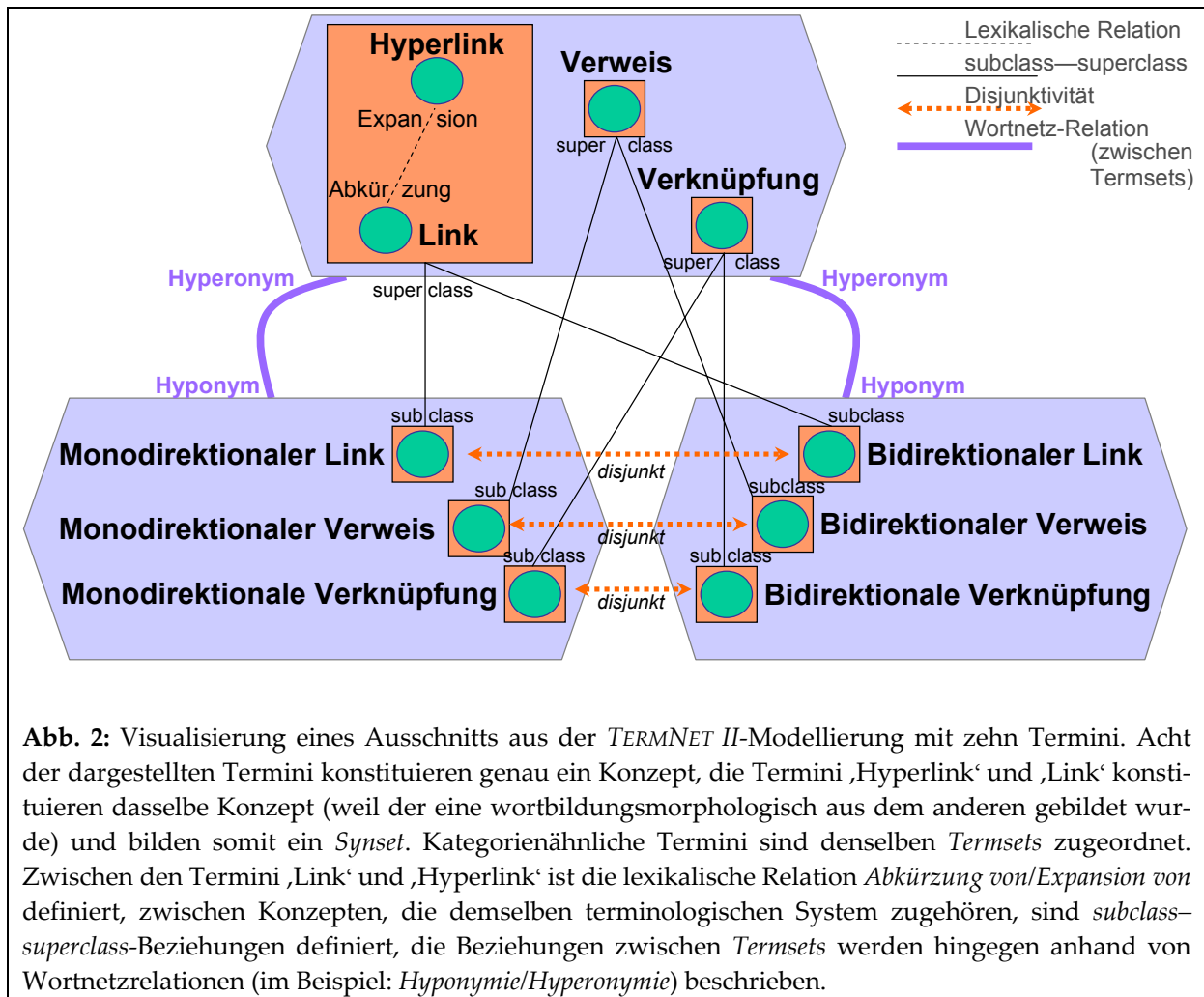


Bezogen auf das o.a. Beispiel aus der Fachsprache der Grammatik, in der diverse konkurrierende terminologische Strukturierungen des Sachverhaltsbereichs „Wortarten“ existieren, würde eine Modellierung anhand des *TERMNET II*-Modells (gegenüber der Modellierung anhand eines allgemeinsprachlich zugeschnittenen Wortnetzmodells) die Repräsentation der folgenden fachsprachlichen Wissensausschnitte zu den Termini ‚Pronomen‘ und ‚Proterm‘ sowie zu terminologischen Systemen im Bereich der Wortartenklassifikation erlauben:

- Dass der Terminus ‚Proterm‘ einem anderen terminologischen System angehört als der Terminus ‚Pronomen‘, wird dadurch dargestellt, dass zu den beiden Termini im Netzmodell kein gemeinsamer übergeordneter Knoten existiert, der ebenfalls den Status eines Terminus hat;
- dass die beiden Termini trotzdem ähnliche Kategorien konzipieren, wird durch die Zuordnung zu demselben Termset deutlich;
- wie die jeweils spezifischen Konzeptualisierungen der Kategorien *PROTERM* und *PRONOMEN* beschaffen sind und worin sie sich (und damit auch die Kategorien) unterscheiden, kann in einer Wortnetzmodellierung nicht explizit ausgedrückt werden; in Kombination mit annotierten Fachtextkorpora (wie sie das *HyTex*-Projekt verwendet) kann auf Basis der *TERMNET*-Ressource aber eine mit Definitionen in den Korpustexten verlinkte Glossaranwendung erzeugt werden; diese erlaubt den Aufruf der zugehörigen Definitionen zu den in einem Termset „vereinten“, aber verschiedenen terminologischen Systemen zugehörigen Termini und somit die Erarbeitung der Feinheiten der unterschiedlichen Konzeptualisierungen



und Kategorisierungen, die zwei (oder mehrere) Systeme zur Strukturierung desselben Sachverhaltsbereichs vornehmen.



**Abb. 2:** Visualisierung eines Ausschnitts aus der *TERMNET II*-Modellierung mit zehn Termini. Acht der dargestellten Termini konstituieren genau ein Konzept, die Termini ‚Hyperlink‘ und ‚Link‘ konstituieren dasselbe Konzept (weil der eine wortbildungsmorphologisch aus dem anderen gebildet wurde) und bilden somit ein *Synset*. Kategorienähnliche Termini sind denselben *Termsets* zugeordnet. Zwischen den Termini ‚Link‘ und ‚Hyperlink‘ ist die lexikalische Relation *Abkürzung von/Expansion von* definiert, zwischen Konzepten, die demselben terminologischen System zugehören, sind *subclass-superclass*-Beziehungen definiert, die Beziehungen zwischen *Termsets* werden hingegen anhand von Wortnetzrelationen (im Beispiel: *Hyponymie/Hyperonymie*) beschrieben.

### 3. Repräsentation von *TERMNET II* in OWL (Web Ontology Language)

Die drei wichtigsten Konstrukte, anhand derer OWL-Ontologien aufgebaut werden, sind *Klassen*, *Individuen* und *Properties*. *Klassen* können miteinander über hierarchische Subklassen-Beziehungen verbunden sein; *Individuen* hingegen werden als konkrete Instanzierungen einer Klasse verstanden. *Properties* sind Relationen zwischen Klassen und können ihrerseits wiederum Instanzen haben (nämlich konkrete Instanzierungen von Klassen-*Properties* für zwei Individuen, die Instanzierungen derjenigen Klassen darstellen, zwischen denen die Property definiert ist).

Für die Repräsentation von *TERMNET II* in OWL wurden die *TERMNET*-spezifischen Modellierungskonstrukte wie folgt anhand der OWL-Konstrukte abgebildet:

- Termsets werden als OWL-Klassen modelliert, die zwischen Termsets etablierten „weichen“ Wortnetz-Relationen werden als *Properties* dargestellt, die die jewei-

- ligen Relationstypen (z.B. *isHyponymOf* mit der Inverse *isHyperonymOf* sowie *isHolonymOf* und *isMeronymOf*) im Namen tragen und die nur zwischen Klassen vom Typ *NounTermSet* etabliert werden können;
- Termini werden ebenfalls als Klassen konstruiert; um sie von den Termset-Klassen unterscheidbar zu halten, werden mit *NounTermSet* und *NounTerm* zwei Superklassen etabliert, die über eine Property *hasMember* (Inverse: *isMemberOf*) miteinander verbunden sind; die Property *hasMember* / *isMemberOf* repräsentiert hierbei die Zuordnung einzelner Termini zu Termsets;
  - Beziehungen innerhalb eines terminologischen Systems werden als „harte“ *subclass/superclass*-Relationen interpretiert, für deren Darstellung in OWL ein vordefinierter Konstruktor existiert; die Unterschiede der Beziehungen zwischen Termsets und zwischen den *Termini* eines einzelnen terminologischen Systems bleiben somit in der OWL-Repräsentation erhalten;
  - wortbildungsmorphologische Beziehungen zwischen Termini (*isAbbreviationOf* mit der Inverse *isExpansionOf*) werden als Properties dargestellt, die nur zwischen Klassen vom Typ *NounTerm* etabliert werden können.

Eine ausführliche Motivation der für die Modellierung von *TERMNET II* gewählten OWL-Repräsentation und ein Vergleich zum OWL-Modell des allgemeinsprachlichen Wortnetzes *GERMANET* findet sich in Kunze, Lemnitzer, Längen & Storrer (2007); eine ausführliche Dokumentation zur OWL-Modellierung von *TERMNET II* gibt Selzam (2008).

Über eine Reihe von „Plugin“-Relationen, anhand derer sich einzelne Termsets aus *TERMNET II* zu ihren konzeptuell – cum grano salis – korrespondierenden Synsets in *GERMANET* in Beziehung setzen lassen, wurde weiterhin eine Möglichkeit geschaffen, die OWL-Modelle von *TERMNET II* und *GERMANET* miteinander zu verknüpfen (vgl. Längen, Kunze, Lemnitzer & Storrer 2008). Verschiedene OWL-Modelle von *TERMNET II* und *GERMANET* als auch ein integriertes Modell („GermaTermNet“), in dem beide Ressourcen verknüpft sind, werden unter der URL <http://www.wordnets-in-owl.de> zur Verfügung gestellt.

## Erwähnte Literatur

- Beißwenger, Michael (2004): Arbeitsbericht: Annotation definitorischer Textsegmente und terminologiesensitives Linking. WWW-Ressource: [http://www.hytext.info/050\\_publicationen/deflink.pdf](http://www.hytext.info/050_publicationen/deflink.pdf)
- Beißwenger, Michael, Angelika Storrer & Maren Runte: (2003): Modellierung eines Terminologienetzes für das automatische Linking auf der Grundlage von WordNet. In: Claudia Kunze, Lothar Lemnitzer & Andreas Wagner (Hrsg.): Anwendungen des deutschen Wortnetzes in Theorie und Praxis (LDV-Forum 19/1-2), 113-125. - PDF-Preprint: [http://www.hytext.info/050\\_publicationen/TermNet-LDV.pdf](http://www.hytext.info/050_publicationen/TermNet-LDV.pdf)

- Beißwenger, Michael & Sandra Wellinghoff (2006): Dokumentation: Inhalt und Zusammensetzung des Fachtextkorpus. WWW-Ressource: [http://www.hytext.info/050\\_publicationen/Korpuszusammensetzung.pdf](http://www.hytext.info/050_publicationen/Korpuszusammensetzung.pdf)
- Dean, Mike & Guus Schreiber (Eds., 2004): OWL Web Ontology Language Reference. W3C Recommendation 10 February 2004. WWW-Ressource: <http://www.w3.org/TR/owl-ref/>
- Fellbaum, Christiane (Ed., 1998): *WordNet: An Electronic Lexical Database*. Cambridge, MA/London: MIT Press.
- Kunze, Claudia (2001): Lexikalisch-semantische Wortnetze. In: Carstensen, K.-U. et al. (eds.): *Computerlinguistik und Sprachtechnologie: Eine Einführung*. Heidelberg/Berlin: Spektrum, 386-393.
- Kunze, Claudia, Lothar Lemnitzer, Harald Lüngen, Angelika Storrer (2007): Repräsentation und Verknüpfung allgemeinsprachlicher und terminologischer Wortnetze in OWL. *Zeitschrift für Sprachwissenschaft* 26, 267-290.
- Lenz, Eva Anna, Michael Beißwenger & Sandra Wellinghoff (2006): Dokumentation: Annotationsschicht: Definitionen und Termverwendungsinstanzen. WWW-Ressource: [http://www.hytext.info/050\\_publicationen/Annotation-Termini-und-Definitionen.pdf](http://www.hytext.info/050_publicationen/Annotation-Termini-und-Definitionen.pdf)
- Lenz, Eva Anna, Benjamin Birkenhake & Jan Frederik Maas (2003): Von der Erstellung zur Nutzung: Wortnetze als XML Topic Maps. In: Claudia Kunze, Lothar Lemnitzer & Andreas Wagner (Hrsg.): *Anwendungen des deutschen Wortnetzes in Theorie und Praxis (LDV-Forum 19/1-2)*, 127-136. - PDF-Preprint: [http://www.hytext.info/050\\_publicationen/tue2-paper.pdf](http://www.hytext.info/050_publicationen/tue2-paper.pdf)
- Lüngen, Harald, Claudia Kunze, Lothar Lemnitzer, Angelika Storrer (2008): Towards an Integrated OWL Model for Domain-Specific and General Language WordNets. In: *Proceedings of the Global WordNet Conference 2008, Szeged, Ungarn*. WWW-Ressource: [http://www.inf.u-szeged.hu/projectdirs/gwc2008/GWC2008\\_Proceedings\\_Final.pdf](http://www.inf.u-szeged.hu/projectdirs/gwc2008/GWC2008_Proceedings_Final.pdf).
- Pepper, Steve & Graham Moore (Eds., 2001): XML Topic Maps (XTM) 1.0. TopicMaps.Org Specification. WWW-Ressource: <http://www.topicmaps.org/xtm/1.0/>
- Runte, Maren (2004): Dokumentation: Modellierung des terminologischen Wissensnetzes *TermNet*. Mit einem Anhang: Der WordNet-Ansatz – ein Überblick. WWW-Ressource: [http://www.hytext.info/050\\_publicationen/termnet-doku.pdf](http://www.hytext.info/050_publicationen/termnet-doku.pdf)
- Selzam, Bianca (2008): Modellierung des *TermNet* in OWL. Technical Report. WWW-Ressource: [http://www.hytext.info/050\\_publicationen/termnet-in-owl.pdf](http://www.hytext.info/050_publicationen/termnet-in-owl.pdf)
- Storrer, Angelika (i.Dr.): Mark-up driven strategies for text-to-hypertext conversion. Erscheint in: Dieter Metzger & Andreas Witt (Eds.): *Linguistic Modelling of Information and Markup Languages*. Dordrecht: Springer. PDF-Preprint: [http://www.storrer.fb15.uni-dortmund.de/10\\_team/as-publicationen/hytext-preprint.pdf](http://www.storrer.fb15.uni-dortmund.de/10_team/as-publicationen/hytext-preprint.pdf)
- Wellinghoff, Sandra (2006): Dokumentation: Manuelle Annotation definitorischer Textsegmente incl. Guidelines Phase I und II. WWW-Ressource: [http://www.hytext.info/050\\_publicationen/Doku%20Manuelle%20Annotation%20Wellinghoff.pdf](http://www.hytext.info/050_publicationen/Doku%20Manuelle%20Annotation%20Wellinghoff.pdf)