



Guidelines für die manuelle Annotation definitorischer Textsegmente

Projekt
**Hypertextualisierung auf
textgrammatischer Grundlage**
(www.hytext.info)

Michael Beißwenger
Sandra Wellinghoff

2005

Als Basis für die Implementierung von Strategien zur Verlinkung von in den Fachtexten des Projektkorpus verwendeten Termini mit den jeweils für ihre korrekte Semantisierung maßgeblichen definitonischen Textsegmenten wurden in Projektphase I (2001-2005) alle in einem Ausschnitt des Fachtextkorpus vorkommenden definitonischen Textsegmente manuell annotiert. Der annotierende Projektmitarbeiter war in Bezug auf die in den Fachtexten behandelten Gegenstände (Hypertextforschung und Texttechnologie) fachkompetent. Die Annotation erfolgte auf der Grundlage eines Annotationsschemas, das im Arbeitsbericht Beißwenger (2004)¹ dokumentiert ist.

Um unsere Strategien zur Verlinkung von im Text vorkommenden Termini mit den zugehörigen Definitionen auf andere Fachtextkorpora übertragen zu können, suchten wir in Projektphase II nach Verfahren, mit denen sich die bislang manuelle Annotation definitonischer Textsegmente automatisieren lässt. Um einen Bezugspunkt für die Evaluation der automatischen Verfahren zu haben, haben hierzu zunächst 5 Hilfskräfte jeweils 11 Fachtexte aus dem HyTex-Korpus manuell annotiert. Die Grundlage für die Annotation bildete ein Annotations-Leitfaden, der die Funktion und Struktur definitonischer Textsegmente so beschreibt, dass sie den bei der manuellen Annotation in Phase I zugrundegelegten Kriterien für das Vorliegen und die Struktur definitonischer Textsegmente entspricht.

Dabei zeigte sich, dass in Zweifelsfällen nur eine niedrige Inter-Annotator-Übereinstimmung erzielt werden konnte, was nicht zuletzt an der Vielfalt der sprachlichen Mittel liegt, die in Fachtexten zur Bedeutungsexplikation eingesetzt werden. Um die Inter-Annotator-Übereinstimmung zu verbessern, wurde der Leitfaden zur Annotation überarbeitet. Insbesondere wurde Wert darauf gelegt, das Konzept ‚Definition‘ besser intersubjektiv vermittelbar zu machen und den Annotatoren mehr Kriterien an die Hand zu geben, die zur Identifikation vertexteter Definitionen herangezogen werden können. Des Weiteren beschränken wir uns auf Definitionen, deren Definiens nach dem Schema *Genus proximum und Differentia specifica* strukturiert ist.

Dieses Dokument umfasst beide Versionen der Guidelines für die manuelle Annotation von Definitionen:

- die in Anlehnung an Beißwenger (2004) formulierten und einem ersten Annotations-experiment zugrunde gelegten Guidelines (S. 3-8);

1 Beißwenger, M. (2004): Annotation definitonischer Textsegmente und terminologiesensitives Linking. Interner Report der DFG-Forschergruppe 437 „Texttechnologische Informationsmodellierung“. <http://coli.lili.uni-bielefeld.de/Texttechnologie/Forschergruppe/publications/hytex-deflink.pdf>.

- die von Sandra Wellinghoff auf Basis der Ergebnisse des ersten Experiments erstellte Neufassung der Guidelines (2005), die den weiteren Annotationsexperimenten zugrunde gelegt werden soll (S. 9-18).

Guidelines für die Annotation von Definitionen: Fassung I (2004)

Ein Textsegment (ein Syntagma, ein Satz, mehrere Sätze) heißt dann *definitorisches Textsegment*, wenn der Autor damit bezweckt, dem Leser offen zu legen, welches Konzept einem von ihm oder anderen Autoren verwendeten Terminus unterliegt.

Wir können das Konzept der *Definition* in Anschluss an Lay (1971) ganz allgemein zunächst einmal wie folgt definieren:

Definitionen sind Aussagen, die in der Absicht formuliert werden, den Sinn und/oder die Bedeutung eines Wortes festzulegen. Mitunter sind Definitionen Verkürzungstechniken: Es wird festgelegt, daß für eine bestimmte Aussage verkürzend ein bestimmter Term stehen soll. In jedem Fall hat aber eine eigentliche Definition zur notwendigen und hinreichenden Bedingung ihrer Legitimität, daß es möglich ist, statt ihrer in allen Aussagen und Aussageverbindungen, in denen sie vorkommt, das definiens einzusetzen, ohne daß sich Sinn und Bedeutung der Aussage wandeln (Kriterium Pascals). (Lay 1971: 260)

Definieren ist in diesem Zusammenhang als ein Zweckbereich sprachlichen Handelns in Fachtexten zu verstehen: Wenn ein Autor einen Terminus *definiert*, dann geht es ihm darum, seinem Adressaten diejenigen Informationen zu geben, die er benötigt, um diesen Terminus begrifflich in seinem Fachwortschatz zu verorten.

Struktur prototypischer Definitionen (und Möglichkeiten ihrer Vertextung)

Prototypische Definitionen bestehen in der Regel aus drei Grundeinheiten:

- (1) dem *definiendum* (Pl.: *definienda*): Das *definiendum* repräsentiert diejenige Einheit, die es zu definieren gilt (also einen sprachlichen Ausdruck oder eine Klasse von Dingen)
- (2) dem *definiens* (Pl.: *definientes*): Das *definiens* gibt eine Beschreibung eines Konzepts oder eine Charakteristik einer Klasse von Dingen
- (3) dem *Definitor*: Der Definitor etabliert die Zuordnungsbeziehung zwischen *definiens* und *definiendum*.

In Texten sind *definiens* und *definiendum* immer sprachliche Einheiten. Der *Definitor* kann eine lexikalische Einheit sein, kann aber auch durch eine bestimmte syntaktische Struktur realisiert werden.

Ist der Definitor lexikalisch realisiert, dann ist er realisiert durch ein Verb im Rahmen einer Prädikativkonstruktion oder (im Rahmen von Parenthesen) durch den Adjunktor *als*.

1. Der Definitor

Ausdrücke bzw. Kennzeichnungen, die in Fachtexten häufig als Definitoren verwendet werden, sind:

bezeichnen als

Definitorisches Textsegment:

Computerverwaltete Zuordnungen zwischen Ankern bezeichne ich im Folgenden als *Links*.

Definiendum	Links
Definitor	<i>bezeichne als</i>
Definiens	Computerverwaltete Zuordnungen zwischen Ankern

verstehen unter**Definitorisches Textsegment:**

Unter einem Link versteht man eine computerverwaltete Zuordnung zwischen Ankern.

Definiendum	Link
Definitor	<i>verstehen unter</i>
Definiens	eine computerverwaltete Zuordnung zwischen Ankern

heißen**Definitorisches Textsegment:**

Eine computerverwaltete Zuordnung zwischen Ankern heiße im Folgenden *Link*.

Definiendum	Link
Definitor	<i>heiße</i>
Definiens	Eine computerverwaltete Zuordnung zwischen Ankern

nennen**Definitorisches Textsegment:**

Computerverwaltete Zuordnungen zwischen Ankern werden von Storrer (2001) *Links* genannt.

Definiendum	Links
Definitor	<i>werden genannt</i>
Definiens	Computerverwaltete Zuordnungen zwischen Ankern

definieren**Definitorisches Textsegment:**

Storrer (2001) definiert Links als computerverwaltete Zuordnungen zwischen Ankern.

Definiendum	Links
Definitor	<i>definiert</i>
Definiens	computerverwaltete Zuordnungen zwischen Ankern

ist**Definitorisches Textsegment:**

Ein Link ist eine computerverwaltete Zuordnung zwischen Ankern.

Definiendum	Link
Definitor	<i>ist</i>
Definiens	eine computerverwaltete Zuordnung zwischen Ankern

sind**Definitorisches Textsegment:**

Links sind computerverwaltete Zuordnungen zwischen Ankern.	
Definiendum	Links
Definitor	<i>sind</i>
Definiens	computerverwaltete Zuordnungen zwischen Ankern

als (im Rahmen von Parenthesen)

Textsegment: Links als computerverwaltete Zuordnungen zwischen Ankern sind zentrale Konstituenten von Hypertexten.	
Definitorisches Teilsegment: Links als computerverwaltete Zuordnungen zwischen Ankern	
Definiendum	Links
Definitor	<i>als</i>
Definiens	computerverwaltete Zuordnungen zwischen Ankern

Ist der Definitor nicht lexikalisch realisiert, so ist er als syntaktische Struktur eines bestimmten Typs (Parenthese / Apposition), inklusive entsprechender Interpunktionsmittel, gegeben.

In solchen Fällen ist die Definition in einen Satz eingebettet, der als Ganzer nicht definitorisch ist. Die Definition erstreckt sich in diesem Fall lediglich auf die Phrase, in deren Rahmen durch die appositive Beigabe eines Definiens der im Phrasenkopf stehende Ausdruck definiert wird. Dies kann prinzipiell sowohl bei substantivischen als auch bei adjektivischen Termini der Fall sein.

Parenthese, gekennzeichnet durch Klammerung

Textsegment: Links (also computerverwaltete Zuordnungen zwischen Ankern) können eine tolle Sache sein.	
Definitorisches Teilsegment: Links (also computerverwaltete Zuordnungen zwischen Ankern)	
Definiendum	Links
Definitor	(also ...)
Definiens	computerverwaltete Zuordnungen zwischen Ankern

Parenthese, gekennzeichnet durch Gedankenstriche

Textsegment: Links – also computerverwaltete Zuordnungen zwischen Ankern – können eine tolle Sache sein.	
Definitorisches Teilsegment: Links – also computerverwaltete Zuordnungen zwischen Ankern –	
Definiendum	Links
Definitor	–also ...–
Definiens	computerverwaltete Zuordnungen zwischen Ankern

Parenthese, gekennzeichnet durch Kommata

Textsegment: Links, also computerverwaltete Zuordnungen zwischen Ankern, können eine tolle Sache sein.	
Definitorisches Teilsegment:	

Links, also computerverwaltete Zuordnungen zwischen Ankern	
Definiendum	Links
Definitor	,also ... ,
Definiens	computerverwaltete Zuordnungen zwischen Ankern

Parenthetische Definition eines adjektivischen Terminus, gekennzeichnet durch Kommata und *also*

Textsegment: Nicht alle XML-Dokumente müssen valide, also konform zu einer DTD, sein.	
Definitorisches Teilsegment: valide, also konform zu einer DTD	
Definiendum	valide
Definitor	,also ... ,
Definiens	konform zu einer DTD

Im Falle von Mehrwort-Termini kann das appositiv beigegebene definiens auch zwischen den Konstituenten des definiendums stehen:

Parenthetische Definition bei Mehrwort-Termini mit diskontinuierlicher Abfolge des definiendums, gekennzeichnet durch Kommata und *also*

Textsegment: In unserem Projekt haben wir es ausschließlich mit validen, also zu einer DTD konformen XML-Dokumenten zu tun.	
Definitorisches Teilsegment: validen, also zu einer DTD konformen XML-Dokumenten	
Definiendum	validen XML-Dokumenten
Definitor	, also ... ,
Definiens	zu einer DTD konformen

In diesem Beispiel ist das definiens, insofern es als Attribut zum selben Nomen realisiert ist, das zugleich einen Teil des definiendums repräsentiert, nicht für sich alleine verständlich. Die Definition würde unverständlich, wenn man sie extrahierte und in folgende Form brächte:

„valide XML-Dokumente“ = „zu einer DTD konformen“

Dies zeigt zugleich die vielfältigen Varianten der Vertextung von Definitionen.

2. Das *definiendum* (Pl.: *definienda*)

In unserer Untersuchung fokussieren wir auf Definitionen mit nominalen (ggf. auch adjektivischen) *Definienda*.

Das *definiendum* ist ein Nomen oder eine Nominalphrase, mit der generisch referiert wird oder ein anaphorischer Ausdruck, der auf ein Nomen oder eine Nominalphrase im Vortext verweist. Es kann der Fall sein, dass *eine* Definition *zwei* (oder mehrere) *Definienda* enthält (die damit zugleich als Synonyme ausgewiesen werden).

Beispiele:

Definitorisches Textsegment mit *einem* Definiendum (Standardfall)

Definitorisches Textsegment: Unter einem Link versteht man eine computerverwaltete Zuordnung zwischen An-

kern.	
Definiendum	Link
Definitor	versteht unter
Definiens	eine computerverwaltete Zuordnung zwischen Ankern

Definitorisches Textsegment mit *zwei* Definienda

Definitorisches Textsegment: Unter einem Link (auch: Knoten) versteht man eine computerverwaltete Zuordnung zwischen Ankern.	
Definienda	Link, Knoten
Definitor	versteht unter
Definiens	eine computerverwaltete Zuordnung zwischen Ankern

Anapher in Definiendum-Position

Textsegment: Ein konstitutives Element von Hypertexten sind die sog. <i>Hyperlinks</i> . Darunter versteht man in der Hypertextliteratur computerverwaltete Zuordnungen zwischen Ankern.	
Definitorisches Teilsegment: Darunter versteht man in der Hypertextliteratur computerverwaltete Zuordnungen zwischen Ankern.	
Definiendum	Darunter
Definitor	versteht man
Definiens	computerverwaltete Zuordnungen zwischen Ankern

3. Das *definiens* (Pl.: *definientes*)

Als Definiens fungiert ein sprachlicher Ausdruck, mit dem ein Konzept charakterisiert oder konstitutive Merkmale für eine Klasse von Dingen gegeben werden; häufig in der Form:

- Angabe eines Oberbegriffs (*genus proximum*) plus
- Angabe eines oder mehrerer spezifizierender Merkmale (*differentia specifica*)

z.B.:

- „ein weißes Pferd“
- „ein Schiff, das zum Hotel umgebaut wurde“
- „eine computerverwaltete Zuordnung zwischen Ankern“

Als Definiens kann auch ein anaphorischer Ausdruck fungieren, der auf einen entsprechenden Ausdruck im Vortext verweist. Nicht selten tritt das Definiens *diskontinuierlich* auf.

Beispiele:

Definitorisches Textsegment mit *kontinuierlichem* Definiens

Definitorisches Textsegment: Unter einem Link versteht man eine computerverwaltete Zuordnung zwischen Ankern.	
Definiendum	Link
Definitor	versteht unter
Definiens	eine computerverwaltete Zuordnung zwischen Ankern

Definitorisches Textsegment mit *diskontinuierlichem* Definiens

Definitorisches Textsegment: Unter „Annotationen“ werden in der Hypertextliteratur Anmerkungen und Notizen verstanden, die ein Hypertextnutzer während des Rezeptionsvorgangs zu den Inhalten eines Moduls anbringt.	
Definiendum	„Annotationen“
Definitor	unter werden verstanden
Definiens	Anmerkungen und Notizen ..., die ein Hypertextnutzer während des Rezeptionsvorgangs zu den Inhalten eines Moduls anbringt

Anapher (rückverweisender Ausdruck) in Definiens-Position

Textsegment: Ein konstitutives Element von Hypertexten sind computerverwaltete Zuordnungen zwischen Ankern. Diese bezeichne ich als <i>Links</i> .	
Definitorisches Teilsegment: Diese bezeichne ich als <i>Links</i> .	
Definiendum	<i>Links</i>
Definitor	bezeichne als
Definiens	Diese

Es kommt auch vor, dass in *einem* definitorischen Textsegment *zwei* unterschiedliche Definiens zu ein- und demselben Definiendum gegeben werden. In solchen Fällen haben wir es mit der Einbettung *zweier unterschiedlicher Definitionen* in *eine syntaktische Struktur* (uin unserer Redeweise: einem *Definitons-Komplex*) zu tun:

Definitions-Komplex: Zwei alternative Definiens zu einem Definiendum

Definitorisches Textsegment: Mit dem Ausdruck „Seite“ wird in der Werkstattsprache des WWW einerseits ein am Bildschirm sichtbares Objekt, andererseits aber auch die von einem WWW-Server verwaltete Dateneinheit bezeichnet.	
Definiendum	„Seite“
Definitor	Mit wird bezeichnet
Definiens 1	ein am Bildschirm sichtbares Objekt
Definiens 2	die von einem WWW-Server verwaltete Dateneinheit

Regeln für die Kennzeichnung von Definitionen in den ausgedruckten Texten:

- Das **Definiendum** wird mit Textmarker **orange** gekennzeichnet
- Das **Definiens** (bzw. seine Teile) werden mit Textmarker **gelb** gekennzeichnet
- Der **Definitor** (bzw. seine Teile) werden mit Kugelschreiber **umrandet**
- **Relevanter Kontext** wird mit Kugelschreiber **unterstrichen**
(„Relevanter Kontext“ sind diejenigen Sätze oder Syntagmen im Vor- oder Nachtext der Definition, die verständnisrelevant sind, wenn entweder im Definiendum oder im Definiens eine Anapher bzw. Katapher steht)
- Im Falle von **Definitions-Komplexen** (s.o.) wird am Rand der Seite (auf Höhe der gekennzeichneten Definition) angemerkt: „Komplex“

Guidelines für die Annotation von Definitionen: Fassung II (2005)

1. Was versteht man unter einer Definition?	- 9 -
2. Definitionsstrukturen.....	- 10 -
3. Definitionstypen.....	- 12 -
3.1 Inhaltsdefinition	- 12 -
3.1.1 Genetische Definition	- 13 -
3.1.2 Operationale Definition	- 13 -
3.2 Umfangsdefinition.....	- 13 -
3.3 Glossardefinition.....	- 14 -
4. Grenzfälle: Erklärung, Beschreibung und Synonymie.....	- 14 -
4.1 Die Erklärung.....	- 15 -
4.2 Die Beschreibung	- 15 -
4.3 Die Synonymie	- 16 -
5. Vorgehen bei der manuellen Annotation	- 17 -
6. Erwähnte Literatur	- 17 -

1. Was versteht man unter einer Definition?

In der klassischen Definitionslehre des Aristoteles, in der erstmals Definitionsregeln festgelegt wurden,

„ ... ist eine D. [Definition] eine Abgrenzung einer Art durch Angabe der diese Art einschließenden Gattung und des artspezifischen Merkmals oder des besonderen Kennzeichens dieser Art.“²

Dieses sehr enge Definitionsverständnis, welches in Definitionen die Differenzierung zwischen Oberbegriff (*genus proximum*) und unterscheidende Merkmale (*differentiae specifica*) fordert, wurde im Rahmen wissenschaftstheoretischer Auseinandersetzungen aufgehoben, da die Auffassung vertreten wurde, dass diese Elemente austauschbar seien und man neutraler von Merkmalen sprechen sollte, aus denen ein Begriff zusammengesetzt sei.³ Tatsächlich gibt es aber noch heute Definitionsformen - wie z.B. die Inhaltsdefinition -, die diese Differenzierung fordern und somit die folgende allgemeine Definition einer Definition in der Anwendung konkretisieren:

Definition (griech. *ὀρισμός*, lat. *definitio*, ursprünglich: Umgrenzung), im weitesten Sinne jede Art der Feststellung oder Festsetzung einer Zeichenverwendung. Das zu definierende (oder definierte) Zeichen heißt *Definiendum* (oder *Definitum*), das definierende Zeichen heißt *Definiens*.“ (Mittelstraß (1995: 439))

Definitionen, in denen die Zeichenverwendung festgesetzt wird, haben zudem eine kommunikative Funktion. Denn

„ als ↑ Sprechakte betrachtet reichen sie von Willensbekundungen (z.B. in einem Vortrag ein bestimmtes Wort stets in einem bestimmten Sinne zu

² Speck (1980: 124).

³ Mittelstraß (1995: 440).

gebrauchen) und Selbstverpflichtungen – soweit der private Sprachgebrauch betroffen ist – über Aufforderungen und Empfehlungen bis zu verbindlichen Wortverwendungsnormen (z.B. in Form juristischer D.en) – soweit der öffentliche Sprachgebrauch betroffen ist.“⁴

Definieren ist in diesem Zusammenhang also als ein Zweckbereich sprachlichen Handelns zu verstehen und auf unser Fachtextkorpus bezogen kann festgehalten werden: Wenn ein Autor einen Terminus *definiert*, dann geht es ihm darum, seinem Adressaten diejenigen Informationen zu geben, die er benötigt, um diesen Terminus begrifflich in seinem Fachwortschatz zu verorten.

Ein Textsegment (ein Syntagma, ein Satz, mehrere Sätze) heißt demnach *definitorisches Textsegment*, wenn der Autor damit bezweckt, dem Leser offen zu legen, welches Konzept einem von ihm oder anderen Autoren verwendeten Terminus unterliegt. Beim Erkennen definitorischer Textsegmente ist es häufig jedoch problematisch, diese in Definitionen implizierte Absicht des Autors zu erkennen. Denn nicht in allen Fällen beginnt ein Autor eine Definition mit: „*Ich verstehe unter ...*“ oder „*Ich definiere ...*“. Zudem werden Definitionen in Fachtexten auf vielfältige Weise versprachlicht, da sich die Autoren nur in seltenen Fällen streng an den Vorgaben und Regeln der Definitionslehre orientieren. Eine weitere Schwierigkeit ist die Abgrenzung einer Definition von einer Erklärung, einer Beschreibung oder von Synonymen (vgl. Kapitel 4). Daher müssen vor diesem theoretischen Hintergrund weitere Kriterien herangezogen werden, die das Auffinden von Definitionen praktisch ermöglichen. Im Folgenden werden daher Definitionsstrukturen und -typen vorgestellt

2. Definitionsstrukturen

Mittelstraß gibt durch seine o.g. Definition zwei Strukturelemente vor, die in jeder Definition nachgewiesen werden können, nämlich das *Definiendum* und das *Definiens*. Die Zuordnungsbeziehung zwischen beiden Elementen wird durch ein Element bestimmt, das wir als *Definitor* oder – im Rahmen von Glossardefinitionen (vgl. Kapitel 3.3) – als *Separator* bezeichnen. Folgende Liste gibt einen zusammenfassenden Überblick über die Eigenschaften dieser Strukturelemente:

- ***Definiendum (Pl.: Definienda):***

Das *Definiendum* repräsentiert das definierte – oder zu definierende – Zeichen. In unserer Untersuchung fokussieren wir auf Definitionen mit nominalen (ggf. auch adjektivischen) *Definienda*. Das *Definiendum* ist ein Nomen oder eine Nominalphrase, mit der generisch referiert wird oder ein anaphorischer Ausdruck, der auf ein Nomen oder eine Nominalphrase im Vortext verweist. Es kann der Fall sein, dass eine Definition zwei (oder mehrere) *Definienda* enthält (die damit zugleich als Synonyme ausgewiesen werden).

- ***Definiens (Pl.: Definientes):***

Im *Definiens* wird das *Definiendum* definiert. Welche weiteren Einschränkungen für das *Definiens* gelten, bestimmt der jeweilige Definitionstyp.

- ***Definitor (Pl.: Definitoren) oder Separator (Pl.: Separatoren)***

Der *Definitor* etabliert die Zuordnungsbeziehung zwischen *Definiens* und *Definiendum*. In der Regel wird er durch ein Verb (evtl. + Hilfsverb) realisiert. Typische *Definitoren* sind z.B. *sein*, *bezeichnen als*, *verstehen unter*, *nennen* und *bestehen aus*. Der *Separator* dagegen etabliert die Zuordnungsbeziehung zwischen *Definiens* und *Definiendum* ausschließlich in Glossardefinitionen und wird durch Interpunktionen (z.B. Doppelpunkt, Gedankenstrich, Komma) realisiert.

Aus diesen Strukturelementen werden alle Definitionen – von einfach bis komplex – konstruiert. Die Anordnung der Elemente innerhalb einer Definition ist dabei variabel: So können

⁴ Mittelstraß (1995: 439).

in einer Definition z.B. zwei synonyme Definienda angegeben werden oder auch zwei alternative Definiendes. Zudem kann das *Definiens* kontinuierlich oder diskontinuierlich – also durch andere Elemente unterbrochen - auftreten. Des Weiteren kann sich das definitorische Textsegment sowohl über nur einen Satz (Ein-Satz-Definition) als auch über mehrere Sätze (Mehr-Satz-Definition) erstrecken; für die manuelle Annotation gilt die Vorgabe, dass ein definitorisches Textsegment incl. erklärenden Kontext (vgl. Kapitel 4.1) nicht mehr als 4 Sätze umfassen sollte (Komplexibilitätsbeschränkung).

Folgende Beispiele sollen einen Eindruck von den verschiedenen strukturellen Umsetzungsmöglichkeiten einer Definition geben:

Beispiel 1:

Ein-Satz-Definition mit prototypischer Struktur, Definiens kontinuierlich

Strukturelemente:		Definiendum	Definitor	Definiens
Definition:	Ein	Link	ist	eine computerverwaltete Zuordnung zwischen Ankern.

Beispiel 2:

Ein-Satz-Definition mit prototypischer Struktur, Definiens diskontinuierlich, da es durch den Definitor unterbrochen wird

Strukturelemente:	Definitor	Definiendum	Definitor	Definiens	Definitor	Definiens
Definition:	Unter	Annotationen	werden	Anmerkungen und Notizen	verstanden,	die ein Hypertextnutzer während des Rezeptionsvorgangs zu den Inhalten eines Moduls anbringt.

Beispiel 3:

Ein-Satz-Definition mit prototypischer Struktur und mehreren Definienda, Definiens kontinuierlich

Strukturelemente:	Definitor		Definiendum (1)	Definiendum (2)	Definitor		Definiens
Definition:	Unter	einem	Link	(auch: Knoten)	versteht	man	eine computerverwaltete Zuordnung zwischen Ankern.

Beispiel 4:

Ein-Satz-Definition mit alternativen Definiendes zu einem Definiendum (= Definitions-Komplex)

Strukturelemente:	Definitor		Definiendum	Definitor		Definiens (1)		Definiens (2)	Definitor
Definition:	Mit	dem Ausdruck	Seite	wird	in der Werkstatt-sprache des WWW einerseits	ein am Bildschirm sichtbares Objekt	andererseits aber auch	die von einem WWW Server verwaltete Dateneinheit	bezeichnet.

Beispiel 5:

Mehr-Satz-Definition mit Anapher als Definiendum-Verweis, Definiens kontinuierlich

Struktur- elemente:		Definiendum auf das verwiesen wird	Anapher als Definiendum- Verweis	Definitor		Definitor	Definiens
Definition:	Ein konstitu- tives Ele- ment von Hypertexten sind die sog.	Hyperlinks.	Diese	bezeichnet	man in der Hy- pertextliteratur	als	Computerver- waltete Zuord- nungen zwi- schen Ankern.

Beispiel 6:

Mehr-Satz-Definition mit Anapher als Definiens-Verweis, Definiens kontinuierlich

Struktur- elemente:		Definiens auf das verwiesen wird	Anapher als Defi- niens- Verweis	Definitor		Definitor	Definiendum
Definition:	Ein konsti- tutives Element von Hyper- texten sind	computerverwaltete Zuordnungen zwi- schen Ankern.	Diese	bezeichne	ich	als	Links.

Für eine genaue Identifizierung der definatorischen Textsegmente reicht das Wissen um die Strukturelemente und ihre mögliche Anordnung jedoch nicht aus. Es werden weitere Informationen benötigt, die durch den *Definitionstyp* vorgegeben werden.

3. Definitionstypen

Für die manuelle Annotation definatorischer Textsegmente haben wir uns für fünf klassische Definitionstypen entschieden, nämlich für die *Inhaltsdefinition* und ihre verwandten Formen (*genetische Definition* und *operationale Definition*), für die *Umfangsdefinition* sowie für die *Glossardefinition*, welche im Folgenden erläutert werden.

3.1 Inhaltsdefinition – auch: *analytical definition, intensional definition, generic definition*

Die Inhaltsdefinition fordert im *Definiens* die aristotelische Unterscheidung zwischen Oberbegriff (*genus proximum*) und unterscheidende Merkmale (*differentiae specifica*), eine hohe Anforderung, der nicht alle Definitionen gerecht werden. Sie „ermöglicht [...] durch Angabe der Merkmale und des Oberbegriffs die Einordnung des Konzepts in das Begriffssystem und die Abgrenzung gegenüber anderen Begriffen.“⁵

⁵ Oberholzer (2002: 9).

Die Inhaltsdefinition soll am o.g. ersten Beispiel veranschaulicht werden:

Im *Definiens* der Definition „*Ein Link ist eine computerverwaltete Zuordnung zwischen Anker*“ ist *Zuordnung* der Oberbegriff zum Definiendum *Link*. Dadurch erhalten wir die Information, dass ein Link Elemente miteinander verbindet, da eine Zuordnung zwei oder mehrere Elemente verknüpft. Da es jedoch zahlreiche verschiedenartige Zuordnungen gibt (z.B. die mathematische oder räumliche Zuordnung), werden weitere Kriterien – Merkmale – benötigt, die das Definiendum näher spezifizieren. Das Definiens liefert diese einschränkenden Merkmale, indem die Zuordnung als *computerverwaltet* bestimmt wird und indem ausgesagt wird, dass die Zuordnung *zwischen Anker* erfolgt. Damit ist das Definiendum bestimmt und kann gegenüber anderen Zuordnungs-Begriffen abgegrenzt werden.

Strukturelemente:	Definiendum	Definitior	Definiens		
Definition:	Ein	Link	ist	eine computerverwaltete	Zuordnung zwischen Anker.
				Merkmal	Oberbegriff
				<i>differentiae specificae</i>	<i>genus proximum</i>
					Merkmal
					<i>differentiae specificae</i>

3.1.1 Genetische Definition – auch: *definitio genetica sive causalis*

Dieser Definitionstyp ist eine spezielle Form der Inhaltsdefinition. Ziel der genetischen Definition ist es, Vorgänge bzw. Ergebnisse von Vorgängen oder die Herkunft eines Gegenstandes zu definieren, was mit dem folgenden Beispiel verdeutlicht werden soll, in welchem die Erzeugung einer History-Liste definiert wird:

Strukturelemente:	Definiendum	Definitior	Definiens		
Definition:	Eine	History-Liste	ist	eine Liste,	die während einer aktuellen Internetsitzung durch die Aufzeichnung bereits besuchter Seiten erzeugt wird.
				Oberbegriff	Merkmale
				<i>genus proximum</i>	<i>differentiae specificae</i>

3.1.2 Operationale Definition

Die operationale Definition ist ebenfalls mit der Inhaltsdefinition verwandt. In diesem Definitionstyp wird das Definiendum durch die Angabe der einzelnen Operationen definiert, die für seine Erzeugung notwendig sind, wie folgendes Beispiel zeigt:

„Ein Verweis zu einer anderen Internetseite (Hyperlink) wird durch die Verwendung des a-Elements in Kombination mit dem href-Attribut im HTML-Quelltext erzeugt. Dem href-Attribut wird dabei das Verweisziel zugeordnet, welches durch einen beliebigen Text zwischen den a-Elementen benannt werden kann. Dieser Text erscheint dem Internetseitenbesucher anschließend als farbiger oder unterstrichener Link.
 Beispiel: `Hier gehts zur Google-Suche!`.“

3.2 Umfangsdefinition

Die Umfangsdefinition ist durch eine ihr immanente Meronymiebeziehung, also einer Ganzes-Teil-Beziehung, gekennzeichnet. So werden zum Definiendum, welches gleichzeitig einen Oberbegriff repräsentiert, alle Unterbegriffe bzw. Teile auf gleicher Unterteilungsebene angegeben, welche das Definiendum umfasst. In vielen Fällen kann der Definitior *bestehen*

aus ein Indiz für eine Umfangsdefinition sein. Zudem wird in diesem Definitionstyp oft eine exakte Angabe über die Anzahl der Unterbegriffe (z.B. *drei Komponenten*) gegeben, bevor diese aufgezählt werden, wie die folgende Definition veranschaulicht:

In der Umfangsdefinition „Nach Kuhlen (1991) besteht ein Hypertextsystem aus drei Komponenten: Produktionskomponente, womit ein Hypertext erstellt werden kann, der Hypertextbasis, gemeinhin bezeichnet als Hypertext, und der Nutzerkomponente, die den HTML-Quelltext in lesbarer und navigierbarer Form anzeigt.“ werden zum Definiendum *Hypertextsystem* (= Oberbegriff) *drei Komponenten* (= Unterbegriffe/Teile) aufgezählt, die ihn definieren.

Strukturelemente:		Definitor		Definiendum	Definitor	Definiens
Definition:	Nach Kuhlen (1991)	besteht	ein	Hypertextsystem	aus	drei Komponenten: Produktionskomponente, womit ein Hypertext erstellt werden kann, der Hypertextbasis, gemeinhin bezeichnet als Hypertext, und der Nutzerkomponente, die den HTML-Quelltext in lesbarer und navigierbarer Form anzeigt.
				Oberbegriff		Merkmal
				<i>genus proximum</i>		<i>differentiae specificae</i>

3.3 Glossardefinition

Die Glossardefinitionen sind von den o.g. Definitionstypen dadurch zu unterscheiden, dass sie in der Regel nicht in den Textfluss integriert sind, sondern sich typographisch/strukturell vom Kontext optisch abheben. Zudem wird hier die Zuordnungsbeziehung durch einen optionalen *Separator* – ein Interpunktionszeichen – realisiert. Optional ist der *Separator* deshalb, da eine Unterscheidung zwischen Definiendum und Definiens in einigen Fällen bereits durch die typographische bzw. strukturelle Realisierung gegeben ist. In den bisherigen Tabellenformen lässt sich die Glossardefinition

Link:

Computerverwaltete Zuordnung zwischen Anker

wie folgt aufschlüsseln:

Strukturelemente:	Definiendum	Separator	Definiens		
Definition:	Link	:	Computerverwaltete	Zuordnung	zwischen Anker.
			Merkmal	Oberbegriff	Merkmal
			<i>differentiae specificae</i>	<i>genus proximum</i>	<i>differentiae specificae</i>

4. Grenzfälle: Erklärung, Beschreibung und Synonymie

Die vorangegangene Vorstellung möglicher Definitionsstrukturen sowie unsere Auswahl an Definitionstypen stellt eine praktische Hilfe für die manuelle Annotation definitorischer Textsegmente dar. Dennoch wird es immer wieder Grenzfälle geben, in denen die Entscheidung schwer fallen wird, ob das vorliegende Textsegment definitorisch ist oder eher als Erklärung, Beschreibung oder Synonym zu werten ist.

4.1 Die Erklärung

Der erste Grenzfall – die Erklärung – lässt sich aufgrund einer Verwandtschaft zur Definition auflösen. So beschreibt z.B. Ritter eine Verwendungsmöglichkeit der Erklärung wie folgt:

„Man erklärt die Bedeutung eines Zeichens, Wortes oder Satzes, die Funktion oder Arbeitsweise eines Systems.“⁶

Im Gegensatz zur definitorischen Festsetzung bzw. die Aufstellung einer Wortverwendungsnorm, zielt die Erklärung in erster Linie auf das *Verstehen* ab. Daher können erklärende Textsegmente umfangreicher als präzise Definitionen sein. Doch es ist möglich, dass durch eine umfassende Erklärung gleichsam die Definition eines Begriffs gegeben ist. Die Grenze zwischen Definition und wissenschaftlicher Erklärung löst sich daher auf und eine Abgrenzung ist nicht mehr möglich. So nennt auch Ritter als eine Bedeutung der wissenschaftlichen Erklärung

„die Explikation (s.d.), d.h. die rationale Rekonstruktion, oder die Definition (s.d.) von Begriffen oder wissenschaftlichen Ausdrücken. Hierbei handelt es sich nicht um die wissenschaftliche E. von Ereignissen oder Gesetzen, sondern um die <<E.>> der Verwendung von Ausdrücken – also um eine ganz andere, semantische Art der <<E.>>. Diese sollte deshalb mit anderen Ausdrücken benannt werden: z.B. <Explikation>, <Definition>, <Exposition>, <Erläuterung>“⁷.

Für die manuelle Annotation definitorischer Textsegmente bedeutet dies, dass ein definitorisches Textsegment durchaus einen erklärenden Charakter besitzen kann oder beide Formen in sich vereint, wie das nachfolgende Beispiel zeigt, in welchem der Begriff *Kataloge (Indizes)* im ersten Satz definiert wird:

„Kataloge (Indizes) sind Sammlungen von URLs, die von einer Redaktion zusammengestellt und gepflegt werden. Die Redaktionsmitglieder besuchen interessante Angebote im Netz und ordnen diese thematischen Kategorien zu. Weiterhin offerieren sie Möglichkeiten, eigene Angebote anzumelden und zu verschlagworten. Ein bekanntes Beispiel für einen solchen Katalog ist Yahoo.“

Die Definition lässt sich wie folgt aufschlüsseln:

Struktur elemente:	Definiendum (1)	Definiendum (2)	Definitior	Definiens	
Definition:	Kataloge	(Indizes)	sind	Sammlungen	von URLs, die von einer Redaktion zusammengestellt und gepflegt werden.
				 Oberbegriff <i>genus proximum</i>	 Merkmale <i>differentiae specificae</i>

In den folgenden drei Sätzen, die sich an diese Definition anschließen, wird die Arbeit der Redakteure erklärt. Dadurch wird dem Rezipienten eine Verstehenshilfe gegeben, die als *erklärender Kontext* bezeichnet wird.

4.2 Die Beschreibung

Um einen Gegenstand, einen Sachverhalt oder eine Idee beschreiben zu können, müssen alle wesentlichen Merkmale genannt werden, die ihn bzw. sie ausmachen.

⁶ Ritter (1971: 693).

⁷ Ritter (1971: 693).

So benötigt man z.B. für eine Wegbeschreibung präzise Richtungsangaben oder Straßennamen, für die Beschreibung einer Person oder eines Gegenstandes dagegen markante Kennzeichen. Begriffe dagegen werden eher *umschrieben* als *beschrieben*, da sie ein Abstraktum darstellen, welches mit konkreten Merkmalen möglichst genau dargestellt werden soll. Hier fehlt der Beschreibung also die Präzision der Definition.

Dennoch weist die Beschreibung eine starke Nähe zur o.g. Umfangsdefinition auf, die zwar nicht durch die Aufzählung von Eigenschaften oder Funktionen motiviert ist, jedoch durch die Aufzählung von Teilen bzw. Unterbegriffen, die das Definiendum unverwechselbar machen. Die Grenze zwischen Beschreibung und Definition wird an dieser Stelle also fließend.

Reine Begriffsumschreibungen – auch wenn sie auf den ersten Blick einen definitiven Charakter haben - bleiben jedoch unpräzise, da sie nur eine Annäherung an einen Begriff darstellen, ihn aber nicht exakt festsetzen können, wie folgendes Beispiel zeigt:

Ein Browser ist ein Programm, welches man auf vielen Rechnern standardmäßig auf der Desktopoberfläche und im Verzeichnis c:\Windows\Programme installiert findet.

In diesem Satz wird der Begriff *Browser* scheinbar definiert: Es gibt einen Oberbegriff (*Programm*) und es wird beschrieben, wo dieses Programm installiert ist. Auf den ersten Blick scheinen die beiden Installationspfade (Desktop und Windows-Programmordner) unterscheidende Merkmale zu sein, die den Oberbegriff *Programme* näher charakterisieren. Tatsächlich fehlt dieser Beschreibung jedoch die Präzision: Denn nimmt man nur die Merkmale, für die *Browser* als verkürzender Term stehen soll, führen sie nicht mehr zu diesem Begriff zurück: Denn außer Internetbrowser gibt es noch zahlreiche andere Programme, die genau diese Installationspfade aufweisen.

Zum Vergleich: Eine korrekte Definition von *Browser* ist z.B.:

Ein Browser ist ein Programm, welches dem Nutzer Orientierungs- und Navigationswerkzeuge für die interaktiven hypertextspezifischen Rezeptionsformen bereitstellt.

4.3 Die Synonymie

Synonymie liegt vor, wenn ein lexikalisches Zeichen die gleiche Bedeutung wie ein anderes lexikalisches Zeichen besitzt. *Wie* das Zeichen verwendet werden soll, wird also nicht festgesetzt. Daher ist der folgende Satz nicht als Definition in unserem Sinne zu bezeichnen, sondern gibt dem Rezipienten lediglich den Hinweis, dass *Link* synonym zu *Verweis*, *Verknüpfung* oder *Knoten* verwendet werden kann.

Link bedeutet Verweis, Verknüpfung, Knoten.

5. Vorgehen bei der manuellen Annotation

Zusammenfassend lässt sich für die manuelle Annotation von Definition folgendes Vorgehen festhalten:

1. Vorstellung entwickeln, was eine Definition ist
Der vorangestellte Text (Abschnitte 1–4) soll dazu dienen, dass Sie eine Vorstellung davon entwickeln, was eine Definition ist.
2. Problembewusstsein entwickeln
Um Definitionen erkennen zu können, muss zuerst ein Problembewusstsein dafür entwickelt werden, was eine Definition *nicht* ist. Dazu gehört, dass
 - a. keine Individuen (z.B. Eigennamen, Firmenprodukte, etc.) definiert werden dürfen, wobei beachtet werden muss, dass es jedoch Grenzfälle gibt, wie z.B. *HTML* oder *XML*, die annotiert werden sollten.
 - b. die Definitions*absicht* eines Autors nur in Ausnahmefällen explizit erkennbar ist,
 - c. Grenzfälle wie Beschreibungen, Erklärungen und Synonyme in manchen Fällen eine eindeutige Bestimmung definitorischer Textsegmente erschweren.
3. Komplexibilitätsbeschränkung einhalten
Es werden in der Regel nur definitorische Textsegmente annotiert, die incl. erklärenden Kontexts max. 4 Sätze umfassen dürfen. Doch auch hier sind Grenzfälle zu beachten: So kann die 4-Satz-Regel aufgehoben werden, wenn der Kontext nicht ausreicht, dass der Nutzer sich den Sinn des Terms erschließen kann.
4. Präzision überprüfen
Ist ein Textsegment als verdächtig identifiziert, wird es auf die vorgestellten Strukturelemente bzw. auf die Bedingungen der o.g. Definitionstypen hin überprüft. Sind alle Strukturelemente und Bedingungen erfüllt, liegt mit hoher Wahrscheinlichkeit eine Definition vor.
5. Faustregel: Austauschertest durchführen (Kriterium der kontextunabhängigen Ersetzbarkeit)
Der Austauschertest ist eine Faustregel, die in Zweifelsfällen hilfreich sein kann: Da Definitionen auch als „Verkürzungstechniken“⁸ verstanden werden können, sollten Definiendum und Definiens ohne Sinnverlust und kontextunabhängig austauschbar sein.
6. Definition annotieren
Sind Punkt 1-4 erfüllbar, kann die Definition als solche gekennzeichnet werden, wobei folgende Regeln für ausgedruckte Texte zu beachten sind:
 - **Definiendum** und **Definiens** (bzw. seine Teile) werden mit unterschiedlichen Textmarker gekennzeichnet.
 - Der **Definitor** (bzw. seine Teile) oder der **Separator** werden mit Kugelschreiber **eingekringelt**
 - **Erklärender Kontext** wird mit Kugelschreiber **unterstrichen**.
 - **Unklarheiten** werden am Rand notiert.

6. Erwähnte Literatur

Lay, Rupert (1971): Grundzüge einer komplexen Wissenschaftstheorie. 1. Band: Grundlagen und Wissenschaftslogik. Frankfurt a.M.: Josef Knecht.

Mittelstraß, Jürgen (Hrsg. in Verbindung mit Gereon Wolters) (1995): Enzyklopädie Philosophie und Wissenschaftstheorie. Band 1: A-G. Stuttgart, Weimar: Metzler.

⁸ Lay (1971: 260).

- Oberholzer, Mirjam (2002): Terminologische Definitionen: Form, Funktion, Extraktion. Universität Zürich: Lizenziatsarbeit der Philosophischen Fakultät, Institut für Computerlinguistik.
- Ritter, Joachim (Hrsg.) (1971): Historisches Wörterbuch der Philosophie. Band 1: A-C. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Speck, Josef (Hrsg. in Verbindung mit Karl Acham) (1980): Handbuch wissenschaftstheoretischer Begriffe. Band 1: A-F. Göttingen: Vandenhoeck & Ruprecht (UTB, 966).