

# Konnotative Textbedeutungen

## Zur Modellierung struktureller Aspekte der Bedeutungen von Texten

Alexander Mehler

### 1. Zusammenfassung

Gegenstand des folgenden Aufsatzes ist der *konnotative* Aspekt der Bedeutungen von *Texten*. Den Ausgangspunkt der Überlegungen zur Konnotation des Textes bildet die Auffassung, wonach Wort- und Textbedeutungskonstitution Ergebnis eines zirkulären Prozesses sind, der für die Emergenz einer Hierarchie ineinander geschachtelter Spracheinheiten verantwortlich zeichnet. Der Prozeß der Zeichenartikulation erfolgt entlang dieser Ebenen und erzeugt durch Verbindung von (konnotativer) Inhalts- und Ausdrucksseite auf Textebene das *Textzeichen*. Im Gegensatz zu einer strikten Interpretation des Fregeschen Kompositionalitätsprinzips, derzufolge die Bedeutungen sprachlicher Einheiten als fixierte, kontextfreie Größen vorauszusetzen sind, behandelt der vorliegende Ansatz bereits die lexikalische Bedeutung als Größe, die in Abhängigkeit von ihrem Kontext variieren kann. Aus semiotischer Perspektive ist es vor allem der Gestaltcharakter, welcher die konnotative Textbedeutung einer Anwendung des Frege-Prinzips entzieht. Anders ausgedrückt: Die konnotative Bedeutung eines Textes ist keineswegs in eine Struktur „atomarer“ Repräsentationen zerlegbar. Die hierarchische Organisation von Texten erweist sich insofern als komplex, als ihre Bedeutungen aus einem *zirkulären* Prozeß resultieren, der bestätigend und/oder verändernd auf die Bedeutungen der Textkonstituenten einwirkt. Diese Zirkularität bedingt, daß Texte nicht nur als Orte der Manifestation von Wortbedeutungsstrukturen anzusehen sind, sondern *zugleich* als Ausgangspunkte für die Modifikation und Emergenz solcher Strukturen dienen.

Im folgenden wird unter Rekurs auf den Kopenhagener Strukturalismus ein Modell der konnotativen Bedeutung von Texten entwickelt, das sich unter anderem an dem glossematischen Begriff der *Konstante* orientiert. Die Formalisierung

des Modells erfolgt mit Hilfe des Konzeptes der unscharfen Menge. Zu diesem Zweck werden die unscharfen Verwendungsregularitäten von Wörtern auf der Basis eines zweistufigen Verfahrens analysiert, welches die syntagmatischen und paradigmatischen Regularitäten des Wortgebrauches berücksichtigt. Die Rolle der Satzebene innerhalb des Prozesses der konnotativen Textbedeutungskonstitution wird angedeutet. Abschließend erfolgt eine Exemplifizierung des Algorithmus anhand der automatischen Analyse eines Textcorpus.

## 2. Einleitung

Innerhalb der Linguistik werden grob gesprochen zwei Bedeutungskomponenten sprachlicher Einheiten unterschieden (Halliday, 1966; s.a. Eco, 1988):

1. Der Begriff der referenziellen Bedeutung geht von der (Prä-)Existenz nicht primär sprachlicher Relationen aus, welche Sprachzeichen mit ihren außersprachlichen Referenten verbinden.
2. Die auf dem konnotativen Bedeutungsaspekt beruhende strukturelle Bedeutung umfaßt demgegenüber sämtliche sprachinternen Relationen (wie etwa Synonymie, Hyperonymie, u.a.), welche die Position eines Zeichens innerhalb des Sprachsystems bedingen.

Im folgenden wird die Hypothese vertreten, daß Texte ebenso wie andere Sprachzeichen eine *konnotative* Bedeutungskomponente aufweisen. So kann etwa die *textuelle* Konnotation eines Textes mit der *Struktur* derjenigen Texte identifiziert werden, die (vor dem Hintergrund eines bestimmten Kontextes) in einen konnotativen Zusammenhang mit dem Ausgangstext gebracht werden können. Dieser konnotative, inhaltliche Zusammenhang von Texten beruht u.a. auf der Ähnlichkeit der Bedeutungen ihrer Konstituenten. Im Zentrum der folgenden Überlegungen steht daher die Entwicklung eines Maßes, welches die Ähnlichkeit von Texten vor dem Hintergrund ihrer strukturellen Charakteristika zu bestimmen sucht. Anstatt hierbei aber die Bedeutungen von Texten und ihrer Konstituenten vorauszusetzen, wird eine empirische Analyse von Textcorpora mit dem Ziel verfolgt, die strukturelle Bedeutung der untersuchten Einheiten als das Ergebnis eines zweistufigen Prozesses zu beschreiben, der Aspekte ihrer syntagmatischen als auch paradigmatischen Abhängigkeiten in Betracht zieht. Diese Vorgehensweise steht in Zusammenhang mit Überlegungen des Harris-schen Distributionalismus, wonach gilt: „[...] difference in meaning correlates with difference of distribution.“ (Harris, 1954:156).

Die Distribution einer sprachlichen Einheit wird u.a. durch ihre Kookkurrenz mit anderen Einheiten des Sprachsystems bestimmt, und es ist, wie Halliday (1966) ausführt, gerade die Ähnlichkeit dieser Kookkurrenzen, welche die Organisation etwa lexikalischer Einheiten zu Strukturen paradigmatisch ähnlicher Ele-

mente bedingt. Gemäß Hjelmslev (1974a) und Harris (1970) bildet der Text als Ganzes den Ausgangspunkt jeder distributionellen Analyse. Hierbei ist aber die Beziehung von Text und zugrunde liegendem Sprachsystem dynamischer Natur (Halliday, 1991; Altmann, 1995): das Sprachsystem wird auf der Basis einer unzähligen Menge von Texten konstituiert und/oder modifiziert. Andererseits bildet das im Sprachsystem organisierte sprachliche Wissen den Ausgangspunkt jeder Textproduktion.<sup>1</sup> Halliday (1977) identifiziert in diesem Zusammenhang Texte als Manifestationen sprachlicher Kommunikationssituationen, die als Prozesse der Auswahl aus einem paradigmatisch organisierten System von Bedeutungspotentialen beschrieben werden können. Auf der einen Seite machen Texte Gebrauch von systematischen strukturellen Abhängigkeiten sprachlicher Einheiten, initiieren andererseits aber *infolge dieses Gebrauches* Modifikationen der Verwendungsregularitäten dieser Einheiten. Aus diesem Grunde sind Texte nicht nur als Manifestationen von Bedeutungsstrukturen zu betrachten, sondern fungieren zugleich als Ausgangspunkt für die Emergenz bzw. Modifikation solcher Strukturen.

Im Rahmen der kognitiven Linguistik wird die konnotative Bedeutung von Texten vor dem Hintergrund des Phänomens des *Textprimings* bzw. unter Rekurs auf das Verfahren der kontrollierten Assoziation untersucht. Nach Sharkey & Sharkey (1992) basiert der Effekt des Textprimings darauf, daß der Kontext einer linguistischen Einheit kontextuell unangemessene (konnotative) Bedeutungsaspekte der Einheit unterdrückt, während passende „Lesarten“ unterstützt werden. Die Kombination von Sätzen zu kohärenten Texten führt zu einer korrespondierenden Beschränkung der lexikalischen Assoziationen, welche aus der Sicht einer strukturalistischen Terminologie Bestandteil der lexikalischen Konnotation eines Textes sind. Der Umstand, daß die jeweils betrachteten Texte (mehr oder weniger) kohärent sind, ist entscheidend für die Beobachtung des Phänomens des Textprimings, da unter Zugrundelegung wahlloser Permutationen der Konstituenten eines Textes der Primingeffect nicht zu beobachten ist (Sharkey & Sharkey, 1992).

Hjelmslev (1974a) hat mit Blick auf eine allgemeine Sprachtheorie eine Systematisierung der Konnotation erarbeitet, welche die Konnotation als eine Variante der Artikulation von Ausdrucks- und Inhaltsform beschreibt, die eine Strukturierung der Ausdrucksseite des jeweiligen Zeichens in ein *System* von Zeichen vorsieht. Die Ausdrucksseite konnotativer Zeichen ist also selbst wieder als Zeichen(system) identifizierbar. Unter Verwendung einer relationalen Notation, wie sie von Barthes (1979) verwendet wird, kann dieser Umstand durch Ausdruck (1) formalisiert werden. Hierbei steht *r* für die Relation der Zeichenartikulation, die allererst aus *A* (Ausdrucksplan) und *I* (Inhaltsplan) ein intern strukturiertes Zeichen erzeugt:

---

<sup>1</sup> Während Hjelmslev die Abhängigkeit von Texten gegenüber dem System betont, wird hier das Verhältnis von System und Diskurs vom Standpunkt der Interdependenz aus thematisiert.

## (1) (A r I) r I

Die Frage erhebt sich, auf der Basis welcher Prozesse die konnotative Bedeutung von Texten emergiert. Die Antwort auf diese Frage kann andeutungsweise unter Rekurs auf das Faktum der hierarchischen Organisation sprachlicher Einheiten erfolgen: Texte integrieren Komponenten einer nächst tieferen Sprachebene, die insofern in Abhängigkeit zum Textganzen stehen, als sie nicht als isolierte Einheiten (ohne Rekurs auf das Textganze) analysierbar sind. Als Ergebnis der rekursiven Anwendung dieser innerhalb der Glossematik systematisierten Analyse­methode (Hjelmslev, 1974a) erhält man eine Stratifikation von Spracheinheiten. Die Elemente einer jeden Ebene dieser Hierarchie sind dabei auf zweierlei Art gekennzeichnet: einerseits integrieren sie eine ausgewählte Menge von Elementen der nächst tieferen Ebene, andererseits werden sie, zusammen mit anderen Elementen derselben Ebene, zu Einheiten der nächst höheren Ebene integriert. Dieser Prozeß der Integration sprachlicher Einheiten zu sogenannten *Superzeichen* heißt *Superisation* (Dörner, 1977). Ein Superzeichen resultiert in der Weise aus der Komposition von Zeichen eines untergeordneten Stratum, als das Superzeichen als ein *Ganzes* wahrnehmbar wird, das seine Teile strukturiert. In diesem Sinne stellen Texte Superzeichen dar; sie bilden geschlossene (komplexe) Einheiten (Ganzheiten), die ihrerseits in strukturellem Kontrast zu anderen (textuellen) Einheiten desselben Stratum stehen.

Einheiten derselben Ebene weisen (wechselseitige) syntagmatische Abhängigkeiten auf, die Resultat *und* Voraussetzung jeder Aktualisierung dieser *Verwendungsregularitäten* darstellen. Kookkurrenzen von Einheiten (derselben Ebene) innerhalb ähnlicher Kontexte konstituieren die paradigmatischen *Gebrauchsregularitäten* dieser Einheiten. Die Kombinationen der Zeichen zu syntagmatischen Ketten lassen, wie angedeutet, komplexere (Super-)Zeichen entstehen, die ihrerseits syntagmatische und paradigmatische Abhängigkeiten (mit zunehmend größeren Freiheitsgraden) ausbilden. Während also Syntagmatik und Paradigmatik die Beziehungen der Elemente (einer Ebene) untereinander kennzeichnen, sind es die Operationen der ausdrucksseitigen Integration und inhaltsseitigen Aggregation, welche die Beziehungen dieser Elemente zu Einheiten höherer Strata bestimmen. In diesem Sinne ist die Superisation (wie schon der Prozeß der Bedeutungskonstitution sprachlicher Einheiten) ein *zweiseitiger* Prozeß, der (via Integration) durch syntagmatische als auch (via Aggregation) durch paradigmatische Regularitäten sprachlicher Einheiten bedingt wird. Der Prozeß der Superisation zu immer komplexeren Einheiten ist im Prinzip offen.<sup>2</sup> In Abbildung 1 wird dies durch Angabe einer weiteren Sprachebene oberhalb der Textebene, deren Ele-

---

<sup>2</sup> Dies zeigt Weizsäcker (1974) allgemein für Systeme, die pragmatische Information verarbeiten.

mente hier in Anlehnung an das entsprechende informationstheoretische Konzept *Hypertexte* genannt werden, angedeutet.

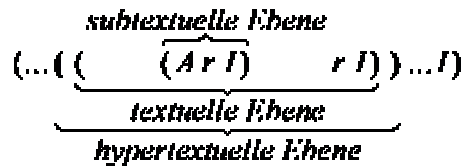


Abb. 1: Konnotation und hierarchische Organisation von Spracheinheiten

## 2.1 Zur Zirkularität von Wort- und Textbedeutungskonstitution

Die bisherigen Ausführungen scheinen zunächst die Vermutung nahezu legen, daß das Frege-Prinzip auf der Ebene der Textbedeutungskonstitution uneingeschränkt Verwendung findet. Demnach wäre die Bedeutung eines Textes als Funktion der Bedeutungen seiner Teile beschreibbar. Ein Text stellt aber alles Andere als eine einfache Komposition von Sätzen und Wörtern dar. Der Prozeß der Textbedeutungsaggregation vollzieht sich vor dem Hintergrund der Struktur des jeweils betrachteten Textes. Die Partizipation der Bedeutung eines Wortes erfolgt dabei *kontextsensitiv* bzw. *aspektabhängig* vor dem Hintergrund aller anderen Wortkonstituenten desselben Textes. Dieser Zusammenhang kann unter Verwendung des Begriffes des *Wirkungssystems* beschrieben werden, der im Rahmen einer logischen Rekonstruktion der Gestalttheorie entwickelt wurde (Rescher & Oppenheim, 1955): zwischen den Konstituenten eines Textes existiert insofern eine Dependenzrelation, als die Bedeutung jeder einzelnen Konstituente erst aus der *Perspektive* der Bedeutungen der übrigen Textkonstituenten wahrgenommen wird. Sämtliche Textkonstituenten werden stets auf der Basis ihrer perspektivischen Restriktion/Mediation durch Einheiten ihres Kontextes (und dessen Struktur) identifiziert. Diese wechselseitige kontextuelle Abhängigkeit der Einheiten eines Textes bedingt, daß der Prozeß der Textbedeutungsaggregation keineswegs dem klassischen Kompositionalitätsprinzip folgt, sondern vielmehr dem *Kontextprinzip*.<sup>3</sup> Als einfaches Beispiel für diesen Zusammenhang kann etwa die Disambiguierung polysemer Wörter in Texten dienen: das Lexem „Bank“ wird in alltagssprachlichen Diskursen innerhalb von Kontexten, die Wörter wie „Geld“, „Investition“, „Crash“, u.a. enthalten, eindeutig unter dem „monetären“ Aspekt verstanden. Daß das Wort „Bank“ polysem ist, erscheint im Rahmen letzterer

<sup>3</sup> Ein Umstand, der unter Bezug auf die referenzielle Satzbedeutung im Rahmen der Klärung der Bedeutung des Zahlbegriffes schon von Frege (1961: XXII) beschrieben wurde: “[...] nach der Bedeutung der Wörter muss im Satzzusammenhange, nicht in ihrer Vereinzelung gefragt werden [...]”. Wie Milne (1986) ausführt, hat Frege diese Interpretation in seinen späteren Werken zugunsten der Kompositionalität der Bedeutung aufgegeben.

Diskurse *nicht* als Problem, sondern erst unter metasprachlicher bzw. sprachsystematischer Perspektive.

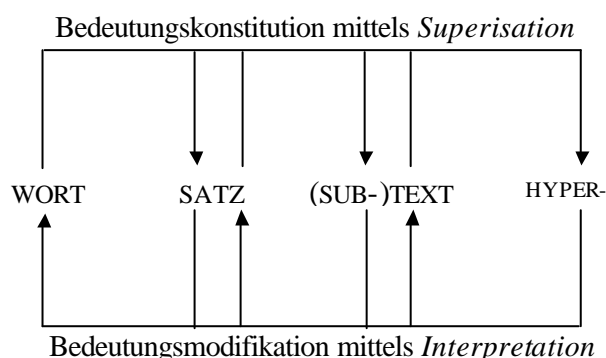


Abb. 2: Zur Superisation des Textzeichens

Auf der anderen Seite ist es aber der Text, der einerseits von strukturellen Beziehungen sprachlicher Einheiten Gebrauch macht und andererseits *infolge dieses Gebrauches* Modifikationen ihrer Verwendungsregularitäten initiiert. Der Gebrauch von Wörtern in einem Text kann eine Bedeutungsmetamorphose<sup>4</sup> dieser Wörter zur Folge haben, was aus strukturalistischer Sicht zugleich eine mittelbare Metamorphose aller anderen Elemente des (lexikalischen) Systems zur Folge hat. Insgesamt erweist sich daher der Zusammenhang von Wort- und Textbedeutungskonstitution als zirkulär. Ein Text stellt ein einzelnes Wort in den Kontext seiner übrigen Wortkonstituenten, wobei diese Art syntagmatischer Einbettung eine aspektabhängige Einschränkung der Wortbedeutung zur Folge hat. Andererseits ist es dasselbe Wort, welches an der Aggregation der Textbedeutung partizipiert. Texte sind folglich sprachliche Repräsentationen von Diskursen, innerhalb derer Wortbedeutungen emergieren und/oder verändert werden, die ihrerseits an der Textbedeutungskonstitution teilhaben.<sup>5</sup>

## 2.2 Zur konnotativen Textbedeutung

Der im folgenden entwickelte Algorithmus zur automatischen Textanalyse betrachtet lediglich Wort- und Textebene. Satz- und Absatzebene bleiben zunächst von der Untersuchung ausgespart, womit aber keineswegs die Aussage verbunden ist, beide Ebenen wären für den Prozeß der Textbedeutungsaggregation irrelevant. Vielmehr wird ihre Wirksamkeit auch für den Aspekt der konnotativen Textbedeutung als unabdingbar angesehen. Unter dieser (vereinfachenden) Per-

<sup>4</sup> Zum Phänomen der Bedeutungsmetamorphose siehe auch Thiopoulos (1992).

<sup>5</sup> Siehe Abbildung 2.

spektive können Texte auf der Basis einer Zeichenrelation beschrieben werden, welche deren Inhaltsform mit einem System lexikalischer Zeichen (Ausdrucksform) verbindet. Die Ausdrucksseite des Textzeichens wird dabei mit dem System (oder einfacher: mit der Menge) all derjenigen Lexeme identifiziert, welche innerhalb des Textes okkurrieren.<sup>6</sup> Vor dem Hintergrund ihrer Kookkurrenz innerhalb des betrachteten Textes *konnotieren* diese Lexeme *gemeinsam* eine (oder mehrere) Position(en) innerhalb des Systems von Wort- und Textbedeutungen. Die derart konnotierte, sprachsystematische Position wird im folgenden *strukturelle Textbedeutung* genannt. Es kann nun definiert werden:

**Definition 1** Die (intern strukturierte) *konnotative Bedeutung eines Textes* umfaßt die Menge aller sprachlichen Einheiten unterschiedlicher Ebenen, die aufgrund ihrer paradigmatischen Gebrauchsregularitäten (mehr oder weniger) in *Nachbarschaft* zur strukturellen Bedeutung dieses Textes stehen. *Intern strukturiert* bedeutet hierbei, daß die bedeutungskonstituierenden Sprachelemente keine ungeordnete Menge bilden, sondern selbst auf der Basis von Nachbarschaftsrelationen (partiell) geordnet sind. Es gilt dann spezieller:

- Die *lexikalische Konnotation* eines Textes enthält sämtliche Lexeme, die ähnliche paradigmatische Gebrauchsregularitäten aufweisen, wie die lexikalischen Konstituenten des Textes.
- Die *textuelle Konnotation* eines Textes umfaßt all diejenigen Texte, die mehr als andere Texte (des untersuchten Corpus) solche lexikalischen Konstituenten enthalten, welche ähnliche paradigmatische Gebrauchsregularitäten aufweisen wie die lexikalischen Konstituenten des Textes.

Letztere Definition wiederholt indirekt die weiter oben angedeutete Auffassung, wonach sich das Sprachsystem vor dem Hintergrund einer unzähligen Menge von Texten konstituiert. Jede corpusgestützte Rekonstruktion von Aspekten dieses Prozesses, die nicht schon die systematischen Beziehungen sprachlicher Elemente voraussetzt, muß diesem Umstand Rechnung tragen. Corpora bilden in diesem Sinne also hypertextuelle Kontexte, vor deren Hintergrund Aussagen über die Bedeutungen beschriebener Einheiten zu relativieren sind.

### 2.3 Implikationen für das weitere Vorgehen

Die bisherigen Überlegungen legen folgende Hypothesen nahe:

---

<sup>6</sup> Texte werden für den Zweck einer statistischen Analyse teils durch Rückgriff auf den Mengenbegriff *modelliert*, was aber nicht bedeutet, daß Texte Mengen im mathematischen Sinne darstellen. Es handelt sich vielmehr um einen Verkürzungseffekt, der sich aus der Wahl des Modelles ergibt. Ob eine solche (partielle) Verkürzung möglich ist, muß u.a. anhand der Analyseresultate bewertet werden.

- Texte haben gleichermaßen wie etwa Wörter eine strukturelle Bedeutung, die Ausgangspunkt ihrer Konnotationen ist. Im Rahmen des Hjelmslevschen Modelles gesprochen bedeutet dies, daß nicht nur die Ausdrucksseite des Textzeichens durch ein System von Zeichen strukturiert ist, sondern zugleich auch ihre Inhaltsseite.
- Die Ähnlichkeit von Texten umfaßt u.a. die Ähnlichkeit ihrer Konnotationen. Strukturelle Ähnlichkeit (relative Nachbarschaft im Bedeutungssystem) bedingt konnotative Ähnlichkeit.
- Die Ähnlichkeitsbeziehungen sprachlicher Einheiten sind dynamischer Natur; sie spiegeln die wechselnden Verhältnisse derjenigen Texte wider, die ihrer Bildung zugrunde lagen. Relationen sprachlicher Einheiten sind daher vielfach vage (Kontextvariabilität, partielle Synonymie, etc.).

Für das gesuchte Modell ergeben sich damit als Konsequenzen:

- Die Definition der Textkonnotation legt eine einheitliche Betrachtung der Bedeutungen sprachlicher Einheiten unterschiedlicher Strata nahe, wodurch eine unmittelbare Bezugnahme dieser Elemente vor dem Hintergrund des gewählten strukturalen Ähnlichkeitskriteriums ermöglicht werden soll. Aus diesem Grunde wird im folgenden ein einheitliches Repräsentationsformat für Wort- und Textbedeutungen entwickelt, das auf einer mathematischen Explikation des in obiger Definition angedeuteten Ähnlichkeitskriteriums beruht.
- Kennzeichnend für den hier verfolgten Analyseansatz ist der Umstand, daß das Wissen über die Abhängigkeitsverhältnisse der untersuchten Einheiten nicht vorausgesetzt, sondern anhand einer automatischen Textanalyse rekonstruiert werden soll. Abweichungen von dieser Prämisse werden geeignet expliziert.
- Als mittelbares Untersuchungsziel erweist sich dabei die Modellierung der *impliziten* konnotativen Ähnlichkeit von Texten: Texte sollten danach auch dann durch den Analysealgorithmus miteinander verknüpft werden, falls sie nur wenige oder keine gemeinsamen Wortkonstituenten enthalten, dafür aber überwiegend solche Lexeme umfassen, die ähnliche paradigmatische Gebrauchsregularitäten aufweisen.

### **3. Zur Modellierung der unscharfen konnotativen Textbedeutung**

Die konnotative Bedeutung eines Textes resultiert aus der Superisation des Textes auf der Basis seiner Konstituenten, wobei derselbe Mechanismus der Bedeu-



tungsaggregation auf verschiedenen sprachlichen Ebenen Anwendung findet. Der Algorithmus zur Modellierung der konnotativen Textbedeutung umfaßt in Anlehnung hieran sieben Schritte, wobei die ersten beiden Stufen zur Abbildung der Zweistufigkeit syntagmatischer und paradigmatischer Regularitäten dienen. Die Bestimmung dieser beiden Schritte als Ausgangspunkte legt (genauso wenig wie alle übrigen Stufen) die Wahl eines konkreten Maßes fest. In diesem Sinne kann der folgende Algorithmus als ein (vorläufiges) *Schema* zur Abbildung von Textkonnotationen verstanden werden.

1. Der sogenannte *Corpusraum* dient als Repräsentationsformat zur Abbildung syntagmatischer Wortverwendungsregularitäten. Er beruht auf einer Analyse der wechselseitigen syntagmatischen (kollokativen) Abhängigkeiten von Lexemen. Die Position eines Corpuspunktes innerhalb des Corpusraumes ist ein Maß für den Grad der Ähnlichkeit der syntagmatischen Regularitäten des betrachteten Lexems verglichen mit denjenigen aller anderen Lexeme des Vokabulars.
2. Syntagmatische Ähnlichkeiten (ähnliche Distributionsmuster) induzieren paradigmatische Verwandtschaften, d.h. relative Austauschbarkeit in den betrachteten Kontexten. Dementsprechend wird die *strukturelle* Bedeutung eines Lexems innerhalb des *Bedeutungsraumes* (auch *semantischer Raum* genannt) über eine Funktion abgebildet, die auf *allen* Unterschieden syntagmatischer Regularitäten des betrachteten Lexems gegenüber *allen* anderen Lexemen des Vokabulars operiert. Eine Gewichtung der Teilhabe von Lexemen an diesem Prozeß bleibt an dieser Stelle noch ausgespart. Die Achsen des Bedeutungsraumes entsprechen den betrachteten Lexemen.

Letztere zwei Stufen sind dem Riegerschen Modell entnommen (Rieger, 1989), das die Bedeutung eines *Wortes* als das Ergebnis des zweistufigen Prozesses der Bedeutungskonstitution analysiert: Paradigmatische Relationen von Lexemen werden mittels Funktionen über ihren syntagmatischen Abhängigkeiten abgeleitet. In der Modellsprache des Bedeutungsraumes gesprochen bedeutet dies, daß der Grad der Nachbarschaft zweier Punkte innerhalb des semantischen Raumes<sup>7</sup> das Ausmaß der paradigmatischen Ähnlichkeit der durch diese Punkte repräsentierten Lexeme reflektiert.

Im folgenden soll der Bedeutungsraum als einheitliches Repräsentationsformat für Wort- und Textbedeutungspunkte dienen, was eine unmittelbare Bezugnahme von Einheiten beider Ebenen (wie auch von Einheiten derselben Ebene) erlaubt:

---

<sup>7</sup> Dieser Grad wird im Riegerschen Modell mittels Euklidischer Metrik abgebildet; s.u.

3. *Analyse struktureller Rollen von Lexemen*: Die Partizipation eines Lexems am Prozeß der Bedeutungsaggregation kann in Abhängigkeit von der strukturellen Rolle (Konstante, Variable), die das Lexem im Corpus einnimmt, eingeschränkt werden. Wörter partizipieren nicht gleichberechtigt am Superisationsprozeß.
4. *Die Superisation von Wörtern zu Texten*: Auf der dritten Stufe erfolgt die *Aggregation* der strukturellen Textbedeutungen auf der Basis von Bedeutungsrepräsentationen der lexikalischen Konstituenten der betrachteten Texte. Den Output dieser Stufe bilden Bedeutungspunkte als Repräsentationen der strukturellen Textbedeutungen.
5. *Generierung der konnotativen Textbedeutung*: Durch Analyse der Umgebungen derjenigen Punkte des semantischen Raumes, die zur Repräsentation der Bedeutungen einer Menge von Texten dienen, gelangt man zu einer strukturellen Positionsbestimmung dieser Texte und damit zu einer Modellierung ihrer Konnotationen.
6. *Die Superisation von Texten zu Hypertexten*: Auf dieser Ebene werden die textuellen Konnotationen der Texte dazu verwendet, eine weitere Ebene zu generieren, die vornehmlich auf intertextuellen Abhängigkeiten beruht. Ihre Elemente werden *Hypertexte* genannt.<sup>8</sup>

Dieses allgemeine Schema kann nun durch Angabe konkreter Maße zur Implementation der Einzelschritte *instanziiert* werden. Im Anhang findet sich ein Analysebeispiel zur Exemplifizierung der Resultate der entsprechenden Instanzierungsschritte.

### 3.1 Wortbedeutungspunkte

Wie bereits angedeutet, werden die ersten beiden Analyseschritte dem Riegerischen Modell entnommen. Zur Abbildung syntagmatischer Regularitäten verwendet Rieger (1989) einen auf der Basis von Schätzwerten modifizierten Korrelationskoeffizienten. Der resultierende Koeffizient basiert auf der Überlegung, daß die Häufigkeit von Kookkurrenzen zweier Wörter als Ausgangspunkt zur Bemessung ihrer syntagmatischen Abhängigkeit verwendet werden kann:

$$(2) \quad \mathbf{a}(x_i, x_j) = \frac{\sum_{k=1}^m (h_{ik} - h_{ik}^*)(h_{jk} - h_{jk}^*)}{\sqrt{\sum_{k=1}^m (h_{ik} - h_{ik}^*)^2 \sum_{k=1}^m (h_{jk} - h_{jk}^*)^2}} \in [-1, 1];$$

wobei gilt:

---

<sup>8</sup> Mehler (1996) enthält einen Algorithmus zur automatischen Bildung von Hypertexten.

$$(3) \quad h_{ik}^* = \frac{H_i}{L} l_k \quad \text{und} \quad h_{jk}^* = \frac{H_j}{L} l_k .$$

$L$  ist die Länge des Corpus  $K = \{T_1, \dots, T_m\}$ ;  $l_k$  steht für die Länge des Textes  $T_k$ .  $h_{ik}$  repräsentiert die Häufigkeit des Lexems  $x_i \in V = \{x_1, \dots, x_N\}$ , und  $H_i$  bezeichnet die Gesamthäufigkeit des Lexems  $x_i$  im Corpus. Mit Hilfe des Korrelationskoeffizienten wird nun zunächst der Corpusraum gebildet. Der Corpuspunkt  $y_i$  eines Lexems  $x_i$  ist dabei definiert durch:

$$(4) \quad y_i = \langle \mathbf{a}(x_i, x_1), \mathbf{K}, \mathbf{a}(x_i, x_N) \rangle .$$

Als Maß für die Ähnlichkeit der syntagmatischen Regularitäten zweier Lexeme verwendet Rieger (1989) eine auf dem Corpusraum basierende Euklidische Metrik:

$$(5) \quad d(y_i, y_j) = \left( \sum_{n=1}^N (\mathbf{a}(x_i, x_n) - \mathbf{a}(x_j, x_n))^2 \right)^{\frac{1}{2}} \in [0, 2\sqrt{N}] ;$$

$y_i, y_j$  sind die Corpuspunkte der beiden Lexeme  $x_i$  und  $x_j$ . Letztere Distanzwerte werden nun dazu benutzt, den Bedeutungsraum  $S \subseteq \mathfrak{R}^N$  zu generieren. Der Bedeutungspunkt eines Lexems wird dabei (als Punkt des semantischen Raumes) über folgende Formel definiert:

$$(6) \quad z_i = \langle c_{i1}, \mathbf{K}, c_{iN} \rangle ;$$

wobei gilt:<sup>9</sup>

$$(7) \quad \forall j \in \{1, \mathbf{K}, N\}: c_{ij} = 1 - \frac{d(y_i, y_j)}{2\sqrt{N}} \in [0, 1] .$$

### 3.2 Textbedeutungspunkte

Im Anschluß an letztere beiden Analyseschritte findet die Repräsentation der Textbedeutung statt. Dabei fungieren (im Einklang mit dem Phänomen der zirkulären Wort- und Textbedeutungskonstitution) die lexikalischen Vorkommen eines Textes als Konstituenten des Prozesses der Bedeutungsaggregation. Das hat zur Folge, daß die Teilhabe eines Lexems an diesem Prozeß mehr oder weniger durch die Teilhabe der übrigen Wortkonstituenten restringiert wird. Die Bedeu-

---

<sup>9</sup> Diese Darstellung weicht insofern von der Darstellung in Rieger (1989) ab, als hier die Distanzwerte in Ähnlichkeitswerte transformiert werden.

tung eines Lexems wird auf diese Weise insofern *aspektabhängig aktualisiert*, als kotextuell passende Bedeutungsvarianten des Lexems Unterstützung finden, während unpassende Lesarten unterdrückt werden.

Die Aggregation der Bedeutung eines Textes vollzieht sich aber nicht auf die Weise, daß sämtliche seiner lexikalischen Konstituenten „gleichberechtigt“ an diesem Prozeß partizipieren. Ganz im Gegenteil hierzu können auf der Basis der Verwendungsregularitäten der Lexeme *Rollen* identifiziert werden, die den Lexemen unterschiedliche Aufgaben im Prozeß der Superisation zuweisen. Die klassische Kategorie der *Funktionswörter* etwa bildet eine Menge von Lexemen, deren Rolle darin besteht, durch ihre Okkurrenz einen *kohäsiven* Textrahmen entstehen zu lassen, innerhalb dessen andere Lexeme *kookkurrieren* können, um einen (mehr oder weniger) *kohärenten* Text zu bilden.

Anstelle einer kategorialen Aufteilung in Funktions- und Inhaltswörter, die immer schon voraussetzt, welches Wort zu welcher Klasse gehört, beansprucht die Glossematik eine strukturelle Analyse sprachlicher Einheiten, welche an ihren *Funktionen* genannten, wechselseitigen Beziehungen orientiert ist (Hjelmslev, 1974a:38). Funktionen konstituieren diejenigen Einheiten, die als ihre Argumente auftreten. Sprachliche Einheiten werden aus dieser Perspektive durch Funktionen definiert, die auf ihnen operieren, und nicht durch Mengen, denen sie angehören.<sup>10</sup> Hjelmslev (1974a) klassifiziert Funktionen auf der Basis ihrer Argumente: eine *Konstante* ist ein Argument, dessen Präsenz als notwendige Bedingung für die Präsenz desjenigen Argumentes dient, mit dem zusammen es im Verhältnis der betrachteten Funktion steht. Vor dem Hintergrund von Sprache als Diskurs meint Präsenz beispielsweise Okkurrenz innerhalb einer syntagmatischen Kette. Eine *Variable* ist ein Argument, dessen Präsenz keine notwendige Bedingung in letzterem Sinne darstellt. Weiterhin gilt: eine Interdependenz ist eine Funktion zwischen zwei Konstanten, eine Determination ist eine Funktion zwischen einer Konstanten und einer Variablen. Schließlich ist eine Konstellation eine Funktion zwischen zwei Variablen.

Das Konzept der Konstante bildet eine Art strukturalistisches Analogon zur Klasse der Funktionswörter. Die Okkurrenz einer Konstanten (Präsenz aus der Perspektive des Diskurses) stellt eine notwendige Bedingung für die Okkurrenz anderer Lexeme dar. Im vorliegenden Rahmen wird die These vertreten, daß die Frage, ob ein gegebenes Lexem als Konstante aufzufassen ist, gerade nicht binär entschieden werden kann. Es ist keineswegs so, daß die Okkurrenz eines Lexems entweder stets notwendige Voraussetzung für die Okkurrenz anderer Lexeme ist, oder ausschließlich das Umgekehrte gilt. Es kann vielmehr gezeigt werden, daß in Abhängigkeit vom Kontext ein und dasselbe Wort (etwa „Bosnien“) in einem Corpus (etwa mit Texten über den Bürgerkrieg im ehemaligen Jugoslawien) als

---

<sup>10</sup> Diese Betrachtungsweise kann unmittelbar mit Grundannahmen der Kategorientheorie in Verbindung gebracht werden. Siehe Goldblatt (1984).

Konstante operiert, in einem anderen Corpus (etwa mit Texten über Urlaubsorte von Europäern in den 70ern) aber nicht. Für eine strukturelle, nicht-kategoriale Rekonstruktion von Konstanten ist letztere, die Kontextabhängigkeit betonende Sichtweise kennzeichnend, derzufolge Konstanten gerade nicht prinzipiell als „bedeutungslose“ Einheiten mit rein syntaktischer Funktion betrachtet werden.

Wie bereits angedeutet, dienen Konstanten im wesentlichen der Herstellung einer kohäsiven Textur: sie erzeugen einen textuellen Rahmen, innerhalb dessen andere Lexeme kookkurrieren. Vom Standpunkt des Prozesses der Bedeutungssaggregation aus betrachtet kann daher ihr Beitrag in desto höherem Maße verringert werden, je deutlicher ihre Rolle als Konstante *vor dem Hintergrund einer geeigneten Corpusanalyse* Bestätigung findet. Wegen der mit dieser Proportionalitätsannahme verbundenen Graduierbarkeit empfiehlt sich die Theorie der unscharfen Mengen als Modellsprache. Zu diesem Zweck wird ein axiomatischer Ansatz verfolgt, der die Determinationsfunktion (eine *Selektion* aus der Sicht des Diskurses) als die für die Konstitution von Konstanten wesentliche Funktion in den Vordergrund rückt. In diesem Zusammenhang wird eine Funktion  $r: V \times V \rightarrow \mathfrak{R}_0^+$  eingeführt, welche die Determinationsbeziehung zweier Lexeme graduell bewertet.  $r$  unterliegt folgenden Axiomen:

**Axiom r1**  $r$  ist nicht symmetrisch. Die Möglichkeit, Konstanten überhaupt (von anderen Lexemen) unterscheiden zu können, setzt Asymmetrie voraus.

**Axiom r2**  $r(x_i, x_j) = \min_{x_l, x_k \in V} r(x_l, x_k) = 0$  iff  $\neg \exists T_n \in K: h_{in}, h_{jn} > 0$ . Falls  $x_i$  und  $x_j$  niemals kookkurrieren, dann ist die Okkurrenz von  $x_i$  in keinem Sinne notwendige Bedingung für die Okkurrenz von  $x_j$ .

**Axiom r3**  $\forall x_i \in V: r(x_i, x_i) = \max_{x_j, x_k \in V} r(x_j, x_k) = 1$ . (Existenz eines maximalen Wertes für  $r$ ).

Eine potentielle Klasse von Funktionen, die letztere Axiome erfüllen und zugleich die *Rekurrenz* von Lexemen (allerdings nicht ihre Abfolgeordnung) innerhalb der Texte berücksichtigen, ist mit Ausdruck (8) gegeben:

$$(8) \quad r(x_i, x_j) = \frac{\sum_{t=1}^{|K|} \min(h_{it}, h_{jt})^p}{\sum_{t=1}^{|K|} h_{it}^p} \in [0, 1].$$

Funktionen dieser Klasse bewerten den Grad, mit dem Okkurrenzen des Typs  $x_j$ , die selektierende Einheit, Okkurrenzen vom Typ  $x_i$ , die *selektierte* Einheit, voraussetzen. Instanzen der Klasse (8) erreichen ihr Maximum, falls die selektierte Einheit (zumindest) in allen Texten okkurriert, in denen die selektierende Einheit okkurriert, und zwar mindestens mit der Häufigkeit der selektierenden Einheit. Erreicht die jeweils betrachtete Instanz dieser Funktionsklasse auch im

Falle umgekehrter Ordnung (falls also  $x_i$  als selektierendes und  $x_j$  als selektiertes Lexem betrachtet wird) ihr Maximum, dann liegt (vom Standpunkt dieser Funktion aus betrachtet) der Fall einer idealen Interdependenz beider Lexeme vor. Ein Lexem wird aber desto eher als Konstante zu bezeichnen sein, je mehr die Bewertung dieser inversen Ordnung vom Maximum der Funktion abweicht. Diese Überlegungen erlauben folgende Charakterisierung der *unscharfen* Mengen der Konstanten: der Zugehörigkeitswert eines Lexems  $x \in V$  zur Menge der Konstanten ist desto höher, je mehr Lexeme (des betrachteten Vokabulars  $V$ ) die Rekurrenz dieses Lexems voraussetzen, damit diese selbst innerhalb der Texte (des betrachteten Corpus) okkurrieren können. Sei etwa die Funktion

$$(9) \quad r(x_i, x_j) = \frac{\sum_{t=1}^{|K|} \min(h_{it}, h_{jt})^2}{\sum_{t=1}^{|K|} h_{jt}^2}$$

gegeben. Sei ferner  $C$  die unscharfe Menge der Konstanten. Dann kann die Zugehörigkeitsfunktion dieser Menge definiert werden als:

$$(10) \quad m_c(x_i) = \frac{\sum_{x_j \in V} r(x_i, x_j)}{\max_{x_k \in V} \left\{ \sum_{x_j \in V} r(x_k, x_j) \right\}} \in [0, 1].$$

Offensichtlich werden durch Ausdruck (10) solche Lexeme, die durch eine Mehrzahl von Wörtern selektiert werden (deren Okkurrenz also durch diese Mehrzahl von Wörtern vorausgesetzt wird) auf hohe Zugehörigkeitswerte abgebildet.

Bezogen auf den Prozeß der Bedeutungsaggregation kann weiterhin die Gruppe der *Solitärwörter* genannten Lexeme identifiziert werden. Hierbei handelt es sich um solche Lexeme, die dazu tendieren, in nur geringer Zahl in nur wenigen Texten zu okkurrieren (starke Konzentration bei gleichzeitig niedriger Frequenz). Da die *Verwendungsregularitäten* solcher Wörter wegen der geringen Zahl ihrer Okkurrenzen nicht ausreichend Bestätigung finden, fehlt die Bezugsbasis zur Bestimmung ihrer *Gebrauchsregularitäten* und damit für die Ableitung ihrer strukturellen Bedeutung. Aus diesem Grunde können Solitärwörter nicht hinreichend von den übrigen Wörtern des Vokabulars separiert werden. Der glossematische Begriff der Substitution (Hjelmslev, 1974b) kann zur Beschreibung dieses Sachverhaltes herangezogen werden. Zwei Spracheinheiten heißen zunächst (mehr oder weniger<sup>11</sup>) substituierbar, falls ihre (wechselseitige) Substitution (beinahe) nicht zu einer Modifikation des Inhaltsplanes derjenigen Einheit führt, innerhalb derer die Substitution vorgenommen wird. Solitärwörter sind mi-

---

<sup>11</sup> Hjelmslev behandelt nicht den graduellen Charakter der Substitution.

nimal substituierbare Einheiten, da sie kaum funktionale Beziehungen aufweisen. Dieser Umstand erlaubt es, ihre Teilhabe am Prozeß der Bedeutungsaggregation einzuschränken.

Die bisherigen Überlegungen können folgendermaßen zusammengefaßt werden: der Betrag struktureller Information von Konstanten fällt gering aus, da ihre Verwendungsregularitäten auf der Basis einer Vielzahl höchst heterogener Texte bestätigt werden. Im Gegensatz hierzu sind die Regularitäten von Solitärwörtern hochgradig *ungewiß*, da ihre (Ko-)Okkurrenzen kaum Bestätigung finden. Im folgenden wird ein Modell zur Repräsentation der unscharfen konnotativen Bedeutung von Texten vorgestellt, welches neben dem Prinzip zirkulärer Bedeutungskonstitution zugleich die Rolle der Konstanten reflektiert.

In der Literatur lassen sich mehrere Klassen von Operationen für die Kombination unscharfer Mengen finden (Klir & Folger, 1988; Klir & Yuan, 1995). Aus Sicht der Modellierung von Bedeutungen mittels Vektoren (als Repräsentationen für Wortbedeutungen<sup>12</sup>) erscheinen Mittelwertoperationen als besonders geeignet, da sie wegen ihrer Interaktivität<sup>13</sup> das Phänomen der perspektivischen Aktualisierung der Bedeutungen sprachlicher Einheiten abzubilden erlauben. Mittelwertoperationen für unscharfe Mengen bilden eine spezielle Klasse von Aggregationsoperationen  $h$ , für die gilt:

$$(11) \quad h: [0, 1]^n \rightarrow [0, 1].$$

Angewandt auf eine Familie unscharfer Mengen  $A_1, \dots, A_n$  über  $X$  resultiert  $h$  in einer unscharfen Menge  $A$ , deren Zugehörigkeitsfunktion  $m_A$  wie folgt definiert ist:

$$(12) \quad \forall x \in X: m_A(x) = h(m_{A_1}(x), m_{A_2}(x), \dots, m_{A_n}(x)).$$

Als Aggregationsfunktion erfüllt  $h$  zumindest die Axiome h1 und h2 (Klir & Folger, 1988):

**Axiom h1**  $h(0, \dots, 0) = 0$ ,  $h(1, \dots, 1) = 1$  (Grenzfälle).

**Axiom h2** Für jedes Paar von Familien  $(a_i | i \in \mathbf{N}_n)$  und  $(b_i | i \in \mathbf{N}_n)$ , mit  $a_i, b_i \in [0, 1]$ , gilt: falls für alle  $i \in \mathbf{N}_n$   $a_i \geq b_i$ , dann ist  $h(a_i | i \in \mathbf{N}_n) \geq h(b_i | i \in \mathbf{N}_n)$ .  $h$  ist eine monotone nicht-fallende Funktion in all ihren Argumenten.

---

<sup>12</sup> Siehe die Ausführungen zum Corpus- und Bedeutungsraum.

<sup>13</sup> Operationen auf unscharfen Mengen heißen interaktiv, falls ihr Operationsergebnis nicht (zwingend) mit einem der Argumentwerte identisch ist, letztere sich also (möglicherweise) gegenseitig beeinflussen. Siehe auch Kruse, Gebhardt & Klawonn (1995).

Mittelwertfunktionen erhält man aus Aggregationsfunktionen durch Einführung eines weiteren Axioms:

**Axiom h3** Für jede Familie  $(a_i \in [0, 1] | i \in \mathbf{N}_n)$  von Zugehörigkeitswerten gilt:

$$\min(a_1, \dots, a_n) \leq h(a_1, \dots, a_n) \leq \max(a_1, \dots, a_n).$$

Eine Klasse von Funktionen, die allesamt diese drei Axiome erfüllen, ist mit Ausdruck (13) gegeben:

$$(13) \quad h(a_1, \dots, a_n) = \left( \frac{\sum_{i=1}^n a_i^a}{n} \right)^{1/a}.$$

Da, wie oben ausgeführt, Lexeme nicht gleichberechtigt an der Bedeutungsaggregation teilhaben und also ihr Beitrag unterschiedlich zu bewerten ist, sind weiterhin (verallgemeinerte) gewichtete Mittelwertoperationen von besonderem Interesse. Solcherlei Funktionen sind definiert durch:

$$(14) \quad h(a_1, \dots, a_n; w_1, \dots, w_n) = \left( \sum_{i=1}^n w_i a_i^a \right)^{1/a},$$

wobei zusätzlich gilt:

$$(15) \quad \sum_{i=1}^n w_i = 1.$$

Die Gewichte  $w_i$  bilden die relative Bedeutsamkeit der aggregierten unscharfen Mengen ab.

Sei nun ein beliebiger Text des Corpus gegeben. Die in diesem Text okkurrierenden Lexeme bilden auf der Basis ihrer Bedeutungspunkte die Konstituenten der Textbedeutungsaggregation. Das Phänomen der aspektabhängigen Aktualisierung der Bedeutung eines Wortes wird dabei über eine geeignete Restriktion seines Bedeutungspunktes modelliert. Dies gelingt durch gleichzeitige Berücksichtigung der Bedeutungspunkte aller übrigen lexikalischen Konstituenten des Textes. Zu diesem Zweck wird ein gewichteter Mittelwert für *Bedeutungspunkte* nach Ausdruck (16) berechnet, der eine Art von „parallel con-straint satisfaction“<sup>14</sup> realisiert:

---

<sup>14</sup> Zu diesem Begriff und seiner gestalttheoretischen Analogie siehe Vanman et al. (1995).



$$(16) \quad z_{T_k} = \sum_{x_i \in L(T_k)} w_{ik} z_i,$$

wobei gilt:

$$(17) \quad w_{ik} = w_{ik}(h_{ik}, \kappa_{ik}) = \frac{h_{ik} \kappa_{ik}}{\sum_{x_j \in L(T_k)} h_{jk} \kappa_{jk}}$$

sowie

$$(18) \quad \sum_{x_i \in L(T_k)} w_{ik} = 1.$$

$h_{ik}$  ist die Frequenz von Lexem  $x_i$  in Text  $T_k$ .  $z_i \in S$  ist der Bedeutungspunkt von  $x_i$ .  $L(T_k)$  ist die Menge der Lexeme aus Text  $T_k$ .

Die Anwendung von Formel (16) generiert einen weiteren Punkt des Bedeutungsraumes, der *Textbedeutungspunkt* genannt wird; es handelt sich um ein Modell der strukturellen Bedeutung eines Textes. Eine beliebige Koordinate dieses Punktes kann als Maß dafür betrachtet werden, inwieweit die Gebrauchsregularitäten desjenigen Lexems, das durch die betrachtete Koordinate repräsentiert wird, durch Wortvorkommen innerhalb dieses Textes widergespiegelt werden. Je höher dieser Koordinatenwert ist, desto höher ist das Ausmaß, in welchem die Gebrauchsregularitäten der lexikalischen Konstituenten des Textes denjenigen des betrachteten Wortes entsprechen (und umgekehrt).

$\kappa_{ik}$  ist ein Bias, der die Rolle von Wörtern im Sinne ihrer Zugehörigkeit zur unscharfen Menge der Konstanten reflektiert. Je deutlicher ein Lexem innerhalb des gewählten Corpus als Konstante fungiert, desto kleiner ist der korrespondierende Wert von  $\kappa_{ik}$  und desto geringer fällt der Beitrag dieses Lexems bezogen auf den Wert von  $z_{T_k}$  aus.  $\kappa_{ik}$  dient zur Bewertung der (in diesem Sinne eingeschränkten) Bedeutsamkeit eines einzelnen Wortes für die Aggregation der Textbedeutung. Unter Rückgriff auf die unscharfe Menge der Konstanten kann  $\kappa_{ik}$  für das Lexem  $x_i$  wie folgt berechnet werden:

$$(19) \quad \kappa_{ik} = \frac{m_{e_c}(x_i)}{e_c}.$$

$e_c$  ist das Equilibrium der Komplementoperation für unscharfe Mengen.  $e_c$  entspricht 0,5, falls  $m_{A^c}(x) = 1 - m_A$  die Zugehörigkeitsfunktion für das Komplement  $A^c$  von  $A$  ist.

Die bisherigen Ausführungen erlauben eine gegenüber der vektoriellen Repräsentation äquivalente Darstellung der strukturellen Bedeutungen sprachlicher

Einheiten mittels unscharfer Mengen. Zu diesem Zweck wird die Zugehörigkeitsfunktion  $\mathbf{m}_{M(x_i)}$  der unscharfen Menge  $M(x_i)$  definiert durch:

$$(20) \quad \mathbf{m}_{M(x_i)}(x_k) = 1 - \frac{d(y_i, y_k)}{2\sqrt{N}} \in [0, 1].$$

Diese Formel repräsentiert die unscharfe strukturelle Bedeutung eines Lexems  $x_i$ . Je größer der Zugehörigkeitswert eines Lexems  $x_k$  ausfällt, desto geringer ist die Distanz der den Lexemen  $x_i$  und  $x_k$  zugeordneten Corpuspunkte und desto mehr ähneln sich schließlich ihre Gebrauchsregularitäten. Auf der anderen Seite induziert ein Zugehörigkeitswert nahe 0 extrem differierende Verwendungsregularitäten von  $x_i$  und  $x_k$ . Um nun auch Textbedeutungspunkte geeignet als unscharfe Mengen über dem Vokabular darstellen zu können, bedarf man eines Analogons für die Mittelwertbildung für Vektoren. Zu diesem Zweck wird das interaktive Produkt  $\otimes$  für eine Familie  $\mathfrak{S} = \{A_i\}$  unscharfer Mengen mit folgender Zugehörigkeitsfunktion eingeführt:

$$(21) \quad \mathbf{m}_{\otimes_{i \in \mathfrak{S}} A_i} = \sum_{i=1}^{\text{Card}(\mathfrak{S})} w_i \mathbf{m}_{A_i}(x).$$

Nun ist es möglich, die unscharfe strukturelle Bedeutung eines Textes als das interaktive Produkt all derjenigen unscharfen Mengen zu definieren, welche die Bedeutungen der lexikalischen Konstituenten des Textes repräsentieren:

$$(22) \quad M(T) = M(x_1) \otimes M(x_2) \otimes \dots \otimes M(x_n) = \otimes_{x_i \in L(T)} M(x_i).$$

Diese Vorgehensweise macht es prinzipiell möglich, die Bedeutung eines Textes in Übereinstimmung mit dem Phänomen struktureller Vagheit als unscharfe Menge über dem Vokabular zu repräsentieren: Ähnlichkeitsrelationen zwischen linguistischen Einheiten werden damit über Zugehörigkeitsfunktionen abgebildet, welche die Zugehörigkeit eines Elementes zu einer Menge mit beliebiger Resolution darstellen können.

### 3.3 Zur Textkonnotation

Textbedeutungspunkte werden wegen der Wahl eines einheitlichen Repräsentationsformats für Wort- und Textbedeutungen auf denselben topologischen Raum abgebildet wie Wortbedeutungspunkte. Dies ermöglicht u.a. die Analyse

von Nachbarschaftsbeziehungen sprachlicher Einheiten unterschiedlicher Strata. In diesem Sinne enthält etwa die lexikalische Umgebung eines Textes  $T$ , seine *lexikalische Konnotation*, all diejenigen Wörter, für deren Bedeutungspunkte folgende Beziehung gilt:

$$(23) \quad U_z(r) = \{z_i \mid \exists x_i \in E: z_i = \mathbf{p}(x_i) \wedge \mathbf{z}(z, z_i) \leq r\}.$$

$\mathbf{z}$  ist eine auf dem Bedeutungsraum definierte Euklidische Metrik; ferner ist  $z$  der Bedeutungspunkt von Text  $T$ .  $E$  ist die Menge  $V \cup K$  der Wörter und Texte. Schließlich ist  $\mathbf{p}: E \rightarrow S \subseteq \mathfrak{R}^N$  eine Funktion, die jeder (untersuchten) sprachlichen Einheit  $x$  einen Punkt des semantischen Raumes als Repräsentation seiner Bedeutung zuordnet. Durch Einschränkung der Menge  $U_z(r)$  auf die Menge der Lexeme ( $U_z(r) - K$ ) bzw. auf das Corpus ( $U_z(r) - V$ ) erhält man die lexikalische bzw. textuelle Konnotation des betrachteten Textes. Die Elemente der Menge  $U_z(r)$  bilden selbst wieder eine Struktur. So kann beispielsweise eine über der Menge  $U_z(r)$  definierte Ordnungsrelation  $\pi \subseteq U_z(r) \times U_z(r)$  bestimmt werden, für die gilt:  $x_i \pi x_j$  gdw.  $\mathbf{z}(z, \mathbf{p}(x_i)) \leq \mathbf{z}(z, \mathbf{p}(x_j))$ .  $\pi$  ordnet (von links nach rechts) sprachliche Einheiten mit zunehmender Distanz zum Bedeutungspunkt des betrachteten Textes. Aus der Gültigkeit der Ordnungsbeziehung  $x_i \pi x_j$  ( $x_i, x_j \in V$ ) kann geschlossen werden, daß die Verwendung des Lexems  $x_i$  im betrachteten Corpus für die in Text  $T$  zum Ausdruck kommenden Verwendungsregularitäten *typischer* ist als die Verwendung des Lexems  $x_j$ . Die Menge  $U_z(r) - K$  läßt sich insofern zur Bestimmung der lexikalischen Konnotation eines Textes heranziehen, als sie nur solche Wörter umfaßt, die hinsichtlich ihrer Verwendungsregularitäten eine größere als durch den Schwellwert  $r$  ausgedrückte minimale Ähnlichkeit bezogen auf Text  $T$  aufweisen. Es handelt sich dabei um solche Wörter, deren Positionen im Bedeutungsraum in hohem Maße durch Okkurrenzen von Wörtern im jeweils betrachteten Text *konnotiert* werden.

Die Zugehörigkeit zur Menge  $L_n(T_k)$  der lexikalischen Konstituenten des Textes  $T$  stellt weder eine notwendige noch hinreichende Bedingung für die Zugehörigkeit zur Menge  $U_z(r)$  dar. Die Generierung des Bedeutungspunktes eines Textes ergibt sich gemäß Ausdruck (16) aus dem gewichteten Mittel derjenigen Bedeutungspunkte, die in  $T$  vorkommende Lexeme repräsentieren. Sei etwa  $z_i = \mathbf{p}(x_i)$  der Bedeutungspunkt eines Lexems  $x_i \in L(T)$ , dann gilt: der Bedeutungspunkt von  $T$  liegt um so weiter entfernt von  $z_i$ , je größer die Distanzen  $\mathbf{z}(z_i, z_j)$  (bezogen auf die Bedeutungspunkte  $z_j$ ) der Lexeme  $x_j \in L(T)$  ausfallen und zugleich je bedeutsamer diese Lexeme (im Sinne des Bias  $\mathbf{k}$ ) sind. Das Vorkommen in einem Text stellt folglich keinerlei Gewähr dafür dar, in der Menge  $U_z(r)$  enthalten zu sein. Die Menge  $U_z(r)$  umfaßt vielmehr (neben Texten) solche Wörter, die unabhängig davon, ob sie im betrachteten Text vorkommen oder nicht, insofern in einen *konnotativen Zusammenhang* mit Text  $T$  zu bringen

sind, als sie (im Sinne des Schwellwertes  $r$ ) bezogen auf die Mehrzahl der Konstituenten von  $T$  ähnliche strukturelle Eigenschaften aufweisen.

### 3.4 Zur Inkorporation der Satzebene

Der bisherige Analyseansatz kann insofern andeutungsweise erweitert werden, als neben Wort- und Textebene gleichermaßen die Ebene von Sätzen und Absätzen inkorporiert wird. Die Einführung weiterer Aggregationsebenen hat zunächst zur Folge, daß die unmittelbaren Textbedeutungskonstituenten (erwartungsgemäß) nicht mehr auf der Ebene von Wörtern angesiedelt sind. Die Aktualisierung der Bedeutung eines einzelnen Wortes wird dadurch gleichfalls nicht mehr unmittelbar durch die übrigen lexikalischen Konstituenten eines Textes restringiert, sondern durch dessen Satz- und Absatz-basierte *Textur*. Unter dieser Perspektive haben dann nur noch diejenigen Wörter unmittelbar restringierende Wirkung auf die aspektabhängige Aktualisierung der Bedeutung eines Wortes, die in demselben Satz vorkommen, wie dieses selbst. Alle übrigen Wörter entfalten ihre aspektstiftende Wirkung mittelbar über den *Vor-* bzw. *Nachtext* desjenigen Satzes, in dem das betrachtete Wort okkurriert.

Der einfachste Ansatz zur Berücksichtigung weiterer Aggregationsebenen besteht darin, ohne Rücksichtnahme auf die Ordnung der Abfolge der Textkonstituenten, denselben Analysealgorithmus auf verschiedenen Ebenen anzuwenden. Man erhält dann Bedeutungspunkte für Wörter, Sätze, Absätze und Texte, die allesamt innerhalb des semantischen Raumes lokalisiert sind. Problematisch ist dieses Verfahren u.a bezogen auf die Absatzebene, da letzteres Verfahren unterstellt, daß Absätze konnotativ homogene Gebilde ohne Satz-übergreifende Struktur sind. Ein alternativer Ansatz besteht folglich darin, die Absatzebene nicht schon vorauszusetzen, sondern selbst als zu rekonstruierende Größe zu betrachten.

## 4. Ein Analysebeispiel

Im folgenden wird ein Analysebeispiel vorgestellt, welches anhand zweier inhaltlich divergierender Texte Ergebnisse des Algorithmus zur automatischen Textanalyse aufzeigt. Beide Texte sind einem Corpus von 502 Artikeln der Süddeutschen Zeitung SZ vom März/April 1996 entnommen. Laut Angaben der SZ entstammen diese Artikel 16 verschiedenen Ressorts.

Tabelle 1  
Die Situation des Analysebeispiels

<i>Variable</i>	<i>Wert</i>
Anzahl der Texte	502

Anzahl aller Wortformen (Token) und Satzzeichen	378.185
Anzahl aller Wortformen	310.771
Anzahl der Lexeme (Types) im Corpus	36.516
Anzahl untersuchter Wortformen	75.488
Anzahl untersuchter Lexeme	1.998
durchschnittliche Textlänge im restringierten Corpus	150
durchschnittliche Worthäufigkeit im restringierten Corpus	38
Umfang der Stopwortliste	2.200

Zum Versuchsablauf: in einem ersten Schritt wurden alle Wortformen aller Texte automatisch auf die ihnen zugeordneten Lexeme reduziert. Nach diesem Schritt wurden 310.771 Wortformen identifiziert, die 36.516 verschiedenen Lexemen zugeordnet werden konnten. Diese Menge wurde im nächsten Schritt durch Bezug auf die absolute Häufigkeit der Lexeme (Nomina mit mindestens 10 Okkurrenzen wurden ebenso berücksichtigt wie Instanzen anderer Wortarten mit mindestens 14 Okkurrenzen) und eine Stopwortliste (Funktionswörter, Monatsnamen) weiter eingeschränkt, so daß schließlich 1.998 Lexeme für die weitere Analyse übrig blieben.

Tabelle 2  
Auszug aus der Stopwortliste

...	heutig	heutigentags
heutzutage	hier	hierüber
hieran	hierauf	...

In einem zweiten Schritt wurden der Corpus- und Bedeutungsraum basierend auf der Menge der 1.998 Lexeme (die Menge  $V$ ) und 502 Texte (das Corpus  $K$ ) gemäß Formel (16) berechnet. In diesem Zusammenhang wurde die unscharfe Menge der Konstanten über dem Vokabular  $V$  auf Grundlage von Ausdruck (10) bestimmt. Tabelle 3 enthält einige Beispiele von Wörtern mit Angabe des Zugehörigkeitswertes bezogen auf die Menge der Konstanten.

Tabelle 3  
Ein Auszug aus der unscharfen Menge der Konstanten  
in Tabellenrepräsentation.

<i>Rang</i>	<i>Lexem</i>	$m_c$
1.	werden	1,0

2.	haben	1,0
3.	können	1,0
...	...	...
64.	kosten	0,745
65.	leben	0,742
66.	unternehmen	0,740
...	...	...

Im folgenden werden zwei Textbeispiele zusammen mit ihren lexikalischen und textuellen Konnotationen betrachtet. Der erste Text hat eine Länge von 546, der zweite Text eine Länge von 242 Wortformen:

Tabelle 4  
Das erste Textbeispiel aus der SZ vom 01.04.1996.

„In der nach Unabhängigkeit strebenden Kaukasusrepublik...  
Jelzin befiehlt Einstellung der Kämpfe in Tschetschenien / Verhandlungen mit Rebellenführer Dudajew angekündigt  
Friedensplan des russischen Präsidenten sieht Abzug der Armee ‘aus friedlichen Regionen’ und Parlamentswahlen vor. Auf Terror-Akte soll weiter militärisch geantwortet werden / Oberbefehlshaber: Sofortiges Ende des Einsatzes ist unmöglich  
tu Moskau (Eigener Bericht) - Der russische Präsident Boris Jelzin hat am Sonntag den russischen Militärverbänden in der Kaukasusrepublik Tschetschenien die Einstellung aller Kampfhandlungen am 1. April um 0 Uhr befohlen. Doch werde auf ‘Terror-Akte’ der Tschetschenen weiter militärisch geantwortet. Im Fernsehen verkündete er einen Friedensplan, der einen Abzug der russischen Armee aus den ‘friedlichen Regionen’, Parlamentswahlen in Tschetschenien sowie die Wiederaufnahme von Verhandlungen mit den tschetschenischen Rebellen vorsehe.  
...“

Tabelle 5  
Das zweite Textbeispiel aus der SZ vom 01.04.1996

„Heereman trumpft auf - zu Recht

Nirgendwo sonst in der EU herrscht solch große Angst vor BSE wie in der Bundesrepublik. In keinem Land hat der Rindfleischverzehr seit den Schreckensmeldungen aus Großbritannien so stark nachgelassen. Bauernverbandspräsident Constantin Freiherr von Heereman verlangt von Brüssel Entschädigung für seine Rindermäster, und er hat doppelten Grund dazu. Zum ersten ist evident, daß die Solidarität, die die 15 Staats- und Regierungschefs der Gemeinschaft just in Turin erst den Briten bekundet haben, gegenüber allen geschädigten Landwirten in der Union geübt werden muß. Die Regeln der EU sind eindeutig. Der Hilfe für die anderen gebührt sogar Vorrang, denn sie sind reine Opfer, hingegen in Großbritannien sich Opfer und Täter mischen.“

Die Menge der 50 Lexeme, welche mittels ihrer Bedeutungspunkte unmittelbare Nachbarn der den Texten zugeordneten Bedeutungspunkte sind, werden in den Tabellen 6 und 7 wiedergegeben. Diese Listen, die von oben nach unten und von links nach rechts zu lesen sind (entsprechend zunehmender Entfernung zum Textbedeutungspunkt), enthalten solche Lexeme, die durch den Text bzw. seine Konstituenten konnotiert werden.

Tabelle 6  
Lexikalische Konnotationen des ersten Textbeispiels

<i>Kreml</i>	<i>Grosny</i>	<i>Armee</i>	<i>IKRK</i>	<i>Tutsi</i>
<i>Posten</i>	<i>Russische</i>	<i>Kampf</i>	<i>Internationale</i>	<i>brechen</i>
<i>Frieden</i>	<i>Süd</i>	<i>Rußland</i>	<i>Westen</i>	<i>Gerichtshof</i>
<i>Nationalist</i>	<i>Sowjetunion</i>	<i>UNO</i>	<i>Präsident</i>	<i>Panmunjom</i>
<i>Kommunist</i>	<i>Waffe</i>	<i>Flüchtling</i>	<i>Hundert</i>	<i>Ruanda</i>
<i>Abzug</i>	<i>sowjetisch</i>	<i>Widerstand</i>	<i>Verbrechen</i>	<i>verpflichten</i>
<i>Abkommen</i>	<i>angreifen</i>	<i>wollen</i>	<i>Regierung</i>	<i>töten</i>
<i>Republik</i>	<i>Kämpfer</i>	<i>Feld</i>	<i>Woche</i>	<i>Verhandlung</i>
<i>verkünden</i>	<i>Russe</i>	<i>aufbauen</i>	<i>Slowenien</i>	<i>versprechen</i>
<i>militärisch</i>	<i>Truppe</i>	<i>Wahl</i>	<i>Land</i>	<i>Nordkorea</i>

Tabelle 7  
Lexikalische Konnotationen des zweiten Textbeispiels.

<i>Entschädigung</i>	<i>Landwirtschaftsminister</i>	<i>Stufe</i>	<i>Kraft</i>	<i>Stabilität</i>
----------------------	--------------------------------	--------------	--------------	-------------------

<i>verbrennen</i>	<i>Creutzfeldt</i>	<i>katholisch</i>	<i>Forderung</i>	<i>Gefängnis</i>
<i>Agrarminister</i>	<i>Verhältnis</i>	<i>verpflichten</i>	<i>Erreger</i>	<i>Einnahme</i>
<i>Maßnahme</i>	<i>Rinderseuche</i>	<i>Italien</i>	<i>Währungsunion</i>	<i>Umwelt</i>
<i>Rinderwahnsinn</i>	<i>staatlich</i>	<i>Euro</i>	<i>streichen</i>	<i>Zentrale</i>
<i>Minister</i>	<i>Bedingung</i>	<i>Datum</i>	<i>Bundesrepublik</i>	<i>Zusammenarbeit</i>
<i>Gemeinschaft</i>	<i>Krise</i>	<i>Regelung</i>	<i>Angabe</i>	<i>Istanbul</i>
<i>Fleisch</i>	<i>Maastricht</i>	<i>gemeinsam</i>	<i>töten</i>	<i>bezahlen</i>
<i>Konferenz</i>	<i>Verbraucher</i>	<i>durchsetzen</i>	<i>Belgrad</i>	<i>Stiftung</i>
<i>Luxemburg</i>	<i>festlegen</i>	<i>deutsch</i>	<i>Europa</i>	<i>übertragen</i>

Wörter der Tabellen 6 und 7 in kursiver Schrift okkurieren *nicht* in den entsprechenden Textbeispielen, sind aber dennoch Bestandteil der entsprechenden Textkonnotationen, da sie verglichen mit den lexikalischen Konstituenten der Texte ähnliche paradigmatische Gebrauchsregularitäten aufweisen. Keines der Wörter „Slowenien“, „Tutsi“, „Gerichtshof“, „Panmunjom“, „Ruanda“ okkuriert innerhalb des Textbeispielen zum Bürgerkrieg in Tschetschenien, dennoch kann jedes dieser Lexeme vor dem Hintergrund des Topos „Bürgerkrieg“ in einen konnotativen, inhaltlichen Zusammenhang mit diesem Text gebracht werden. Alle diese Wörter weisen Gebrauchsregularitäten auf, welche mehr oder weniger typisch für Wörter im Zusammenhang der Schilderung von Bürgerkriegsereignissen sind. Daher stehen diese Wörter in Zusammenhang mit Lexemen wie „Armee“, „Kampf“, „Verbrechen“, die allesamt in erstem Textbeispiel vorkommen. Der vorgestellte Textanalysealgorithmus ist somit durchaus in der Lage, implizite konnotative Beziehungen sprachlicher Einheiten abzubilden, die nicht direkt durch die Okkurrenz im Text vermittelt werden. Darüber hinaus fällt auf, daß die lexikalischen Konnotationen beider Textbeispiele kaum Wörter gemeinsam haben (lediglich „verpflichten“ und „töten“). Die inhaltliche Differenz beider Artikel findet auf der Ebene der rekonstruierten lexikalischen Konnotationen erwartungsgemäß eine Entsprechung. Die im folgenden aufgeführten textuellen Konnotationen (Tabelle 8 und 9) bestätigen dieses Ergebnis. Es handelt sich dabei jeweils (bezogen auf Distanzen der entsprechenden Bedeutungspunkte im semantischen Raum) um die 10 nächst benachbarten Texte der entsprechenden Textbeispiele:

Tabelle 8

Die textuelle Konnotation des ersten Textbeispielen.

1. Jelzins Plan läßt Raum für Willkür und Waffenruhe [Über den Bürgerkrieg in Tschetschenien]; 01.04.1996.
--



2. Im Profil [Portrait des Oberkommandierenden der russischen Streitkräfte, Wjatscheslaw Tichomirow, in Tschetschenien]; 03.04.1996.
3. Seit Dezember 1994 mehr als 40.000 Tote [Über den Bürgerkrieg in Tschetschenien]; 02.04.1996.
4. Rußlands neuer Zar [Über den Bürgerkrieg in Tschetschenien]; 09.04.1996.
5. Tschetschenien: 'Der Krieg ist hier noch lange nicht vorbei'; 09.04.1996.
6. Präsident Jelzin und Lukaschenko unterzeichnen Unionsvertrag; 03.04.1996.
7. Ehe unter ungleichen Partnern [Über den Unionsvertrag zwischen Rußland und Weißrußland]; 03.04.1996.
8. Der nette Herr Stalin [Über die Propagandapolitik der Sowjetunion in den 30ern]; 11.04.1996.
9. Die Scheidelinie am 38. Breitengrad [Über den Grenzkonflikt zwischen Nord- und Südkorea]; 09.04.1996.
10. Korea: wem nutzen die Provokationen des Nordens?; 10.04.1996.

Tabelle. 9

Die textuelle Konnotation des zweiten Textbeispiels

1. Thema des Tages [Über BSE]; 04.04.1996.
2. In Brüssel hält jeder die Hand auf [Über BSE]; 01.04.1996.
3. Die Rechnung für den Rinderwahnsinn; 03.04.1996.
4. BSE und die Fleischindustrie; 03.04.1996.
5. Tauziehen um die EU-Hilfe [Über BSE]; 02.04.1996.
6. Krisentreffen in Luxemburg [Über BSE]; 02.04.1996.
7. Verfütterung von Tiermehl an Wiederkäuer; 04.04.1996.
8. Bundestag soll Produzenten in die Pflicht nehmen [Über BSE]; 12.04.1996.
9. Fleischwirtschaft will Vertrauen der Verbraucher zurückgewinnen [Im Zusammenhang mit dem BSE-Skandal]; 02.04.1996.
10. Nach 40stündigen Verhandlungen der Landwirtschaftsminister [Über BSE]; 04.04.1996.

## 5. Schlußfolgerungen

Der Algorithmus zur automatischen Repräsentation der konnotativen Bedeutung von Texten ist in der Lage, konnotative (Un-)Ähnlichkeiten von Texten, die auf Gebrauchsregularitäten lexikalischer Einheiten beruhen, abzubilden. Die vorgestellten Ergebnisse zeigen, daß der Algorithmus inhaltlich verwandte Texte einander zuordnet, während inhaltlich nicht zusammenhängende Texte dissoziiert werden. Die inhaltliche Ähnlichkeit von Texten wird in diesem Zusammenhang nicht auf der Basis von Okkurrenzen identischer, sondern *paradigmatisch ähnlicher* Wörter definiert. Texte werden somit nicht nur dann als inhaltlich verwandt und also konnotativ zusammenhängend erkannt, wenn sie eine größere Menge gemeinsamer Wortkonstituenten aufweisen, sondern falls sie paradigmatisch verwandte Wortkonstituenten umfassen. Darüber hinaus erlaubt gerade die Verwendung des Bedeutungsraumes als eines einheitlichen Repräsentationsformates die unmittelbare, wechselseitige Bezugnahme von Einheiten verschiedener Ebenen. Erweiterungen des Algorithmus sehen die Inkorporation weiterer Sprachebenen, so etwa die Ebene von Sätzen und Absätzen, vor. Zu diesem Zweck wird der Begriff der konnotativen Satzbedeutung als mittelbarer Konstituente der Textbedeutung erarbeitet.

## 6. Literatur

- Altmann, G.** (1995). Die Natur der Spracheinheiten. In M. Boroda (Hg.), *Musikometrika 6* (S. 1-12), Bochum: Brockmeyer.
- Barthes, R.** (1979). *Elemente der Semiologie*. Frankfurt a.M.: Syndikat.
- Barwise, J., & Perry, J.** (1987). *Situationen und Einstellungen. Grundlagen der Situationssemantik*. Berlin: de Gruyter.
- Dörner, D.** (1977). Superzeichen und kognitive Prozesse. In R. Posner & H.-P. Reinecke (Hg.) *Zeichenprozesse. Semiotische Forschung in den Einzelwissenschaften* (S. 73-82), Wiesbaden: Akademische Verlagsgesellschaft Athenaion.
- Eco, U.** (1988). Wer ist schuld an der Konfusion von Denotation und Bedeutung? Versuch einer Spurensicherung. *Zeitschrift für Semiotik*, 10, 3, 189-207.
- Eco, U., Santambrogio, M., & Violi, P.** (1988). *Meaning and Mental Representations*. Bloomington: Indiana University Press.
- Frege, G.** (1961). *Die Grundlagen der Arithmetik. Eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Hildesheim: Georg Olms.
- Frege, G.** (1994). *Funktion, Begriff, Bedeutung. Fünf logische Studien*. Göttingen: Vandenhoeck & Ruprecht.
- Goldblatt, R.** (1984). Topoi. The Categorical Analysis of Logic. *Studies in Logic and the Foundations of Mathematics*, 98.
- Halliday, M.A.K.** (1966). Lexis as a Linguistic Level. In C. Bazell, J. Catford, M.A.K. Halliday & R. Robins (Hg.), *In Memory of J. R. Firth* (S. 148-162), London: Longman.
- Halliday, M.A.K.** (1977). Text as Semiotic Choice in Social Context. In T.A. van Dijk & J.P. Petöfi (Hg.), *Grammars and Descriptions* (S. 176-225), Berlin, New York: de Gruyter.
- Halliday, M.A.K.** (1991). Towards probabilistic interpretations. In E. Ventola, (Hg.), *Functional and Systemic Linguistics* (S. 39-61), Berlin, New York: de Gruyter.
- Harris, Z.S.** (1954). Distributional Structure. *Word*, 10, 146-162.
- Harris, Z.S.** (1970) Discourse Analysis. In Z.S. Harris (Hg.), *Papers in Structural and Transformational Linguistics* (S. 313-348), Dordrecht: Reidel.
- Hjelmslev, L.** (1974a). *Prolegomena zu einer Sprachtheorie*. München: Max Hueber.
- Hjelmslev, L.** (1974b). Für eine strukturelle Semantik. In Hjelmslev, L.: *Aufsätze zur Sprachwissenschaft* (S. 105-119), Stuttgart: Klett.
- Klir, G.J., & Folger, T.A.** (1988). *Fuzzy Sets, Uncertainty, and Information*. New Jersey: Prentice Hall.

- Klir, G.J., & Yuan, B.** (1995). *Fuzzy Sets and Fuzzy Logic*. New Jersey: Prentice Hall.
- Kruse, R., Gebhardt, J., & Klawonn, F.** (1995). *Fuzzy-Systeme*. Stuttgart: Teubner.
- Labov, W.** (1978). Denotational Structure. In D. Farkas, W.M. Jacobson & K.W. Todrys (Hg.), *Papers from the Parasession on the Lexicon, Chicago Linguistic Society April 14-15* (S. 220-260), Chicago: University of Chicago.
- Lakoff, G.** (1988). Cognitive Semantics. In U. Eco, M. Santambrogio & P. Vio-li (Hg.), 119-154, Bloomington: Indiana University Press.
- Mehler, A.** (1996). A Multiresolutional Approach to Fuzzy Text Meaning. *Journal of Quantitative Linguistics*, 3, 2, 113-127.
- Mehler, A.** (1998). Modelling the Circularity of Connotative Word and Text Meaning Constitution. *Proceedings of the International Workshop on Computational Semiotics, IWCS'1997*, Int. Institute of Multimedia, Pôle Universitaire Léonard de Vinci, Paris, May 1997.
- Milne, P.** (1986): Frege's context principle. *Mind*, 491-495.
- Rescher, N., & Oppenheim, P.** (1955). Logical Analysis of Gestalt Psychology. *The British Journal for the Philosophy of Science*, VI, 22, 89-106.
- Rieger, B.** (1981). Feasible Fuzzy Semantics. On Some Problems of How to Handle Word Meaning Empirically. In H.J. Eikmeyer & H. Rieser (Hg.), *Words, Worlds, and Contexts. New Approaches in Word Semantics* (S. 193-209), (Research in Text Theory, Vol. 6), Berlin, New York: de Gruyter.
- Rieger, B.** (1989). *Unscharfe Semantik: die empirische Analyse, quantitative Beschreibung, formale Repräsentation und prozedurale Modellierung väger Wortbedeutungen in Texten*. Frankfurt a.M.: Lang.
- Rieger, B.** (1996). Situation Semantics and Computational Linguistics: towards Informational Ecology. In K. Kornwachs & K. Jacoby (Hg.), *Information: New Questions to a Multidisciplinary Concept* (S. 285-315), Berlin: Akademie.
- Sharkey, A.J.C., & Sharkey, N.E.** (1992). Weak Contextual Constraints in Text and Word Priming. *Journal of Memory and Language*, 543-572.
- Thiopoulos, C.** (1992). *Semiosis und Topoi*. Pfaffenweiler: Centaurus.
- Vanman, E.J., Read, S.J., & Miller, L.C.** (1995). Gestalt Principles and Parallel Constraint Satisfaction Processes: The Parallels. *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society 1995*, 766-770.
- Weizsäcker, E. von** (1974). Erstmaligkeit und Bestätigung als Komponenten der pragmatischen Information. In E. von Weizsäcker (Hg.), *Offene Systeme I. Beiträge zur Zeitstruktur von Information, Entropie und Evolution* (S. 82-113), Klett: Stuttgart.