

Unschärfe (fuzzy) Modellierung natürlichsprachlicher Bedeutung

Zu einer computerlinguistischen Neuorientierung der Semantik

*Burghard Rieger*¹

Lehrstuhl für Computerlinguistik

Fachbereich II: Linguistische Datenverarbeitung – Universität Trier

1 Einleitung

In der Computerlinguistik zeichnen sich derzeit weltweit Veränderungen ab, die ihre **Untersuchungsmethoden** ebenso wie ihre **Forschungsgegenstände** betreffen und damit eine Neubestimmung auch ihres **Erkenntnisinteresses** kennzeichnet, das die traditionellen Ansätze der computerlinguistischen Forschung wie der sprachverarbeitenden künstlichen Intelligenz in sehr grundsätzlicher Weise in Frage stellt (Rieger 98a).

Bekanntlich basieren nicht triviale, maschinelle Analyse- und Synthesysteme in der Sprachverarbeitung auf formalen Beschreibungen syntaktischer Strukturen und den diesen zuordenbaren semantischen Interpretationen. Diese sind in geeignete Formalismen zu überführen, damit eine Verarbeitung durch Automaten möglich wird. Hierbei lassen sich zwei Bereiche formaler Wissensrepräsentationen unterscheiden: das (syntaktische und lexikalische) **Sprachwissen** in Form von Produktionen (Ersetzungsregeln) sowie geeigneten Prozeduren zur kontrollierten Abarbeitung (Grammatikformalismen) einerseits und das (referenzielle und situative) **Weltwissen** in Form von prädikatenlogisch motivierten Strukturen (Propositionen) sowie Prozeduren zur inferenzfähigen Aufarbeitung (Wissensrepräsentationsformalismen) andererseits. Derartige Repräsentationsformalismen sind Bestandteil der eben deswegen **wissensbasiert** genannten Modellbildungen in der kognitiver Linguistik, die auf der Grundlage und mit Hilfe symbolanalytischer Techniken (monotone Logiken, symbolische Repräsentationen, regelbasierte Operationen, serielle Verarbeitung etc.) zunehmend komplexere Systeme entwickelte. Diese Modellbildungen sehen sich inzwischen von neueren Ansätzen herausgefordert, deren nach **Methodik** und **Gegenstand** verändertes und erweitertes **Interesse**, operable Lösungen in Bereichen zu entwickeln und anzuwenden möglich erscheinen läßt, die bisher als Problemfelder computerlinguistischer Forschung ausgespart blieben.

1.1 Die Herausforderung

Vor allem die Dynamik und anpassungsfähige Variabilität natürlicher kognitiver Systeme, welche selbst unvollständige oder vage Informationen sprachlicher Daten problemlos verarbeiten, sind deswegen zu Prüfsteinen geworden, die von *konnektionistischen Modellen* durch den Einsatz prozeßsimulierender Techniken (auf der Basis der Theorie dynamischer

¹E-mail: rieger@ldv.uni-trier.de

Systeme, sub-symbolischer oder verteilter Repräsentationen, kontinuierlich-numerischer Operationen, paralleler Verarbeitung) tatsächlich nachgebildet werden können (Nadel et.al. 1989; Pfeiffer et.al. 1989; Carpenter/Grossberg 1991). Gleichzeitig gelang es den *quantitativ-linguistischen Ansätzen* der empirischen Untersuchung sehr großer Sprachkorpora (auf wahrscheinlichkeits- und possibilitäts-theoretischer Basis, über stochastische bzw. fuzzy Modellbildungen, numerisch-mathematische und fuzzy-logische Operationen, durch strenge Verfahren der Hypothesenbewertung), vor allem den Reichtum an funktionalen Zusammenhängen sprachlicher Regularitäten und Strukturen bis hin zu universalen Gesetzmäßigkeiten erkennen zu lassen (Armstrong-Warwick 1993; Köhler/Rieger 1993), der große Textmengen tatsächlicher Sprachverwendung beobachtbar auszeichnet und Rückschlüsse auf Eigenschaften des menschlichen Sprachvermögens durchaus zu begründen vermag.

1.2 Die Aufgaben

Obwohl sich die Diskussion der Unterschiede konnektionistischer vs. regelbasierter Ansätze zunächst vor allem auf die mit den unterschiedlichen, auch *hybriden* Modellbildungen verbundenen Erklärungsansprüche (Pfeiffer et.al. 1989; Schnelle 1990; Hinton 1992) konzentrierte, hat die Auseinandersetzung mit den empirisch-quantitativen Ansätzen, deren Rezeption auch wegen der für Linguisten neuen und lernaufwendigen Untersuchungsmethodik verzögert wird, gerade erst eingesetzt (vgl. Rieger 1998). Immerhin scheint sich unter dem Eindruck des formalen Aufwands wie der praktischen Beschränktheit wissensbasierter Modelle einerseits, der überraschenden Leistungsfähigkeit stochastischer Parser (Church 1988; Briscoe 1993) und statistischer Übersetzungssysteme (Brown 1988; Brown 93) andererseits eine Revision der für die kognitive Linguistik verbindlichen Grundhypothesen abzuzeichnen. Diese besagen, daß die natürliche Sprache vor allem als Problem der Struktur und des Erwerbs von *Sprachvermögen (Kompetenz)* zu erforschen ist und daß dies – ohne die Kenntnis empirisch überprüfbarer Parameterwerte aus der Anwendung und Realisation solchen Vermögens in Situationen kommunikativer *Sprachverwendung (Performanz)* – auch (theoretisch) zu analysieren möglich und (formal) zu charakterisieren nötig sei (vgl. Chomsky 1988).

Ein Ergebnis dieser Auseinandersetzungen (Schnelle 1994) könnte sein, daß die Basis bisheriger kognitiv-linguistischer Forschung als eine zu weit gehende Abstraktion der sprachlich-kommunikativen Realität erkannt wird. Angesichts der Defizite komplexer regelbasierter Modellbildungen (Rieger 1988) in der kognitiven Linguistik (KL), Computerlinguistik (CL) und der sprachverarbeitenden künstlichen Intelligenzforschung (KI), und vor dem Hintergrund der zwar empirisch nachweisbaren, durch traditionelle Analysen mithilfe linguistischer Kategorien aber nicht erfaßten sprachlichen Strukturen,²

²Phänomene etwa der linearen Nah-Ordnung sprachlich performativer Einheiten (z.B. Kokkurenzen), deren Regularitäten sich regelbasierten Notationen entziehen aber in numerischen Ausdrücken von Korrelationen, mutual information Werten, etc. beliebig genau notiert und verarbeitet werden können, können als Resultate von Prinzipien verstanden werden, die kompetenztheoretisch bisher nur deswegen unbeachtet blieben, weil sie keine der regelbasierten Verkettungen linguistischer Kategorien betreffen, wie sie gängige Grammatikformalismen erfassen (Fucks 54), (Church 90), (Hindle 93).

wird eine empirisch-performative Linguistik bedeutsam gerade aufgrund ihrer komplexen Forschungsgegenstände und Untersuchungsmethoden. Durch ihre quantitativ-statistischen wie auch fuzzy-theoretischen Modellbildungen könnte so eine Neu- oder Umorientierung des kognitiv-linguistischen Erkenntnisinteresses eingeleitet werden, das den zeichentheoretisch motivierten Grundlagen einer funktionalen Semiotik näher steht als bisher.

Die Modellierung *performativer* Prozesse unterscheidet sich von der Wert-zuweisenden Auffüllung schon vorgegebener, weil *kompetenztheoretisch* vorausgesetzter Repräsentationsstrukturen dadurch, daß (diesen vergleichbare, neuartige) Repräsentationen als Strukturierungsleistung von den prozeßmodellierenden Prozeduren selbst erbracht werden. Ihre Resultate werden zunächst in Form von Verteilungen oder Vektoren notiert, die in der Regel hochdimensionierte (metrische) Raumstrukturen (*semiotische Räume*) bilden und deren Elemente als *unscharfe (fuzzy)* Mengen gedeutet werden können. So ließen sich anhand von (*morphischen*) Vektoren, die auf n-Grammen von Buchstaben basieren (Oerder 1994; Schneider 1994; Zubak 1994) und anhand von (*semischen*) Vektoren, die aus Korrelationen von Wörtern (Rieger 1982a; 1989; Reichert 1999) berechnet wurden, erstmals Zusammenhänge aufdecken und modelltheoretisch nutzen, die erst durch die vektoriiell repräsentierten *fuzzy* Elemente und Relationen des *morphologischen* bzw. *semantischen* Raumes erkennbar wurden. Inwieweit dabei Eigenschaften des Modells den Eigenschaften der Modellobjekte (Originale) entsprechen, wird im Einzelfall zu prüfen und sicherlich auch schwierig zu entscheiden sein. Sie sind aber eine Voraussetzung des besseren Verständnisses der semiotischen Funktionen solcher sprachlicher Entitäten, die unterschiedliche Ordnungen, Strukturen und Bedeutungen in natürlichsprachlichen Texten konstituieren, und die zu erfassen, zu analysieren und zu repräsentieren den bloß regelbasierten Ansätzen der traditionellen linguistischen Beschreibungen nicht hinreichend gelang.

2 Zur kognitiven Perspektive

Im Rahmen der *Kognitionswissenschaften*, deren Erkenntnisinteresse sich auf die Erweiterung der Kenntnisse über das Wissen (seine Formen, seine Strukturen, seines Erwerbs, seiner Anwendungen und seiner Realität im Denken) richtet (Johnson-Laird/Wason 1977), wird unter durchaus uneinheitlichen Wissenschaftsauffassungen anhand unterschiedlichster Forschungsgegenstände und mit den verschiedensten Untersuchungsmethoden an der theoretischen wie praktischen Präzisierung dessen gearbeitet, was *Kognition* ausmacht.

2.1 Kognitive Prozesse

Weitgehend unkontrovers scheint ein prozessuales Verständnis von Kognition zu sein. *Kognitive Prozesse* können danach – im weiten Sinn: systemtheoretisch – als adaptive Leistung von informationsverarbeitenden Systemen (*Agenten*) verstanden werden, die ihre Umgebungen in jeweils dem Maße in strukturierte Repräsentationen zu überführen

vermögen, wie dies ihre jeweils eigene Strukturiertheit erlaubt (Rieger 1995b; 1995d).

In den Kognitionswissenschaften heißen diese Prozesse *mentale Prozesse*. Sie gelten als Bedingung dafür, daß Regularitäten, Ordnungen und Strukturen nicht nur entstehen, sondern auch repräsentiert werden. Sofern es sich dabei um (nicht notwendigerweise auch zeichenhaft-symbolische) *Repräsentationen* von für das Zustandekommen von Erkenntnis notwendigen (Zwischen-)Resultaten handelt, wird angenommen, daß auch ihnen (andere) *Prozesse* zugeordnet sind, die einerseits das Entstehen solcher *Repräsentationen*, andererseits ihre Verarbeitung kontrollieren (Harnad 1982). Die *Semiotisierung* der kognitionswissenschaftlichen Sicht besteht – verkürzt gesagt – in der Suche nach solchen Prozessen (und ihren Modellen), die durch die Verarbeitung (von schon repräsentierten) Einheiten der einen Stufe die Repräsentationen (von Einheiten) einer anderen Stufe erst schaffen.

Während die **kognitive Linguistik** von formal konstruierbaren Repräsentationen ausgeht, die als Bedingung der Möglichkeit von auf ihnen operierenden mentalen Prozessen gelten, konfrontiert die **kognitive Sprachverarbeitung** kompetenztheoretisch relevante Ergebnisse und Resultate der kognitiven Linguistik mit der performativen Praxis der kommunikativen Produktion und Rezeption natürlichsprachlicher Texte. Hierzu fordert sie eine *ökologisch* orientierte Kognitionstheorie (Rickheit/Strohner 1993), die sich als *semiotischer* Ansatz für informationsverarbeitende Systeme darstellt (Rieger 1989) und thematisch für eine Weiterentwicklung solcher Systeme wurde. Diese Erweiterung bezieht konsequenterweise die Bedingungen wissenschaftlicher Kommunikation und Modellpluralität als besondere, weil explizierte Form situierter Produktion und Rezeption von Zeichen- und Symbolaggregationen ein.

Der Forderung eines ökologischen Paradigmas (Bateson 1979; Maturana/Varela 1980) auch für die kognitiv-wissensbasierte Computerlinguistik entspricht dabei eine dynamische Konzeption von Modellierung, deren Überprüfbarkeit weitgehend in der methodischen Realisierung kontextuell situierter *Prozesssimulationen* begründet ist. Diese können als *semiotische* Erklärung für das Entstehen von Strukturen, Ordnungen und Einheiten aus Vielheiten dann gelten, wenn sie – unabhängig von anderen kognitiven Erklärungsparadigmen – einen durch *Berechenbarkeit* kontrollierbaren, durch *Algorithmisierbarkeit* modellierbaren und durch seine *Prozeduralität* vermittelten Zusammenhang (Marr 1982) herstellen zwischen den Repräsentationen unterschiedlicher Analyse- und Darstellungsebenen. Während das Prinzip der Berechenbarkeit (*computational level*) mögliche Formate, Einheiten und Operationen über die Repräsentationen aller Beschreibungsebenen sprachlicher Phänomene festzulegen erlaubt, sind aus der Menge ihrer möglichen Algorithmisierungen (*algorithmic level*) nur diejenigen *semiotisch* interessant, die durch Verarbeitung der Einheiten einer Repräsentationsebene die Einheiten einer anderen Ebene erst konstituieren. Sie bilden offenbar eine Teilklasse der Algorithmen, die in Modellen symbolischer Repräsentationen von Einheiten und der regelbasierten (syntaktischen) Festlegung ihrer Konkatenationen gar nicht vorkommen, sondern überhaupt nur in Modellen mit verteilt repräsentierten Einheiten und nicht-syntaktischen Agglomerationen benötigt werden. Denn während die regelverarbeitenden Algorithmen in den symboli-

schen Modellen den Bereich der *Zuordnungen* abdecken, vermögen Algorithmen, die auf sub-symbolischen oder verteilten Repräsentationen operieren, offenbar den Bereich der *Korrespondenzen* zu modellieren. Diese *semiotischen* Algorithmen setzen im wesentlichen Einheiten unterschiedlicher Repräsentationsebenen derart zueinander in Beziehung, daß sie (mindestens) eine dieser Ebenen mit ihren Einheiten erst schaffen. Die Rede ist von *emergenten* Strukturen, welche bisher nicht-unterscheidbare Einheiten dadurch unterscheiden lassen, daß sie als Resultate von Prozessen erscheinen, welche die Daten, die sie verarbeiten, systematisch verändern.

2.2 Kognitive Modellierung

Weitgehend ungeklärt ist bisher, ob – und gegebenenfalls wie – *semiotische* Modellbildungen ein stufenweises Entstehen von Strukturen aus Ordnungen und dieser Ordnungen aus Regularitäten von Vielheiten beobachtbarer Entitäten erklären können. Es kann aber aufgrund vorliegender Untersuchungen vermutet werden, daß diese Prozesse der Identifikation von Regelmäßigkeiten und deren Zusammenfassung in solchen (Zwischen-)Repräsentationen, denen wiederum Eigenschaften von beobachtbaren Entitäten zukommen, für das Entstehen und die Verwendung zeichenhaft-funktionaler Strukturen in natürlichsprachlichen Systemen verantwortlich, wenn nicht mit ihnen identisch sind. Eine fundierte Hypothese dazu ist, daß sie einen durch *Berechenbarkeit* **formal kontrollierbaren**, durch *Algorithmisierbarkeit* **prozedural modellierbaren** und durch ihre *Implementationen* **praktisch realisierten** Zusammenhang überprüfbar herstellen zwischen Repräsentationen verschiedener Ebenen, die sie selbst erzeugen. Für kognitive Prozesse des Sprachverstehens bietet sich eine Modellierung in Form von mehrstufigen Verarbeitungs- und Repräsentationsebenen an, weil sie auf Gegebenheiten aufsetzen können, die als sprachliche Manifestationen selber schon (Zwischen-) Repräsentationen sind.

2.2.1 Quantitative Verfahren

Die quantitative Beschreibung und numerische Analyse sprachlicher Elemente, Einheiten und Strukturen bietet sich an, wenn es darum geht, Eigenschaften ihrer Verwendung, ihres Gebrauchs und der damit verbundenen Zusammenhänge zu ermitteln, die als (nicht unmittelbar beobachtbare) abgeleitete Funktionen ihres (beobachtbaren) Vorkommens beschrieben werden können. In Verbindung mit den *fuzzy-theoretischen* Möglichkeiten der Modellierung (Kruse et.al. 1993; Novák 1989) erlauben diese Verfahren die Definition von elastischen Einheiten (Zadeh 1975) – den *soft constraints* (Smolensky 1989) in sub-symbolischen Modellen entsprechend – durch numerische Spezifizierungen und erhöhtes Auflösungsvermögen von Zugehörigkeitsgraden verbunden mit größeren Toleranzen der Kategorisierung und der Verarbeitung (Zadeh 1994). *Weiche Kategorien* lassen sich so in ihren Reichweiten (Umfang) über die gleichen Formen numerischer Bestimmung kennzeichnen wie die Arten der ihnen subsumierbaren Elemente (Inhalt). *Unschärfe (fuzzy) Kategorien* heißen dabei solche abstrakten Zuordnungen, deren (leere) Strukturen ebenso wie deren mögliche Füllungen als Resultate von Prozessen erscheinen, die in Form von *Prozeduren* dargestellt werden können. Die *prozedurale* Form erlaubt dabei,

- *dynamische*, strukturelle Zusammenhänge als sich veränderliche Resultate von Prozessen zu erklären, deren (wiederholte) Durchläufe Veränderungen der (metrischen, topologischen) Struktur der Daten, auf denen sie operieren, zur Folge hat;
- diese Prozesse als Modell *kognitiver* Leistungen zu deuten, wodurch die Elemente (und Elementverbindungen) einer Ebene Strukturen anderer Ebenen zugeordnet werden, die durch diese Zuordnung erst entstehen;
- diese Zuordnung als prozedurale Explikation einer *Bedeutungskonstitution* zu verstehen, insofern bisher verborgenen (*hidden*) und uninterpretierten Einheiten und Strukturen durch die sie konstituierenden Prozeduren erkennbare (repräsentierende) Funktionen zukommen.

Kategoriale Konstrukte und Begriffsbildungen der strukturellen Linguistik, die immerhin aufgrund operationaler (wenn auch nicht streng algorithmisierbarer) Ansätze gefunden wurden, dienen hierbei als Leitlinie für die Suche nach algorithmisierbaren Prozeduren. Sie haben das Ziel, operational definierte, prozedural (re)konstruierende Konzepte zu entwickeln, denen überprüfbare, gegebenenfalls *weiche Kategorien* entsprechen, die in einer semiotischen Theorie linguistischer Performanz relevant werden.

Die Verbindung eines *semiotisch* fundierten Modells vektorieller Funktionsrepräsentationen linguistischer Einheiten mit einer dem semiotischen Gegenstandsbereich angemessenen mathematischen Analyse dieses Repräsentationsformats eröffnet begründete Aussicht darauf, neuartige Zusammenhänge zwischen den Erscheinungen *performativer* Sprachverwendung und den Prinzipien *kompetenten Sprachvermögens* aufzuweisen, deren quasi empirische Überprüfbarkeit durch die prozedurale Modellierung und experimentelle Simulation Zeichen-konstitutiver Prozesse gewährleistet werden kann.

2.2.2 Rekonstruktive Modelle

Für die Modellbildung kann auf ein Grundprinzip sprachlicher Strukturbildung zurückgegriffen werden, dessen Universalität³ in der spezifischen Form der Einschränkung liegt, welche die in beobachtbaren sprachlichen Regularitäten tatsächlichen realisierten Kombinationen von den theoretisch möglichen Kombinationen dieser sprachlichen Einheiten unterscheidet. Diese als lineare Verkettungsrelationen (**Syntagmatik**) und als selektive Ersetzungsrelationen (**Paradigmatik**) von (eben hierdurch unterschiedenen) linguistischen Einheiten wurde schon von den Begründern der strukturellen Linguistik in ihrer Systematik erkannt und zur Konstitution verschiedener Beschreibungsebenen sprachlicher Erscheinungen und ihrer Kategorisierung (Segment und Klasse) genutzt. Diese Unterscheidung kann durch den fuzzy-theoretischen Ansatz für die semiotische Modellbildung

³Es ist keine Sprache dieser Welt bekannt, die nicht den Strukturierungsprinzipien der *syntagmatischen* und *paradigmatischen* Restriktionen unterläge, wenn gleich wesentliche sprachtypologische Unterschiede auf dem unterschiedlichen Gebrauch beruhen, der in den verschiedenen Sprachen von diesen Prinzipien gemacht wird.

verschärft und numerisch präzisiert werden.⁴ Die hierbei zu verarbeitenden primären Eingabedaten sind Mengen von (hier: eindimensionalen) Verteilungen (Ketten) von Elementen, die aufgrund *syntagmatischer* und *paradigmatischer* Restriktionen erst als Einheiten unterschieden werden. Verarbeitungsergebnisse geeigneter Prozeduren werden wiederum als (diesmal: zwei-dimensionale) Verteilungen (Matrizen) ausgegeben, welche als Relationen, Vektoren oder hochdimensionierte Raumstrukturen (*semiotische Räume*) interpretiert werden können⁵.

Die auf *semantischer* und *morphologischer* Ebene schon durchgeführte Untersuchungen, die auf der Basis schriftsprachlicher Textdaten des Deutschen durchgeführt wurden, haben die (graphischen) Analoga so unterschiedlicher linguistischer Konzepte wie *Wortbedeutung* und *Silbe* als vektorbasierte Strukturierungsleistung von Prozeduren modellieren können. Grundlage hierfür bildeten Korrelationsmessungen und Rekurrenzanalysen, die auf linearen Folgen von Elementen (*Wörter* und *Graphen* bzw. Buchstaben) in natürlich-sprachlichen Texten großer Korpora aufsetzten. Die in Trier entwickelten Implementierungen solcher prozeduralen Modellierungen von *weichen* (*fuzzy*) linguistischen Kategorien beruhen dabei auf verteilten Repräsentationen, welche die Zustände eines dynamischen Systems möglicher Graphen-Agglomerationen bzw. Wort-Zusammenhänge partiell festlegen. In zwei unterschiedlichen Analysebereichen – dem *morphologischen Raum* mit (*morphischen*) Vektoren aus Prozeduren auf n-Grammen von Buchstaben (Rieger 1996) und dem *semantischen Raum* mit (*semischen*) Vektoren aus Prozeduren auf Korrelationen von Wörtern (Rieger 1991; 1997) – hat sich gleichermaßen gezeigt, daß die Fülle der in diesem Repräsentationsformat enthaltenen Informationen nur in dem Maße zugänglich und nutzbar ist, wie die Verfügbarkeit von Prozeduren zu ihrer Überführung in Strukturen, die (potentiell) mathematisch und/oder linguistisch⁶ interpretierbar sind.

2.3 Der semantische Raum

Anknüpfend an schon früher vorgelegte Operationalisierungen von Prinzipien der WITTGENSTEINschen Gebrauchssemantik und der Situationssemantik von BARWISE/PERRY, wonach sich die Bedeutungen der (Inhalts-) Wörter einer natürlichen Sprache als deren Gebrauch analysieren lasse, den die Sprecher in Situationen kommunikativer Sprachverwendung von ihnen machen, bietet auf der Ebene lexikalisch-semantischer Einheiten die Theorie der unscharfen (*fuzzy*) Mengen in Verbindung mit statistischen Verfahren der quantitativen Analyse großer Textkorpora die Möglichkeit, natürlich-sprachliche *Wortbedeutungen* in Form von Vektoren zu repräsentieren, deren Komponenten als Funktions-

⁴Daß damit kein Widerspruch zur *fuzzy* Modellierungstechnik entsteht, ist in der Konzeption *unscharfer Mengen* (Zadeh 65) begründet, für die die Verbindung von algebraischen Termen diskreter Entitäten mit numerischen Termen mathematischer Kontinua kennzeichnend ist.

⁵Diese *semiotischen Räume* werfen in ihrer Topologie wie in ihrer Metrik noch beträchtliche Probleme auf, die bei der Deutung und Interpretation (durch veranschaulichende Transformationen) selbst kleiner Ausschnitte – möglicherweise aber auch wegen dieser Fragmentierung – auftreten.

⁶Diese Strukturen analysierenden und generierenden Algorithmen, welche Konzepte und Techniken der *unscharfen* Modellierung bei der Transformation tatsächlicher Sprachdaten in (Zwischen-)Repräsentationen verschiedener Stufen nutzen, gehören in den Bereich der empirisch, (re-)konstruktiv arbeitenden Computerlinguistik, die als *Fuzzy Linguistics* (Rieger 98a) sich derzeit erst abzuzeichnen beginnt.

werte *syntagmatischer* und *paradigmatischer* Restriktionen von Wortverwendungsweisen berechnet wurden (Rieger 1981a; 1982a; 1982b).

2.3.1 Die textanalytischen Verfahren

Die hierzu auf der Wortebene angewandten Verfahren sind deskriptiv-statistischer Natur und beruhen im wesentlichen auf einer Korrelationsmessung α von (Wort-)Token in großen Korpora *pragmatisch-homogener* Texte sowie auf der *verteilten* Repräsentation der Verwendungsregularitäten ihrer (Wort-)Typen $y_i = \{\alpha(x_i, x_1) \dots \alpha(x_i, x_N)\}$. Deren Unterschiede – über ein Distanzmaß δ als Zwischenrepräsentation $z_i = \{\delta(y_i, y_1) \dots \delta(y_i, y_N)\}$ im *Corpusraum* $\langle C, \delta \rangle$ numerisch präzisiert – konstituieren den *semantischen Raum* $\langle S, \zeta \rangle$, der als ein (Mengen-)System abstrakter Bedeutungsrepräsentationen $z_i \in S$ (*Bedeutungspunkten*) bestimmt wird.

Für ein Vokabular $V = \{x_n\}, n = 1, \dots, i, j, \dots, N$ von Lexikoneinträgen (*Types*) werden die diesen zugeordneten Bedeutungen $z_n \in \langle S, \zeta \rangle$ als zweistufige Abbildungsfunktion der Verteilung der Gebrauchsweisen jeden Types zu allen übrigen

$$\alpha|x_n : V \rightarrow C; \{y_n\} =: C \quad (1)$$

und der Differenzen-Verteilung

$$\delta|y_n : C \rightarrow S; \{z_n\} =: S \quad (2)$$

berechnet. Dies geschieht mit Hilfe eines Koeffizienten, der die Intensitäten kookkurrierender Wörter in Texten mißt und in numerischen Werten des reellen Intervalls zwischen $[-1, +1]$ ausdrückt

$$\alpha(x_i, x_j) = \frac{\sum_{t=1}^T (h_{it} - e_{it})(h_{jt} - e_{jt})}{\sqrt{\sum_{t=1}^T (h_{it} - e_{it})^2 \sum_{t=1}^T (h_{jt} - e_{jt})^2}}; \quad -1 \leq \alpha(x_i, x_j) \leq +1 \quad (3)$$

Dabei sind die beiden Erwartungswerte e_{it} und e_{jt} , mit denen zwei Worttypen x_i und x_j theoretisch im Text k_t der Länge l_t des analysierten Korpus K vorkommen können

$$e_{it} = \frac{H_i \cdot l_t}{L} \quad \text{und} \quad e_{jt} = \frac{H_j \cdot l_t}{L}, \quad (4)$$

abhängig vom analysierten Textkorpus

$$K = \{k_t\}; \quad t = 1, \dots, T \quad (5)$$

und von seinem Umfang und der Länge seiner Texte, gemessen in der Anzahl der Wörter (Tokens)

$$L = \sum_{t=1}^T l_t; \quad 1 \leq l_t \leq T \quad (6)$$

sowie von den Häufigkeiten h_{it} bzw. h_{jt} , mit denen die paarweise betrachteten einzelnen Worttypen des Vokabulars $x_i, x_j \in V$ tatsächlich in jedem t -ten Text vorkommen

$$H_i = \sum_{t=1}^T h_{it}; \quad 0 \leq h_{it} \leq H_i \quad \text{bzw.} \quad H_j = \sum_{t=1}^T h_{jt}; \quad 0 \leq h_{jt} \leq H_j \quad (7)$$

Um die Unterschiede bemessen und numerisch präzisieren zu können, welche die Verwendungsweisen jeden Wortes zu allen anderen des Vokabulars unterscheidbar macht, wurde die EUKLIDische Distanz verwendet

$$\delta(y_i, y_j) = \left(\sum_{n=1}^N (\alpha(x_i, x_n) - \alpha(x_j, x_n))^2 \right)^{1/2} ; \quad 0 \leq \delta(y_i, y_j) \leq 2\sqrt{n} \quad (8)$$

Der Berechnung zunächst der α -Werte (3) aus den Eingabetexten und nachfolgend aus den Ergebnissen die Berechnung der δ -Werte (8) läßt die konsekutive Anwendung der beiden in (1) und (2) gegebenen Abbildungsfunktionen auch numerisch präzisieren. Die beiden Schritte der α - und δ -Abstraktion, die über Zwischenrepräsentationen – wie in **Abb. 1** als Matrizen – oder über Morphismen – wie in **Abb. 2** als Diagramm – die formale Rekonstruktion und empirische Berechnung verbinden, erweisen sich so als Kern einer auf syntagmatische und paradigmatische Restriktionen der Kombinierbarkeit von Zeichenaggregaten fußende Rekonstruktion der Bedeutungen von Wörtern als deren Gebrauch in Texten.

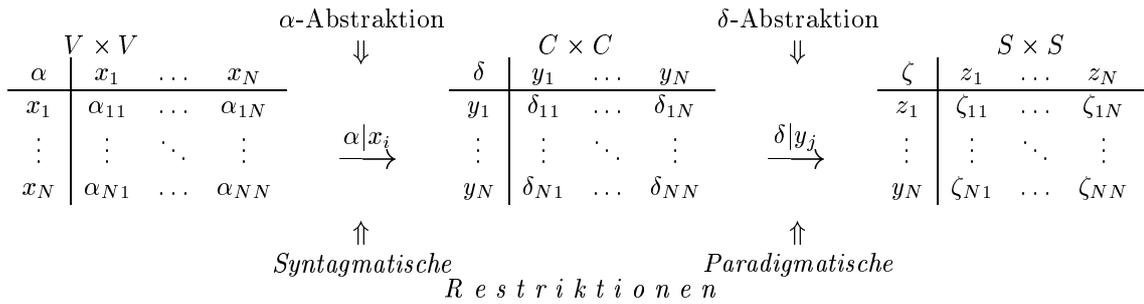


Abbildung 1: Formalisierung (*syntagmatischer/paradigmatischer*) Restriktionen (*constraints*) durch die zweistufige (α - und δ -) Abstraktion von Verwendungs-Regularitäten x_i über deren Ähnlichkeiten/Unterschieden y_i zu den Bedeutungspunkten z_i .

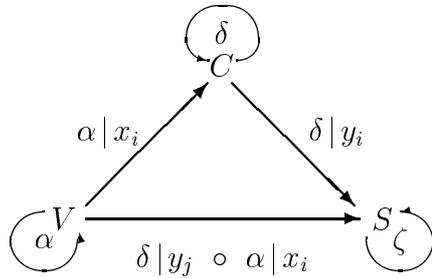


Abbildung 2: Abbildungsrelationen α and δ zwischen den strukturierten Mengen des Vokabulars $x_n \in V$, seiner Verwendungs-Regularitäten $y_n \in C$ und seiner Bedeutungspunkte $z_n \in S$.

Als Resultat des zweifachen Abstraktionsschritts (**Abb. 1**) läßt sich die Lage und Position jedes Bedeutungspunkts $z \in S$ in der vieldimensionalen metrischen Struktur des *semantischen Raum* $\langle S, \zeta \rangle$ auch als Funktion aller Unterschiede (δ - oder Distanzwerte) aller Verwendungsregularitäten (α - oder Korrelationswerte) der in den untersuchten Texten verwendeten Wörter des Vokabulars $x \in V$ deuten (**Abb. 2**), die als Komposition $\delta|y_j \circ \alpha|x_i$ zweier restringierter Abbildungen (*fuzzy mengentheoretische*) erklärt wird.

Als Zuordnung von *Wörtern* zu ihren *Bedeutungspunkten* stellt diese Funktion eine (mögliche) formale Repräsentation und empirisch überprüfbare Operationalisierung der *syntagmatischen* und *paradigmatischen* Restriktionen dar, welche Folgen von Wörtern in Texten pragmatisch homogener Korpora erkennen lassen. Die Visualisierung *unscharfer* Konzepte mithilfe eines Bäume generierenden Algorithmus zur Auffindung dispositio-neller Abhängigkeiten, der die Bedeutungspunkte nach ihrer Relevanz (*Kriterialität*) für die Bedeutung eines Konzepts (Wurzelknoten) reorganisiert, stellt dabei bisher eine einzigartige Möglichkeit dar, um eine anschauliche Vorstellung *relevanter* Ausschnitte des semantischen Raumes zu vermitteln (**Abb. 3**).

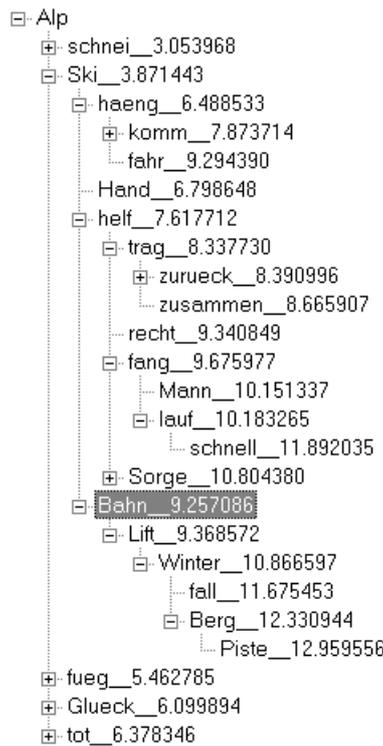


Abbildung 3: Fragment des DDS-Baumgraphen mit Wurzelknoten des Bedeutungspunktes $z_i = \text{ALPEN}$ in $\langle S, \zeta \rangle$, ermittelt anhand eines Teilkorpus deutscher Zeitungstexte (DIE WELT, 1964). Die numerischen Werte sind Punktdistanzen, die mit + markierten Knoten sind nicht-explizierte Teilbäume des generierten Graphen.

Die besonderen (formalen und inhaltlichen) Merkmale derart modellierter lexikalisch-

semantischer Beziehungen haben eine Reihe sehr interessanter Eigenschaften, deren Grundlage in der primär Distanz-bestimmten Repräsentationsweise zu sehen ist, welche ihnen die Charakteristika von *Stereotypen* (Rieger 1985b; 1985d; 1985e) verleiht. Deren semantische Verwandtschaften lassen sich anhand struktureller Eigenschaften des vieldimensionalen Raumes explizieren und auch zur Modellierung von *semantischen* – im Unterschied zu *syntaktischen* – Verknüpfungen nutzen, die von der vektoriellen Repräsentationsform der Bedeutungen ausgeht.

Sie erlaubt eine Interpretation der die Vektorkomponenten bestimmenden numerischen Werte – nach geeigneter Transformation – als Zugehörigkeitswerte von so definierten unscharfen (fuzzy) Mengen. Auf diesen sind – als Verallgemeinerungen der mengentheoretischen Explikationen prädikatenlogischer Operatoren \wedge (*und*), \vee (*oder*) und \neg (*nicht*) – numerische Operationen definiert, die unscharfe Mengen A und B als Repräsentationen vager Ausdrücke in analoger Weise – $\mu_{A \wedge B}$, $\mu_{A \vee B}$ und $\mu_{\neg A}$ – miteinander zu verknüpfen gestatten. Dies erlaubt die Konjunktion, Disjunktion und Negation von so repräsentierten Bedeutungen im semantischen Raum, als deren Resultate sich unscharfe (fuzzy) Mengen ergeben, die **neue Bedeutungspunkte** im semantischen Raum definieren (Rieger 1979a). Sie lassen sich anhand ihrer DDS-Baumgraphen veranschaulichen (**Abb. 4** und **Abb. 5**).

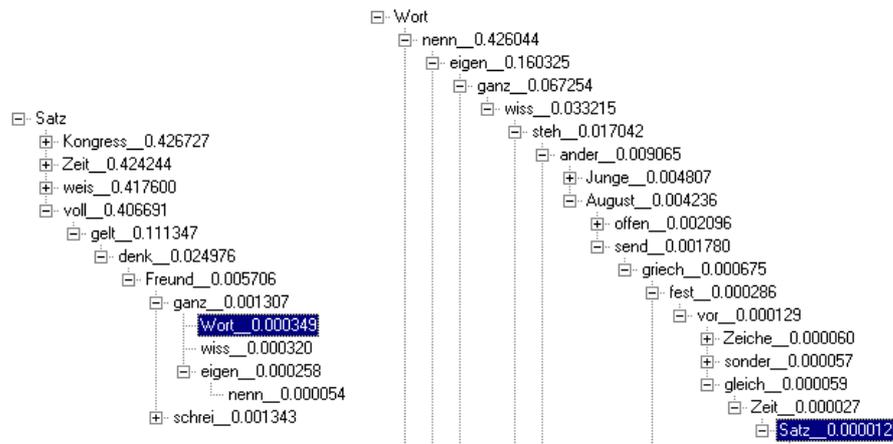


Abbildung 4: Fragmente der DDS-Bäume $z_i = \text{SATZ}$ und $z_i = \text{WORT}$, welche auch die schwache Symmetrie der wechselseitigen Abhängigkeiten erkennen lassen, d.h. nicht alle Knoten finden sich in beiden Abhängigkeitspfaden. Die Zifferangaben bei den Knoten sind ein Maß der Kriterikalität $[0,1]$, mit der jeder der Knoten im Baum zur Bedeutung der Wurzel beiträgt. Mit + markierte Knoten kennzeichnen verborgene Teilbäume.

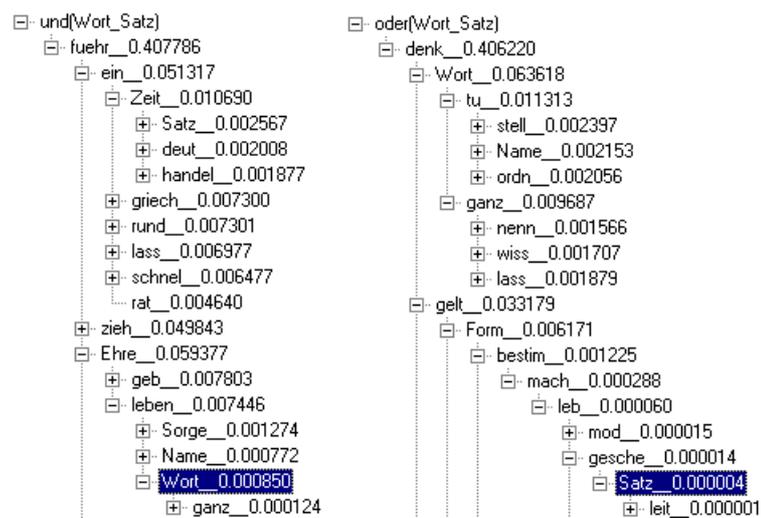


Abbildung 5: Fragmente der DDS-Bäume $z_{(i \wedge j)} = \text{WORT} \wedge \text{SATZ}$ bzw. $z_{(i \vee j)} = \text{WORT} \vee \text{SATZ}$ mit Kriteriäritäts-Werten der Knoten in den jeweiligen Dependenz-Pfaden. Man beachte, daß in beiden Bäumen die beiden durch den jeweiligen Operator verknüpften Bedeutungspunkte vertauscht erscheinen. Mit + markierte Knoten kennzeichnen wieder nicht-expandierte Teilbäume.

2.3.2 Der systemtheoretische Zusammenhang

Entscheidender als diese auf dem Repräsentationsformat aufsetzenden Operationalisierungen sind aber Organisationsprozesse, die – als **Prozeduren** modelliert – abstrakte Repräsentationsstrukturen erst liefern und zwar als variable Resultate der textanalytischen Verarbeitung sprachlicher Eingabeketten. Dies geschieht durch die algorithmische Rekonstruktion der *syntagmatischen* und *paradigmatischen* Beschränkungen, welche die formalen (Zwischen-)Repräsentationen über einen zweistufigen Prozeß als eine Art **Emergenz** von Zusammenhängen liefern. Solche emergierenden Zusammenhänge können als **Bedeutungen** deswegen gelten (Rieger 1991), weil sie vektorieLL repräsentierte und unterscheidbare Strukturen gleichzeitig mit bestimmten Zeichen (und Zeichenketten) verbinden, deren systematischen Gebrauch sie darstellen. Auf diese Weise ist der Vektor (oder die unscharfe Menge) nicht etwas, das dem Zeichen – gleichsam *extrinsisch* – zugeordnet würde, sondern diese Zuordnung ergibt sich – gleichsam *intrinsisch* – aus den systematischen Unterschieden der (Verwendungs-)Regularitäten dieser Zeichen in den analysierten Texten (Rieger 1991; 1994b; Thiopoulos 1992; Rieger/Thiopoulos 1993).

Die den Meß- und Abbildungsfunktionen α , δ und ζ dabei zugeordneten Analyse- und Repräsentationsalgorithmen können daher in ihrer aufeinander aufbauenden Mehrstufigkeit als rekonstruktiv-prozedurales Modell der *Verstehens*-Fähigkeit gedeutet werden.

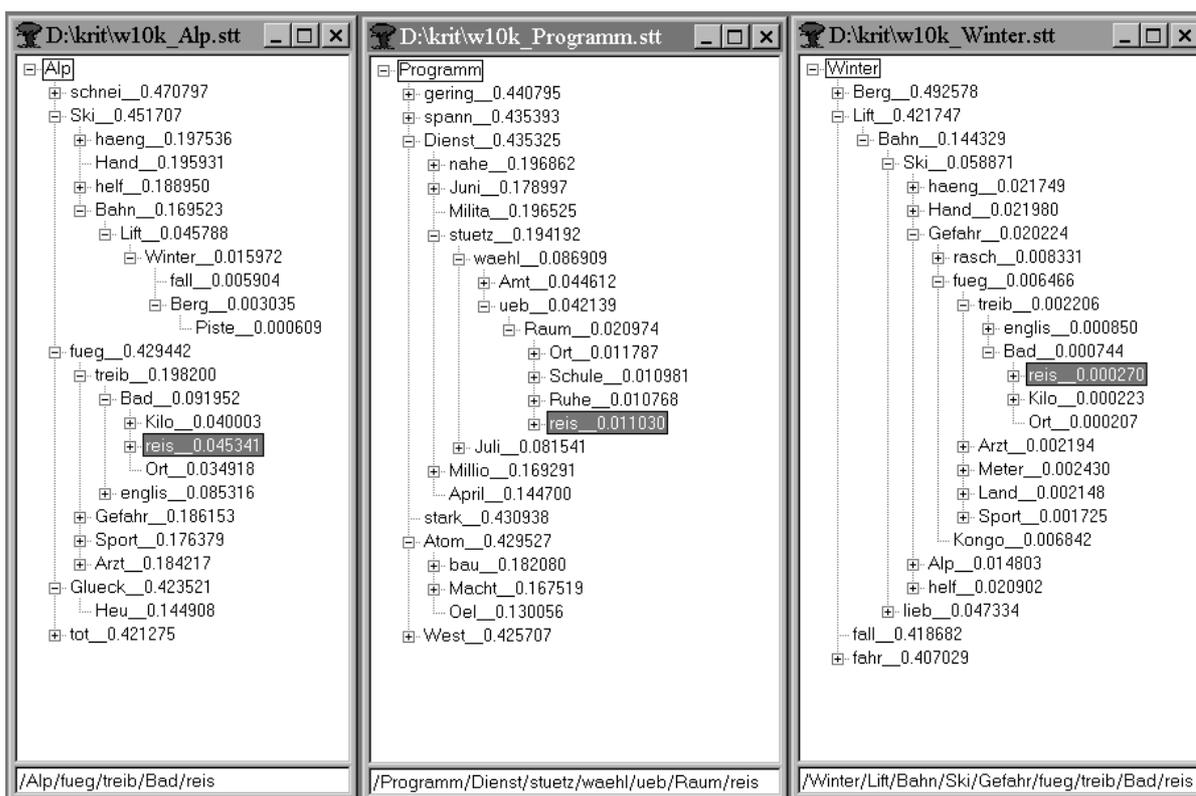


Abbildung 6: Semantische Inferenz von den Bedeutungspunkten ALP, PROGRAMM, WINTER (Prämissen) auf REIS(en) (Konklusion) mit den während des Schlußprozesses aktivierten *Dependenzpfaden* (im unteren Rahmen).

Das wird deutlich, wenn man sich klarmacht, daß die in der vektoruell bestimmten Struktur des mehrdimensionalen *semantischen Raumes* enthaltene Information bei weitem reicher ist, als dies die Lage und Position jedes einzelnen Bedeutungspunktes erkennen läßt. Da dieser Raum definitionsgemäß als Beziehungsstruktur sämtlicher Unterschiede aller Verwendungsregularitäten jedes einzelnen Wortes mit sämtlichen anderen des in den analysierten Texte verwendeten Vokabulars bestimmt ist, kann seine komplexe Topologie der Lagen und Positionen solcher Bedeutungsrepräsentationen vielfältige Nachbarschaften, Gemeinsamkeiten und Unterschiede erkennen lassen, die sich – wie durch veränderte Perspektiven und semantische Selektionen – algorithmisch veranschaulichen und visualisieren lassen.

Hierzu wurde eine rekursiv definierte Prozedur entwickelt und algorithmisiert, die Hierarchien von Bedeutungspunkten (Pfadknoten) in Abhängigkeit eines vorgegebenen Aspekts (Startknoten) als sog. dispositionelle Dependenzstrukturen (DDS) erzeugt. Der implementierte Algorithmus operiert auf der Datenstruktur des hochdimensionalen semantischen Raumes und überführt einen durch den Startknoten determinierten Teilraum in eine zweidimensionale Baumstruktur (vgl. **Abb. 3 – 5**). Die Prozedur induziert dabei auf der Basis der δ -Distanzen im *semantischen Raum* $\langle S, \zeta \rangle$ eine reflexive, schwach symmetrische und kontextuell transitive, semantische *Relevanzrelation* zwischen den selegierten Bedeu-

tungspunkten z_i (Rieger 1998c). Dadurch werden die in DDS-Bäumen re-organisierten Bedeutungspunkte zur geeigneten Grundlage einer kontextsensitiven Modellierung (Rieger 1998b) **semantischer Inferenzen** bzw. analoger Schlüsse (**Abb. 6**). Für jeden der m als **Prämissen** fungierenden Bedeutungspunkte (hier: $m = 3$) wird dabei je ein DDS-Prozeß initiiert. Während der parallelen Abarbeitung der Daten im *semantischen Raum* durch diese drei Baum-generierenden Prozeduren wird jeder der dabei verarbeiteten *Bedeutungspunkte* mit einer Marke versehen. Der Prozeß endet, sobald eine der m Prozeduren auf einen Bedeutungspunkt trifft, der schon von den anderen $m - 1$ Prozeduren (hier: 2) markiert wurde; dieser Bedeutungspunkt bildet gleichzeitig die semantische **Konklusion** des Inferenzprozesses.

Obwohl hierbei weder Propositionen und Sätze noch Prädikate und Eigenschaften wahrheitsfunktional bestimmt oder im Rahmen einer traditionell formal-semantischen Theorie analysiert werden, kann einem künstlichen kognitiven System, das über Analyse-, Repräsentations- und Kontroll-Mechanismen verfügt, die den α -, δ - und ζ -Funktionen entsprechen, sowie über die DDS- und Inferenz-Prozeduren, die auf den Bedeutungspunkten operieren, ein quasi kognitives Vermögen zugesprochen werden. Seine Fähigkeit, aus natürlichsprachlichen Eingabetexten und ihrer algorithmischen Verarbeitung eine interne Repräsentationsstruktur selbstorganisierend aufzubauen, die es zur Ableitung von Erwartungen (Dispositionen) nutzen und durch weitere Eingaben verändern und verfeinern kann (Lernen), macht ein solches System zu einem – wenn auch nicht-propositionalen – flachen Verstehensmodell.

Anders als alle Regel-basiert, Satz-semantisch und Wahrheits-funktional begründeten Verstehenskonzepte vermag ein auf den genannten Funktionen und Prozeduren aufgebautes, künstliches System der semiotischen kognitiven Informations-Verarbeitung (SCIPS) – ausschließlich durch Verarbeitung von natürlichsprachlichen Texteingaben – Bedeutungsrepräsentationen aus Zeichenverwendungen zu berechnen, die es ihm erlauben, *strukturelle Information* seiner (textuellen) Umgebung zu erkennen, zu verarbeiten und als sein (semantisches) *Wissen* in *Schlußprozessen* zu nutzen. Daß das System dabei in der Lage ist, aufgrund veränderter (textueller) Umgebungen auch sein eigenes (semantisches) Wissen kontinuierlich zu modifizieren, macht die *Dynamik* dieses Modells lernenden *Bedeutungserwerbs* aus. Dieser Prozeß kann als *Verstehens* gedeutet werden und kennzeichnet die adaptive Leistung des SCIP-Systemen, durch die es seine sprachliche Umgebung in strukturierte Repräsentationen überführt, und zwar in jeweils dem Maße, wie seine eigene Strukturiertheit dies erlaubt.

Aufgrund früherer Untersuchungen zur Bedeutungskonstitution (Rieger 1979b, 1980) und der prozeduralen Verarbeitung unscharfer (fuzzy) Bedeutungsrepräsentationen (Rieger 1983, 1984, 1985b), erscheint es auch im Licht der neueren Untersuchungen plausibel anzunehmen, daß *Bedeutungen* sprachlicher Einheiten als Resultate einer Klasse kognitiver Prozesse verstanden werden können, deren prozedurale Modellierung darin besteht, durch die Verarbeitung von schon repräsentierten Einheiten (der einen Stufe) die Repräsentationen von Einheiten (einer anderen Stufe) erst zu schaffen (Rieger/ Thiopoulos 1989; Rieger 94).

3 Zusammenfassung

Akzeptiert man ein eher **semiotisches Konzept der Semantik**, wie es auch – anders als in der **kognitiven Linguistik** – von performanztheoretischen Ansätzen der **kognitiven Sprachverarbeitung** nahegelegt wird, dann läßt sich Bedeutung kaum mehr als bloß relationale Zuordnung mehr oder weniger fixierter, jedenfalls weitgehend statischer Gegebenheiten analysieren und erklären, sondern sie wird eher als **dynamisch** und **kontextabhängig** sich veränderndes Resultat **mehrstufiger**, auf **lernend** sich verändernden Wissensbasen operierender **Konstitutionsprozesse** verstanden werden müssen, deren Relationalität in der Zuordnung von etwas besteht, das ohne diese Prozesse nicht faßbar ist: **Bedeutungen von Zeichen**, die nicht definiert oder verbal expliziert, sondern aufgrund ihres **kommunikativen Gebrauchs** in Texten **berechnet und dargestellt** werden.

Literatur

Armstrong-Warwick, Susan (Hg.) (1993): *Using Large Corpora I & II*. [Special Issues: *Computational Linguistics*, 19:1 & 19:2]. – Boston, MA: MIT Press for ACL.

Bateson, Gregory (1979): *Mind and Nature: A Necessary Unit*. – New York, (Dutton).

Briscoe, Ted; Carroll, John (1993): *Generalized Probabilistic {LR} Parsing of Natural Language (Corpora) with Unification-Based Grammars*. – In: *Computational Linguistics* (1993)1, 25-59.

Brown, Peter F.; DellaPietra, Stephen; DellaPietra, Vincent; Mercer, Robert L. (1993): *The Mathematics of Statistical Machine Translation: Parameter Estimation*. – In: *Computational Linguistics* (1993)2, 263-311.

Church, Kenneth W.; Hanks, Patrick W. (1990): *Word association norms, mutual information, and lexicography*. In: *Computational Linguistics* (1990) 1, 22-29.

Chomsky, Noam (1988): *Language and Problems of Knowledge. The Managua Lectures*. [Current Studies in Linguistics 16]. – 2. Aufl. – Cambridge, MA, London: MIT Press.

Fucks, Wilhelm (1954): *On Nahordnung and Fernordnung in samples of literary texts*. – In: *Biometrika* (1954) 2, 116-132.

Harnad, Stevan (1982): *A Unifying Theme for the Cognitive Science*. – In: T. Simon, R. Scholes (Hg.): *Language, Mind, and Brain*. – Hillsdale, NJ: Erlbaum, 1-12.

Hinton, Geoffrey (Hg.) (1992): *Connectionist Symbol Processing*. – Cambridge, MA: MIT Press.

- Hindle, Donald; Rooth, Mats (1993): *Structural Ambiguity and Lexical Relations*. – In: *Computational Linguistics* (1993) 1, 103-120.
- Johnson-Laird, Philip N.; Wason, Peter (Hg.) (1977): *Thinking. Readings in Cognitive Science*. [Open University Set Book]. – Cambridge, London, New York: Cambridge UP.
- Kruse, Rudolf; Gebhardt, Jörg; Klawonn, Frank (1993): *Fuzzy Systeme*. [Leitfäden und Monographien der Informatik]. – 1. Aufl. – Stuttgart: Teubner.
- Köhler, Reinhard; Rieger, Burghard B. (Hg.) (1993): *Contributions to Quantitative Linguistics. Proceedings of the 1st Quantitative Linguistics Conference – QUALICO-91*. – Dordrecht: Kluwer Academic Publishers.
- Marr, David (1982): *Vision*. – San Francisco: Freeman.
- Maturana, Humberto R.; Varela, Francesco J. (1980): *Autopoiesis and Cognition*. [Boston Studies in the Philosophy of Science 42]. – Dordrecht, Boston, London: Reidel.
- Novák, Vilaem (1989): *Fuzzy Sets and their Applications*. – Bristol, Philadelphia: A. Hilger.
- Oerder, Beate (1994): *Zum Problem linguistischer Kategorien: Untersuchungen der Silbenkonstitution im Deutschen und ihre unscharfe Modellierung*. – Magister-Arbeit (LDV/CL), FB II, Universität Trier.
- Pfeiffer, R.; Schreter, Z.; Fogelman-Soulié, F.; Steels, L. (Hg.) (1989): *Connectionism in Perspective*. – Amsterdam: (North Holland).
- Reichert, Maria (1999): *Konnektionistische Modellierung lexikalischer Prozesse der Bedeutungskonstitution*. – Diss., Universität Trier.
- Rickheit, Gert; Strohner, Hans (1993): *Grundlagen der kognitiven Sprachverarbeitung*. [UTB 1735]. – Tübingen, Basel: Francke.
- Rieger, Burghard B. (1979a): *Linguistic Semantics and the Problem of Vagueness: on analysing and representing word meaning*. – In: D.E. Ager, F.E. Knowles, J. Smith (Hg.): *Advances in Computer-aided Literary and Linguistic Research*. – Birmingham: AMLC University of Aston, 271-288.
- Rieger, Burghard B. (1979b): *Fuzzy Structural Semantics. On a generative model of vague natural language meaning*. – In: R. Trappl, P. Hanika, F. Pichler (Hg.): *Progress in Cybernetics and Systems Research*. – Washington, New York, London: Wiley & Sons, 495-503.

- Rieger, Burghard B. (1980): *Fuzzy Word Meaning Analysis and Representation in Linguistic Semantics. An empirical approach to the reconstruction of lexical meanings in East- and West-German newspaper texts.* – In: M. Nagao, K. Fuchi (Hg.): *COLING-80 8th International Conference on Computational Linguistics, International Committee on Computational Linguistics.* – Tokyo: ICCL, 76-84.
- Rieger, Burghard B. (1981a): *Unscharfe Semantik natürlicher Sprache. Zum Problem der Repräsentation und Analyse vager Wortbedeutungen.* – In: J. Scharff (Hg.): *Naturwissenschaftliche Linguistik. Leopoldina Symposium 1976.* – Halle: Barth, 251-276.
- Rieger, Burghard B. (1982a): *Fuzzy Representation Systems in Linguistic Semantics.* – In: R. Trappl, N. Findler, W. Horn (Hg.): *Progress in Cybernetics and Systems Research.* – Washington, New York, London: McGraw-Hill Intern., 249-256.
- Rieger, Burghard B. (1982b): *Procedural Meaning Representation. An empirical approach to word semantics and analogical inferencing.* – In: J. Horecky (Hg.): *COLING-82 Proceedings 9th International Conference on Computational Linguistics.* – Amsterdam, New York: North Holland, 319-324.
- Rieger, Burghard B. (1983): *Generating Dependency Structures of Fuzzy Word Meanings in Semantic Space.* – In: S. Hattori, K. Iounu (Hg.): *Proceedings of the XIIIth International Congress of Linguists, Comité International Permanent des Linguistes.* – Tokyo: CIPL, 543-548.
- Rieger, Burghard B. (1984): *Semantic Relevance and Aspect Dependency in a Given Subject Domain.* – In: D. Walker (Hg.): *COLING-84 10th International Conference on Computational Linguistics, International Committee of Computational Linguistics.* – Stanford: ACL-ICCL, 298-301.
- Rieger, Burghard B. (1985a): *Semantische Dispositionen. Prozedurale Wissensstrukturen mit stereotypisch repräsentierten Wortbedeutungen.* – In: B. Rieger (Hg.): *Dynamik in der Bedeutungskonstitution.* – Hamburg: Buske, 163-228.
- Rieger, Burghard B. (1985b): *Inducing a Relevance Relation in a Distance-like Data Structure of Fuzzy Word Meaning Representations.* – In: R. Allen (Hg.): *Data Bases in the Humanities and Social Sciences. 4th International Conference on Data Bases in the Humanities and Social Sciences, ICDBHSS/83.* – Osprey, FL: Paradigm Press, 374-386.
- Rieger, Burghard B. (1985c): *Lexical Relevance and Semantic Disposition. On stereotype word meaning representation in procedural semantics.* – In: G. Hoppenbrouwes, P. Seuren, T. Weijters (Hg.): *Meaning and the Lexicon.* – Dordrecht: Foris Publications, 387-400.
- Rieger, Burghard B. (1985d): *Stereotype representation and dynamic structuring of fuzzy word meanings for contents-driven semantic processing.* – In: J. Agrawal, P. Zunde

(Hg.): *Empirical Foundations of Information and Software Science*. – New York, London: Plenum Press, 273-291.

Rieger, Burghard B. (1988): *Theoretische Grundlagen der Computerlinguistik: eine Panneldiskussion mit M. Bierwisch, C. Habel, B. Rieger, H. Uszkoreit und W. Wahlster*. – In: I. Bátori, U. Hahn, M. Pinkal, W. Wahlster (Hg.): *Computerlinguistik und ihre theoretischen Grundlagen*. – Berlin, Heidelberg, New York: Springer, 192-218.

Rieger, Burghard B. (1989): *Unscharfe Semantik. Die empirische Analyse, quantitative Beschreibung, formale Repräsentation und prozedurale Modellierung vager Wortbedeutungen in Texten*. – Frankfurt/M., Bern, Paris: Peter Lang.

Rieger, Burghard B. (1991): *Distributed Semantic Representation of Word Meanings*. – In: J.D. Becker, I. Eisele, F.W. Mündemann (Hg.): *Parallelism, Learning, Evolution. Evolutionary Models and Strategies, {WOPPLOT}-89*. – Berlin, Heidelberg, New York: Springer, 243-273.

Rieger, Burghard B. (1994b): *Fuzzy Computational Semantics*. – In: H. Zimmermann (Hg.): *Fuzzy Systems. Proceedings of the Japanese-German-Center Symposium*. – Berlin (JGCB), 197-217.

Rieger, Burghard B. (1995b): *Meaning Acquisition by SCIPS*. – In: B.M. Ayyub (Hg.): *ISUMA-NAFIPS-95, IEEE-Transactions: Joint Intern. Conf. on Uncertainty Modeling and Analysis, North American Fuzzy Information Processing Society*. – Los Alamitos, CA: IEEE Computer Society Press, 390-395.

Rieger, Burghard B. (1995d): *Situations, Language Games, and Semiotic Agents*. – In: A. Meystel, N. Nerode (Hg.): *Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems, (1995 IEEE Monterey Workshop)*. Bala Cynwyd, PA: AdRem, 130-138.

Rieger, Burghard B. (1996): *Fuzzy Modellierung linguistischer Kategorien*. – In: H. Feldweg, E.W. Hinrichs (Hg.): *Lexikon und Text (Lexicographica Series Maior 73)*. – Tübingen: Niemeyer, 155-169.

Rieger, Burghard B. (1997): *Computational Semiotics and Fuzzy Linguistics. On Meaning Constitution and Soft Categories*. – In: A.M. Meystel (Ed.): *A Learning Perspective. Intern. Conf. On Intelligent Systems and Semiotics (NIST Special Publication 918)*. – Washington, DC: US Gov. Printing Office, 541-551.

Rieger, Burghard B. (1998a): *Warum Fuzzy Linguistik? Überlegungen und Ansätze einer computerlinguistischen Neuorientierung*. – In: D. Krallmann, H.W. Schmitz (Hg.): *Perspektiven einer Kommunikationswissenschaft*. – Münster: Nodus, 153-183

Rieger, Burghard B. (1998b): *Tree-like Dispositional Dependency Structures for non-*

propositional Semantic Inferencing. – In: B. Bouchon-Meunier, R.R. Yager (Eds.): *Proceedings of 7th Intern. Conf. on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU)*. – Paris: EDK Publ., 351-358.

Rieger, Burghard B. (1998c): *A Systems Theoretical View on Computational Semiotics. Modeling text understanding as meaning constitution by SCIPS*. – In: J.S. Albus (Ed.): *Proceedings of the Joint Conf. on the Science and Technology of Intelligent Systems (ISIC/CIRA/ISAS)*. – Piscataway, NJ: IEEE Publ., 840-845.

Rieger, Burghard B.; Thiopoulos, Constantin (1989): *Situations, Topoi, and Dispositions. On the phenomenological modelling of meaning*. – In: J. Retti, K. Leidlmair (Hg.): 5. *Österreichische Artificial-Intelligence-Tagung, Innsbruck-Igls*. – Berlin, Heidelberg, New York: Springer, 365-375.

Rieger, Burghard B.; Thiopoulos, Constantin (1993): *Semiotic Dynamics: a self-organizing lexical system in hypertext*. – In: R. Köhler, B.B. Rieger (Eds.): *Contributions to Quantitative Linguistics. Proceedings of the 1st Quantitative Linguistics Conference - QUALICO-91*. – Dordrecht: Kluwer Academic, 67-78.

Schnelle, Helmut (1994): *Welcher Sprachwissenschaft auf der Spur? Plädoyer für größere Klarheit und Toleranz*. – In: *Zeitschrift für Sprachwissenschaft* (1994) 1, 110-120.

Smolensky, Paul (1989): *Connectionist Modelling: Neural Computation / Mental Connections*. – In: L. Nadel, L. Cooper, P. Culicover, R. Harnish (Hg.): *Neural Connections, Mental Computation*. – Cambridge, MA, London: MIT Press, 49-68.

Thiopoulos, Constantin (1992): *Semiosis und Topoi*. – Diss. Universität Trier.

Zadeh, Lofti A. (1965): *Fuzzy Sets*. – In: *Information and Control* (1965) 8, 338-353.

Zadeh, Lofti A. (Hg.) (1975): *Fuzzy Sets and their Application to Cognitive and Decision Processes*. – New York, San Francisco: Academic Press.

Zadeh, Lofti A. (1994): *Fuzzy Logic, Neural Networks, and Soft Computing*. – In: *Comm. of the ACM*, 37(1994) 3, 77-84.