Semantische Analyse von Texten und Situationen (SATUS) – Semiotic Cognitive Information Processing System (SCIPS) – Language Learning and Meaning Acquisition (LLAMA)

SHOE related project research in SATUS, SCIPS, and LLAMA*

Burghard B. Rieger
Universität Trier FB II: Linguistische Datenverarbeitung/Computerlinguistik

February 7, 1992

1 Introduction

The common ground and widely accepted frame for modelling the semantics of natural language is to be found in the dualism of the rationalistic tradition of thought as exemplified in its notions of some independent (objective) reality and the (subjective) conception of it. According to this *realistic* view, the meaning of a language term (i.e. text, sentence, phrase, word, syllable) is conceived as something being related somehow to (and partly derivable from) certain other entities, called signs, a term is composed of. As a sign and its meaning is to be related by some function, called interpretation, language *terms*, composed of *signs*, and related *meanings* are understood to form some structures of entities which appear to be at the same time part of the (objective) reality and its (subjective) interpretation of it. In order to let signs and their meanings be identified as part of language terms whose interpretations may then be derived, some knowledge of these structures has to be presupposed and accessible for any symbolic information processing. Accordingly, *understanding* of language expressions can basically be identified with a process of matching some input strings with supposedly predefined configurations of word meaning and/or world structure whose *symbolic* representations have to be available to the (natural or artificial) understanding system's particular (though limited) *knowledge*. The so-called *cognitive paradigm* followed in advanced computational linguistics and artificial intelligence research can easily be traced back to stem from this fundamental duality, according to which natural language understanding will have to be modelled as *knowledge-based* processing of information.
Subscribing to this notion of understanding, however, tends to be tantamount to accepting certain unwarranted presuppositions of theoretical linguistics (and particularly some of its model-theoretical semantics) which have been exemplified elsewhere[1] by way of the formal and representational tools developed and used so far in cognitive psychology (*CP*), artificial intelligence (*AI*), and computational linguistics (*CL*). In accordance with these tools, *word meaning* and/or *world knowledge* is uniformly represented as a directed (more or less complex) graph with the (tacid) understanding that associating its vertices and edges with symbols from some established system of sign-entity-relationship (like e.g. that of natural language) will render such graph-theoretical configurations a model of structures or properties which are believed to be either those of the sign-system that

*Published in: Daelemans, W./Powers, D. (Eds.): Background and Experiments in Machine Learning of Natural Language. Proceedings First SHOE Workshop Tilburg University, Tilburg (Institute for Language Technology and AI) 1992, pp. 161–170
[1] Rieger 1991

provided the graph's labels, or those of the system of entities depicted by way of referential identification. Obviously, these representational formats are not meant to model the *emergence* of structures and the *processes* that constitute such structures as part of word meaning and/or world knowledge. Instead, these represenations are making use of them[2].

2 The Representational Issue

It has long been overlooked that relating arc-and-node structures with sign-and-term labels in symbolic knowledge representation formats is but another illustration of the traditional *mind-matter*-duality presupposing a realm of *meanings* very much like the structures of the *real world*. This duality does neither allow to explain where the structures nor where the labels come from. Their emergence, therefore, never occurred to be in need of some explanatory modelling because the existence of *objects, signs* and *meanings* seemed to be out of all scrutiny and hence was accepted unquestioned. Under this presupposition, fundamental *semiotic* questions of *semantics*—simply did not come up, they have hardly been asked yet[3], and are still far from being solved.

2.1 In following a more *semiotic* approach, this inadequacy can be overcome, allowing to avoid (if not to solve) a number of spin-off problems, which originate in the traditional distinction and/or the methodological separation of the meaning of a language's term from the way it is employed in discourse. It appears that failing to mediate between these two sides of natural language semantics, phenomena like *acquisition, creativity, dynamism, efficiency, learning, vagueness,* and *variability* of meaning—to name only the most salient—have fallen in between, stayed (or be kept) out of the focus of interest, or were being overlooked altogether, sofar. Moreover, there is some chance to bridge the gap between the formal theories of language description (*competence*) and the empirical analysis of language usage (*performance*) that is increasingly felt to be responsible for some unwarranted abstractions of fundamental properties of natural languages.

2.2 Modelling the meaning of an expression along reference-theoretical lines has to presuppose the structured sets of entities to serve as range of the denotational function which provided the expression's interpretation in order to let such an symbolic experession be understood. However, it appears feasible to have this very range be constituted as a result of exactly those cognitive functions by way of which understanding is produced as a process of emergence of structure. It may have to be modelled dynamically as the interaction of some system and its environment which reconstructs the possible structural connections as an identity (*structural coupling*) between the structures of expressions and those of the cognitive systems depending on the expressions' and the systems' pragmatics as specified by their *situational* setting.

3 Towards a Cognitive Semiotics

Approaching the problem from a *cognitive* point-of-view, identification and interpretation of external structures has to be conceived as some form of *information processing* which (natural/artificial) systems—due to their own structuredness—are (or ought to be) able to perform. These processes or the structures underlying them, however, ought to be derivable from—rather than presupposed to—procedural models of meaning[4]. Based upon a phenomenological reinterpretation of the analytical concept of *situation* as expressed by Barwise/Perry (1983) and the synthetical notion of *language game* as advanced by the late Wittgenstein (1958), the combination of both lends itself easily

[2] For illustrative examples and a detailed discussion see Rieger 1989, pp.103–132.

[3] see however Rieger (1977)

[4] It has been argued elsewhere (Rieger 1990, 1991) that *meaning* need not be introduced as a presupposition of *semantics* but may instead be derived as a result of semiotic modelling.

to operational extensions in empirical analysis and procedural simulation of associative meaning constitution which may grasp essential parts of what PEIRCE named *semiosis*[5].

3.1 In phenomenological terms, the set of structural constraints defines any cognitive (natural or artificial) system's possible range in constituting its schemata whose instantiations will determine the system's actual interpretations of what it perceives. As such, these cannot be characterized as a domain of objective entities, external to and standing in contrast with a system's internal, subjective domain; instead, the links between these two domains are to be thought of as *ontologically fundamental*[6] or pre-theoretical. They constitute—from a *semiotic* point-of-view—a system's primary means of access to and interpretation of what may be called its "world" as the system's particular apprehension of its environment[7]. Being fundamental to any cognitive activity, this basal identification appears to provide the grounding framework which underlies the duality of categorial-type rationalistic mind-world or subject-object separation.

3.2 From a systems-theoretical point-of-view, this is tantamount to a shift from *linear* to *non-linear* systems in modelling cognitive and semiotic behaviour. The simplest way to distinguish these approaches is by identifying the behaviour of *linear systems* as being equal to the sum of the behaviour of its parts, whereas the behaviour of *non-linear systems* is more than that of its parts. FREGEs principle of *compositionality* as well as CHOMSKEYs hypotheses of independance of syntax are concepts in point of the *linear*-systems'-view: by studying first the parts of a system in isolation will then allow for a full understanding of the complete system by composition of these parts. This collides with the *non-linear*-systems'-view according to which the primary interest is not in the behaviour of parts as properties of a system but rather in the behaviour of the *interaction* between parts of a system. Such interaction-based properties necessarily disappear when the parts are studied in isolation. This can be witnessed in referencial and model-theoretic semantics where phenomena like *vagueness, contextual variability* and *creative dynamism* cannot be dealt with, as well as in competence theoretical syntax where grades of *grammaticality, adaptive change* and *discourse adequacy* cannot be addressed. The *self-organizing* property of a non-linear, semiotic system has formally been derived elsewhere[8] and in some detail from mathematical *topos theory*[9] and *category theory*[10]. A first implementation of the system and its organisation as a dynamic *hypertext* structure has successfully been made to simulate the emergence of lexical meaning structures on the basis of—for that purpose rather coarsely measured—word co-occurrences in natural language texts[11].

4 Exploiting Syntagmatic Constraints

During my sabbatical 1991, I spent several months as visiting scholar at the International Computer Science Institute (ICSI) in Berkeley, University of California (UCB) with affiliation also to the Center for the Study of Language and Information (CSLI) at Stanford University. In consequence of the numerous discussions with members of these institutions—among which *Lotfi Zadeh* of the UCB Computer Science Department, *Jerry Feldman* of the ICSI, and *David Israel* of the CSLI have to be mentioned separately—my general interest in non-linear models of complex semiotic processes from SATUS now focusses on the investigation into aspects of *semiotic and cognitive information*

[5]"By *semiosis* I mean [...] an action, or influence, which is, or involves, a coöperation of *three* subjects, such as sign, its object, and its interpretant, this tri-relative influence not being in any way resolvable into actions between pairs." (Peirce 1906, p.282)

[6]Heidegger (1927)

[7]Maturana/Varela 1980

[8]Rieger/Thiopoulos 1989; Thiopoulos 1992

[9]Goldblatt 1979

[10]Bell 1981; Lambek/Scott 1986

[11]Rieger/Thiopoulos 1992

processing systems (SCIPS) and *language learning and meaning acquisition* (LLAMA) in respect of which presently appear to be particularly promising:

- miniature language acquisition studies in a non-referential environment,
- numerical exploitation of sub-symbolic constraints in NL discourse,
- model construction using memory augmented multi-layered networks.

4.1 Our earlier empirical approaches towards a system theoretical analysis and representation of word meaning from NL-texts emphasized the independence of any sentence parsing techniques. So far, this approach provides the procedural means of representing word meanings as a result of statistical and fuzzy-set-theoretical methods, which transform the linearity of strings of vocabulary items as used in discourse into the multi-dimensionality of their associated meanings. These could topologically be represented by points or vectors in a semantic space, allowing to be organized dynamically as tree-like semantic *dispositional dependency structures (DDS)*[12]. Based upon correlational analyses of co-occurring vocabulary items in texts, *DDS* do not consider their string distances. Thus, the approach is accounting for limited (paradigmatic) aspects of textual data only and gives away some of the linear (syntagmatic) structuredness inherent in any natural language string of items.

4.2 In a first approach, the correlational measure used so far had to be modified in order to allow for an *incremental* processing of texts, i.e. the computation of affinities and/or repugnancies of lexical items—text by text—in order to augment their overall computation as being exercised—all texts pooled—in text corpora.
Let K be the corpus of texts t and $\bar{K}$ the corpus increment

(1) $$K = \{k_t\},\ t = 1, \ldots, T$$

(2) $$\bar{K} = K \cup \{k_{\bar{t}}\},\ \bar{t} = T+1, \ldots, Inc$$

the length L or $\bar{L}$ respective for the increments given by the number of words summed up for all texts:

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

(4) $$\bar{L} = L + \sum_{\bar{t}=T+1}^{Inc} l_{\bar{t}}; \qquad 1 \leq l_{\bar{t}} \leq \bar{L}$$

the vocabulary V as the number of different words (types):

(5) $$V = \{x_n\} \ ; \ n = 1, \ldots, i, j, \ldots, N$$

and the frequency H with which each of these word types is found (tokens):

(6) $$H_i = \sum_{t=1}^{T} h_{it} \ ; \qquad 0 \leq h_{it} \leq H_i$$

(7) $$\bar{H}_i = H_i + \sum_{\bar{t}=T+1}^{Inc} h_{i\bar{t}} \ ; \qquad 0 \leq h_{i\bar{t}} \leq \bar{H}_i$$

then the modified correlation measure reads:

[12] Rieger 1985, 1990

(8)
$$\alpha(x_i, x_j) \;=\; \frac{\sum_{t=1}^{T}(h_{it} - h_{it}^*)(h_{jt} - h_{jt}^*)}{\sqrt{\sum_{t=1}^{T}(h_{it} - h_{it}^*)^2 \sum_{t=1}^{T}(h_{jt} - h_{jt}^*)^2}}\;;$$
$$-1 \leq \alpha(x_i, x_j) \leq +1$$
$$\text{with } h_{it}^* = \frac{H_i}{L} \cdot l_t \text{ and } h_{jt}^* = \frac{H_j}{L} \cdot l_t$$

with its (**bold-face**) 1st-increment:

(9)
$$\alpha_{Inc_1}(x_i, x_j) \;=\; \frac{\sum_{t=1}^{T}(h_{it} - h_{it}^*)(h_{jt} - h_{jt}^*)}{\sqrt{\sum_{t=1}^{T}(h_{it} - h_{it}^*)^2 \sum_{t=1}^{T}(h_{jt} - h_{jt}^*)^2}} + \mathrm{Inc}_1(i, j)$$

whose numerator reads:

$$\sum_{t=1}^{T}(h_{it} - h_{it}^*)(h_{jt} - h_{jt}^*) + (\mathrm{h}_{iT^{+1}} - \mathrm{h}_{iT^{+1}}^{\star})(\mathrm{h}_{jT^{+1}} - \mathrm{h}_{jT^{+1}}^{\star})$$

and whose denominator reads:

$$\sqrt{\left(\sum_{t=1}^{T}(h_{it} - h_{it}^*)^2 + (\mathrm{h}_{iT^{+1}} - \mathrm{h}_{iT^{+1}}^{\star})^2\right)\left(\sum_{t=1}^{T}(h_{jt} - h_{jt}^*)^2 + (\mathrm{h}_{jT^{+1}} - \mathrm{h}_{jT^{+1}}^{\star})^2\right)}$$

This will give the **incremental** correlation measure:

(10)
$$\alpha_{Inc}(x_i, x_j) \;=\; \frac{\sum_{t=1}^{T}(h_{it} - h_{it}^*)(h_{jt} - h_{jt}^*)}{\sqrt{\sum_{t=1}^{T}(h_{it} - h_{it}^*)^2 \sum_{t=1}^{T}(h_{jt} - h_{jt}^*)^2}} + \mathrm{Inc}_{\bar{t}}(i, j) =$$

$$\frac{\sum_{t=1}^{T}(h_{it} - h_{it}^*) \cdot (h_{jt} - h_{jt}^*) + \sum_{\bar{t}=T+1}^{Inc}(\mathrm{h}_{i\bar{t}} - \mathrm{h}_{i\bar{t}}^{\star})(\mathrm{h}_{j\bar{t}} - \mathrm{h}_{j\bar{t}}^{\star})}{\sqrt{\left(\sum_{t=1}^{T}(h_{it} - h_{it}^*)^2 + \sum_{\bar{t}=T+1}^{Inc}(\mathrm{h}_{i\bar{t}} - \mathrm{h}_{i\bar{t}}^{\star})^2\right)\left(\sum_{t=1}^{T}(h_{jt} - h_{jt}^*)^2 + \sum_{\bar{t}=T+1}^{Inc}(\mathrm{h}_{j\bar{t}} - \mathrm{h}_{j\bar{t}}^{\star})^2\right)}}$$
$$-1 \leq \alpha_{Inc}(x_i, x_j) \leq +1$$

where
$$\mathrm{h}_{i\bar{t}}^{\star} = \frac{H_i + \mathrm{h}_{i\bar{t}}}{L + \mathrm{l}_{\bar{t}}} \cdot \mathrm{l}_{\bar{t}} \quad \text{and} \quad \mathrm{h}_{j\bar{t}}^{\star} = \frac{H_j + \mathrm{h}_{j\bar{t}}}{L + \mathrm{l}_{\bar{t}}} \cdot \mathrm{l}_{\bar{t}}$$

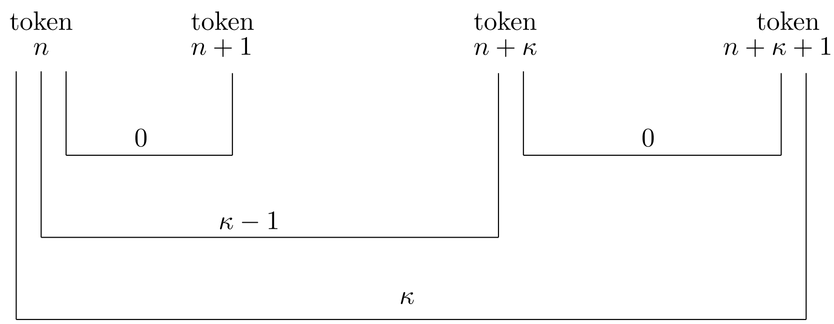

Figure 1: For any string of $n+2$ units, transitions of the order of κ can be defined

4.3 In a second vein inspired by stochastic processes as represented by *Markov Models (MM)*, their basic idea was generalized by way of higher order dependencies. Whereas *MM* make formation of any strings n dependent only on their $n-1$th element, the observable dependencies in string formation by linguistic entities of higher semiotic structures (phrases, clauses, sentences, texts) call for higher orders of control by the $n-2,\ldots,n-(n-1)$th units in each string extending step κ—corresponding to states in *Hidden Markov Models (HMM)*.
As the probability distributions of these state transitions are unknown (albeit all attempts to approximate them theoretically by conditional probabilities) and as they are furthermore subject to dynamic changes depending on *semiotically* constrained parameters, a procedural approach has been envisaged, that operates on empirically ascertained relative transition frequencies or Ω-matrices (*RTFN*s) whose order κ capture (in our case) each items' i differing (syntagmatic) influence on any other item j by the relative values $\bar{\omega}$ of absolute ω-transient frequencies according to

$$\bar{\omega}^{\kappa}_{ij} = \frac{\omega^{\kappa}_{ij}}{H_i \cdot H_j} \quad \text{where} \quad \omega^{\kappa}_{ij} = \frac{1}{L}\sum_{t=1}^{T} h^{\kappa}_{ij,t} \cdot l_t \tag{11}$$

$$\begin{array}{c|ccc} \Omega_0 & x_1 & \ldots & x_N \\ \hline x_1 & \bar{\omega}^0_{11} & \ldots & \bar{\omega}^0_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ x_N & \bar{\omega}^0_{N1} & \ldots & \bar{\omega}^0_{NN} \end{array} \quad ,\ldots, \quad \begin{array}{c|ccc} \Omega_{\kappa-1} & x_1 & \ldots & x_N \\ \hline x_1 & \bar{\omega}^{\kappa-1}_{11} & \ldots & \bar{\omega}^{\kappa-1}_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ x_N & \bar{\omega}^{\kappa-1}_{N1} & \ldots & \bar{\omega}^{\kappa-1}_{NN} \end{array} \quad , \quad \begin{array}{c|ccc} \Omega_{\kappa} & x_1 & \ldots & x_N \\ \hline x_1 & \bar{\omega}^{\kappa}_{11} & \ldots & \bar{\omega}^{\kappa}_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ x_N & \bar{\omega}^{\kappa}_{N1} & \ldots & \bar{\omega}^{\kappa}_{NN} \end{array}$$

Table 1: Ω-transition matrices of different order $0,\ldots\kappa-1$ and κ

4.4 The algorithm developed so far is still under testing. It produces tree-like graphs representing any vocabulary item's (root) tendency (numerical weight) in a decreasing top-to-bottom, left-to-right order which displays syntagmatic string regularities with other items (dependent notes on different levels).

5 Multi-layered and Simple Recurrent Networks

The expertise in neural networking and connectionist research assembled at the ICSI, in particularl that of *Joachim Diederich* and *Andreas Stolcke* drew my attention to a type of multi-layered network (MLN) which seems particularly suited for string processing and context sensitive, memory augmented adaptation to string regularities. Using propagation as the adapting mechanism—as most of the multi-layered architectures do—the Simple Recurrent Network (SRN) was inspired by a model studied by Jordan (1986) and further developed and modified by Elman (1988, 1989, 1990).

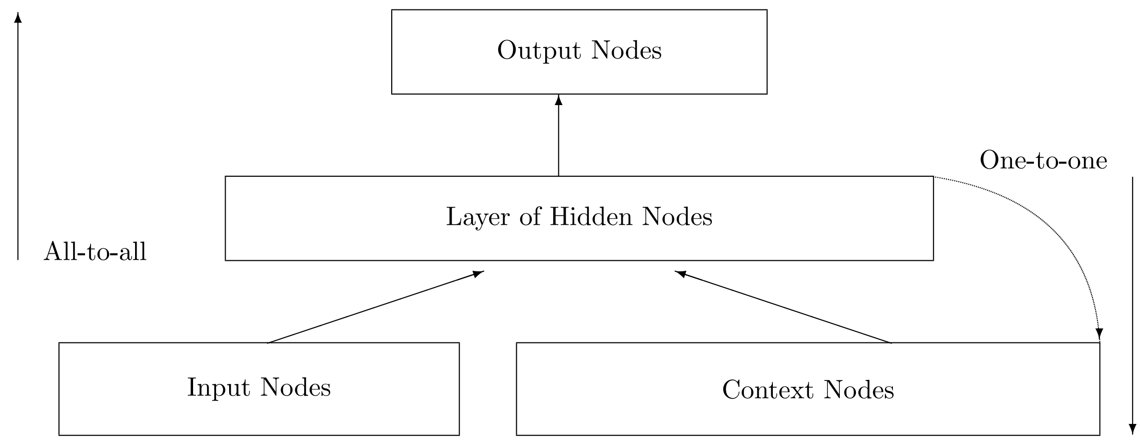

Figure 2: Schema of Elman-SRN (Elman 1989)

5.1 In addition to the input units, hidden units, and output nodes common to *MLN*, the *SRN* feature a set of context units which hold a copy of the hidden units activated from the prior cycle. On the next cycle this context units then feed back into the hidden units. These have the task of mapping the input to the output, and as the input includes their prior state of activation, these hidden units' states may record syntagmatic regularities emerging from contextual constraints. Thus, they can well be understood as the sort of memory the SRN is enhanced with.
It is this architecture's distributed rather than localist representation and its special form of recording sub-regularities in its hidden layers' states what makes *SRN* an attractive candidat for the remodelling of semantic state structures and semantic *DDS* (Rieger 1991). Both these constructs employ a distributed notion of memory where its "contents" is not associated with individual notes but rather with state vectors on the item types of the vocabulary which lends itself readily—at least in the current state of investigation—to be incorporated as context units in a *SRN* setup.

5.2 One of the problems we are faced with and are working on at present (without having solved it yet) is the great number of units needed with their increasing amount of fully connected networking (as the set of additional context units is to bear the one-to- one copy of the layer of hidden units). Reflections are underway whether—and if so, how—a recursively multi-layered architecture can be considered a realistic possibility to overcome this anticipated difficulty. The increasing numbers of units and the equally rising amount of necessary computing in an architecture which could cope with larger contexts being memorized is still beyond immediate realization for a *cognitive information processing system* that may, however, be envisaged for the future.
In case the basic idea of using such a "generalized" *SRN* for the remodelling of semantic *DDS* in a connectionist architecture proves to be feasable, when even the "syntactic" *DDS* (as outlined above) could be handled and processed by the same type of network. It would allow both, syntagmatic and paradigmatic constraints, to be used and modelled dynamically in an artificial system that would process natural language strings cognitively, i.e. in a way that is much more similar to the processing that natural cognitive systems perform.

References

Barwise, J./ Perry, J.(1983): Situations and Attitudes. Cambridge, MA (MIT)

Bell, J. L. (1981): "Category theory and the foundation of mathematics" *British Journal of the Philosophy of Science,* 32,1981:349–358

Elman, J.L. (1989): Structured representation and connectionist models. (CRL-Report-No. 8901) *Center for Research in Language* UCSD

Elman, J.L. (1990): "Finding Structure in Time", *Cognitive Science* 14,1990:179–211; also: (CRL-Report-No. 8801) of April, 1988

Goldblatt, R. (1984): Topoi. The Categorial Analysis of Logic. (Studies in Logic and the Foundations of Mathematics 98), Amsterdam (North Holland)

Heidegger, M. (1927): Sein und Zeit. Tübingen (M.Niemeyer)

Jordan, M.I. (1986): Serial Order: a parallel distributed processing approach. (ICS-Report-No. 8604) *Institute for Cognitive Science* UCSD

Lambek, J./ Scott P. J. (1986): Introduction to higher order categorical logic. Cambridge (Cambridge University Press)

Maturana, H./ Varela, F. (1980): Autopoiesis and Cognition. The Realization of the Living. Dordrecht (Reidel)

Peirce, C.S. (1906): "Pragmatics in Retrospect: a last formulation" (CP 5.11 – 5.13), in: The Philosophical Writings of Peirce. Ed. by J. Buchler, NewYork (Dover), pp.269–289

Rieger, B. (1977): "Bedeutungskonstitution. Einige Bemerkungen zur semiotischen Problematik eines linguistischen Problems" *Zeitschrift für Literaturwissenschaft und Linguistik* 27/28, pp.55–68

Rieger, B.B. (1985b): "On Generating Semantic Dispositions in a Given Subject Domain" in: Agrawal, J.C./ Zunde, P. (Eds.): Empirical Foundation of Information and Software Science. NewYork/ London (Plenum Press), pp.273–291

Rieger, B. (1989): Unscharfe Semantik. Die empirische Analyse, quantitative Beschreibung, formale Repräsentation und prozedurale Modellierung vager Wortbedeutungen in Texten. Frankfurt/ Bern/ NewYork (P. Lang)

Rieger, B.B. (1990): "Situations and Dispositions. Some formal and empirical tools for semantic analysis" in: Bahner, W. / Schildt, J./ Viehweger, D. (Ed.): Proceedings of the XIV.International Congress of Linguists (CIPL), Vol. II, Berlin (Akademie Verlag), pp.1233–1235

Rieger, B.B. (1991): "On Distributed Representation in Word Semantics" (ICSI-Report TR-91-012) International Computer Science Institute, Berkeley, CA

Rieger, B.B. (1991): "Reconstructing Meaning from Texts. A computational view on natural language understanding", Proceedings of the *2nd German Chinese Electronic Congress (GCEC–91) at Shanghai*, Berlin/ Offenbach (VDE-Verlag), pp.193–200

Rieger, B.B./ Thiopoulos, C. (1989): "Situations, Topoi, and Dispositions. On the phenomenological modelling of meaning." in: Retti, J./ Leidlmair, K. (Eds.): 5th Austrian AI-Conference ÖGAI-89. (KI–Informatik Bd.208) Berlin/ NewYork (Springer), pp.365–375

Rieger, B.B./ Thiopoulos, C. (1992): "Semiotic Dynamics: A self-organizing lexical system in hypertext", in: Köhler, R./ Rieger, B.B. (Eds.): Proceedings of the *1st Quantitative Linguistics Conference – QUALICO-91*, Amsterdam (Elsevier Science) 1992 [forthcoming]

Rieger, B.B./ Badry, P./ Reichert, M. (1992): "Sub-symbolic Control Structures: synthesizing constraints from syntagmatic and paradigmatic sub-regularities in NL-texts for language understanding and generation" [in preparation]

Thiopoulos, C. (1990): "Meaning metamorphosis in the semiotic topos." *Theoretical Linguistics* 16:2/3, pp.255–274

Thiopoulos, C. (1992): Semiosis and Topoi. *PhD Diss. 1991*: Dept. of Computational Linguistics, FB II: LDV/CL, Universität Trier [forthcoming]

Winograd,T./ Flores, F. (1986): Understanding Computers and Cognition: A New Foundation for Design. Norwood, NJ (Ablex)

Wittgenstein, L. (1958): The Blue and Brown Books. Ed. by R. Rhees, Oxford (Blackwell)