Semiotics and Computational Linguistics

On Semiotic Cognitive Information Processing*

Burghard B. Rieger**
FB II: Department of Computational Linguistics
University of Trier, Germany

Abstract

Signs, which are the domain of inquiry in *semiotics*, have a complex ontology. Apart from being used—adequate knowledge provided—by communicators, and recognized as being decomposable into smaller elements and aggregatable to larger structures, they are also meant to be understood. This is a consequence of their manifold identity as compound physical *objects* with real world extensions in space-time-locations a n d as activators for complex mental *processes* which tend to be identified with some mind and/or brain activities responsible for their *understanding*. In the *cognitive sciences* all processes of perception, identification, and interpretation of (external) structures are considered *information processing* which (natural or artificial) systems—due to their own (internal) structuredness or knowledge—are able (or unable) to perform. Combining the *semiotic* with the *cognitive* paradigm in *computational linguistics*, the processes believed to constitute natural language sign structures and their understanding is modeled by way of procedural, i.e. computational (re-)constructions of such processes that produce structures comparable to those that the understanding of (very large) samples of situated natural language discourse would imply. Thus, *computational semiotic* models in cognitive linguistics aim at simulating the constitution of meanings and the interpretation of signs without their predicative and propositional representations which dominate traditional research formats in syntax and semantics so far. This is achieved by analyzing the linear or *syntagmatic* and selective or *paradigmatic* constraints which natural languages impose recursively on the formation and structure of (strings of) linguistic entities on different levels of systemic distinction. It will be argued (and illustrated) that *fuzzy* modeling allows to derive more adequate representational means whose (numerical) specificity and (procedural) definiteness may complement formats of categorial type *precision* (which would appear phenomenologically incompatible) and processual *determinateness* (which would seem cognitively inadequate). Several examples from *fuzzy linguistic* research will be given to illustrate these points.

 Revised version of a paper presented at the Pre-ISAS Tutorials for the *Intelligent Systems and Semiotics* conference, National Institute of Standards and Technology (NIST), Gaithersburg, Ma. USA, on September 22–25, 1997

**E-Mail: `rieger@ldv.Uni-Trier.de`

The author is indebted to *Kathrin Gieseking, Alexander Mehler, Beate Oerder* and *Maria Reichert* for numerous discussions of ideas expressed in this paper. As always, the errors are his own.

Although we plainly can say nothing about matters that lie beyond our current understanding, it is difficult to say why one should retain the faith that traditional conceptions will somehow be applicable there, even though we find them generally useless to the extent that we come to understand some aspects of the nature of organisms, in particular, the mental life of humans.[1]

1 Introduction

Anything we know or believe about the world can—more or less precisely—be communicated verbally. We do so by language means, employing words, forming sentences, producing texts whose meanings are understood to convey, stand for, designate, refer to or deal with topics and subjects, entities and domains, structures and processes in the real world. Natural language texts (still) are the most flexible and as that highly efficient means to represent knowledge for and convey learning to others. What appears to be conditional for this kind of text understanding is humans' language faculty, i.e. the (*performative*) ability to identify, recognize, produce, and structure some fragments of real world stimuli according to some internal—though externally conditioned—principles (*competence*). Traditional approaches in linguistics proper (LP), computational linguistics (CL) and artificial intelligence research (AI) have developed structural and procedural conceptions for (parts of) the process of language understanding. Their notational systems employed allow for the distinction of *linguistic knowledge* formally represented in rule based formats, and of *world knowledge* whose structuredness is mediated by symbol representational formats which are combined to model language processing by machine. However, important features characteristic of natural language understanding processes, like e.g. vagueness, robustness, adaptivity, dynamism, etc. had to be overlooked or intentionally put aside because of the representational formats chosen and their processing possible.

Computational Semiotics (CS) neither depends on rule-based or symbolic formats for (linguistic) knowledge representations, nor does it subscribe to the notion of (world) knowledge as some static structures that may be abstracted from and represented independently of the way they are processed. Instead, knowledge structures and the processes operating on them are modeled as procedures that can be implemented as algorithms. *Semiotic Cognitive Information Processing* (SCIP) systems are defined as collections of cognitive information processing devices whose semiotics consists in their multi-level representational performance of (working) structures emerging from and being modified by such

[1]Chomsky: The Managua Lectures, 1986, p. 48

processing. The emergence of sign structures as a self-organizing process may in particular be studied on the basis of combinatorial and selective constraints universal to all natural languages. Both, (linguistic) entity formation and (semiotic) function acquisition may thus be reconstructed from *syntagmatic* constraints observed in linear agglomeration, and *paradigmatic* constraints on selectional choice of elements of natural language sign structures in discourse. Their regularities are exploited by text analyzing algorithms operating on and defining levels of *morpho-phonemic, lexico-semantic, phraseo-syntactic* and *situational* or *pragma-semantic* representation. Initially, the algorithms accept natural language discourse as input and produce vector space structures as output. These may be interpreted as intermediate (internal) representations on different levels of a semiotic system's states of adaptation to the (external) structures of its environment as signaled and mediated by the natural language discourse processed.

2 Modeling Cognition

The alliance of logics and linguistics, mediated mainly by (language) philosophy in the past and by (discrete) mathematics since the first half of this century, has long been (and partly still is) dominating the way in what terms natural languages expressions should be explicated and how their processing could be modeled. It is ironic that the dramatic increase of computational power and symbol manipulation means has changed the fundamentals of many scientific disciplines, creating even new ones, but has left linguistically oriented disciplines, even new ones, adhere to the lore of seemingly well grounded and traditionally dignified concepts (like *phrase* and *sentence*, *predicate* and *proposition*, *grammatical correctness* and *formal truth*, etc.) in describing natural language structures and their processing. Considering our as yet very limited understanding of natural language understanding which explicatory cognitive models as well as implemented operational systems of computational natural language processing demonstrate, it may well be suspected that some of the problems encountered by these model constructions are due to the representational formats they employ in depicting and manipulating entities (elements, structures, processes, and procedures) considered to be of interest or even essential to the understanding of the communicative use of natural languages by humans.

2.1 Information Systems View

Following a systems theoretical paradigm of information processing and accepting the cognitive point-of-view according to which information processing is knowledge based, humans appear to be far from being just another species of natural information processing systems with some higher cognitive abilities.

Instead, they have to be considered very particular cognitive systems whose outstanding plasticity and capability to adapt their behavior more rapidly to changing environmental conditions than others, is essentially tied to their use and understanding of natural languages in communicative discourse. It seems that the language faculty expands their learning potential well beyond experimental experience into realms of virtual reality (*Gedankenexperimente*) which allows for the recognition of consequences inferred from *potentially real* instead of *factually existing* conditions for survival[2]. The basic idea of model construction in terms of such an ecological theory of information systems [39] is that the processing structure of an information system is a correlate of those structures which that system is able to process in order to survive. For cognitive models of natural language processing the systems theoretical view suggests to accept natural language discourse in situations of communicative interaction as analyzable and empirically accessible evidence for tracing such processes. Thus, situated natural language communication might reveal essential parts of the particularly structured, multi-layered information representation and processing *potential* to a system analyzer and model constructor in rather the same way as this potential is accessible to an information processing system. The difference here, however, between the system and its analyzer is that they are active in and part of different information processing situations of which only the former—and not the latter—can be said to be properly attuned. It is this lack of attunement to the *semiotics* underlying situational natural language understanding which prompts cognitive linguists as system analyzers to fall back on their attunement to *situations* of language understanding [2]. But whereas in language understanding one can, and even has to take the semiotic dimension of sign and meaning constitution for granted (and beyond questioning) in order to let any particular sign or meaning be understood, the purpose of modeling that very process must not in aiming at the conditions for the possibility of such processes of understanding. Hence, a system analyzer and model constructor in semiotics should not rely (solely) on linguistic categories in describing and modeling semiotic entities. She/he has to make any provision that her/his ideas about the modeling of both, the representation a n d the processing are not unduly pre-defined by long established, but possibly inadequate concepts and related formats. Rule-based models of syntactic processing as well as truth-functional models of (sentence) meaning appear to be as inadequate as predicative and propositional formats of semiotic entity representation and processing.

[2] This expansion—however *genotypically* advantageous—has *phenotypical* drawbacks for individual semiotic systems whose *virtual* realities may become *vicious* to the extend they might replace instead of complement the systems' factual situatedness. Illustrating the case is a tendency to blur the distinction of *reality* from *fiction* due to the (semiotic) media's increasing degree of experiential (situational) density: print literature < theatre < film < television < VR environments.

2.2 Semiotic Attunement

In a systems theoretic approach, *attunement* obviously replaces the notion of *static* knowledge structures as realized in cognitive information processing models so far, by a *dynamic* conception of structuredness which defines knowledge as an open, modifiable, and adaptive system whose organization can be conceived as a function of the system's own processing results (*knowledge acquisition*). This, however, can only be achieved by allowing *semiotic entities* to have their own[3] (perhaps yet unknown) ontology which is not (or not fully) accounted for by predicative and propositional representations or rule-based and truth-functional formats which tacitly make believe that semiotic entities can be characterized and their functions be modeled exclusively by these categorial structures and associated processing of symbol manipulation. Instead, *semiotic modeling* is to find and employ representational formats and processing algorithms which do not prematurely decide and delimit the range of semiotically relevant entities, their representational formats and procedural modes of processing. One of the advantages of semiotic models would be that the entities considered relevant would not need to be defined prior to model construction but should emerge from the very processing which the model simulates or is able to enact. It appears that—if any—this property of models does account for the intrinsic (co- and contextual) constraining of the meaning potential characteristic of natural language discourse which renders them *semiotic* in a (meaning or function) constituting sense which is the core of *understanding*. Representing a system's environment (or fragments thereof) in a way, that such representations not only take part in that system's direct (*immediate*) environment (via language texts) but may moreover be understood as virtual in the sense that new (*mediate*) environments (via textual meanings) can also be processed, has been explicitly introduced elsewhere [41] [37] [38]. It is again dependent on a system's attunement to these kinds of discourse situations.

2.3 Discourse Situations

These situations (comprising *system*, *environment*, and *processing*) are considered *cognitive* inasmuch as the system's internal (formal and procedural) knowledge has to be applied to identify and to recognize structures external to the system (allowing *meaning interpretation*). These situations become *semiotic* whenever the internal knowledge applied to identify and interpret environmental structures is derived from former processes of structure identification and interpretation as the result of self-organizing feedback through different levels of (inter-)mediate representation and organization. This process (of *meaning* constitution or structure *understanding*) has its procedural analogue in the mul-

[3]i.e. with reference to *intrinsic* interdependencies within the *system-environment-processing* situation (see Peirce's *tri-relative influence* below).

tiple enactment of the threefold relation which is called—following PEIRCE—*semiosis*[4]. The triadic relation allows for the different ontological abstractions of *language*

▷ as a component (*sign*) in a system's external environment, i.e. material *discourse* as a physical space-time location;
▷ as a constituent of virtuality which systems properly attuned experience as their environment (*object*), i.e. structured *text* as an interpretable potential of meanings, and
▷ as a process of actualization (*interpretant*) in a particular system-environment situation, i.e. *understanding* as the constitution of meaning.

Under these preliminary abstractions, the distinction between (the formats of) the representation and (the properties of) the represented is not a prerequisite but an outcome of *semiosis*, i.e. the semiotic process of sign *constitution* and *understanding*. Hence, it should not be a presupposition or input to, but a result or output of the processes which are to be modeled procedurally and called *semiotic*.

2.4 Cognitive Models

An earlier attempt [46] to classify model constructions as forwarded in the cognitive sciences had distinguished three types of modeling approaches:

▷ the *cognitive* approach presupposes the existence of the external world, structured by given objects and properties, and the existence of representations of (fragments of) that world internal to the system, so that the cognitive systems' (observable) behavior of action and reaction may be modeled by processes operating on these structures;
▷ the *associative* approach is described as a dynamic structuring based on the model concept of self-organization which cognitive systems constantly apply to adapt to changing environmental conditions and to modify their internal representations of them;
▷ the *enactive* approach may be characterized as being based upon the notion of *structural coupling*. Instead of assuming an external world and the systems' internal representations of it, some unity of structural relatedness is considered to be fundamental of—and the (only) condition for—any abstracted or acquired duality in notions of the external and the internal, object and subject, reality and its experience.

Whereas the first two approaches apparently draw on the traditional rationalistic paradigm of mind-matter-duality—*static* the former, *dynamic* the latter—by assuming the existence of *external* world structures and *internal* representations of them, the third type does not. Considering the importance that the notions

[4]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. ([21], p.282)

of formatting and representation (both internal and external to an information processing system) have gained in tracing processes on the grounds of their observable or resulting structures, it appears to be justified to add the fourth type: [40]:

▷ the *semiotic* approaches focus on the notion of *semiosis* and may be characterized by the process of *enactment* too, complemented, however, by the representational impact. It is considered fundamental to the distinction of e.g. *cognitive processes* from their *structural results* which—due to the traces these processes leave behind—may emerge in some form of *knowledge* whose different representational modes comply with the distinction of *internal* or *tacit* knowledge (i.e. *memory*) on the one hand and of *external* or *declarative* knowledge (i.e. *language* expressions) on the other.

According to these types of cognitive modeling, *computational semiotics* can be characterized as aiming at the dynamics of meaning constitution by simulating processes of multi-resolutional representation [19] within the frame of an ecological information processing paradigm [39]. When we take human beings to be *systems* whose knowledge based *processing* of represented *information* makes them *cognitive*, and whose sign and symbol generation, manipulation, and understanding capabilities render them *semiotic*, we will do so due to our own daily experience of these systems' outstanding ability for representing results of cognitive processes, organize these representations, and modify them according to changing conditions and states of system-environment adaptedness.

3 Natural Language Processing

Computational systems for natural language processing which are based upon relevant CL and AI research are presently undergoing some fundamental scrutiny. It may broadly be characterized by challenges concerning some of the founding assumptions and basal hypotheses implied in the research *goals* (*Erkenntnisinteresse*), the critical evaluation of methodological standards and their possible completion by new research *methods* (*Untersuchungsmethoden*), and a re-definition of the linguistic domain of language research *objects* in general (*Forschungsgegenstände*).

3.1 Challenging Representational Formats

As is well known, computational systems for natural language analysis and generation are based upon correct structural descriptions of input strings and their semantic interpretations. This is made possible by rule based representations of (syntactic and lexical) knowledge of a language and of (referential and situative) world knowledge concerned in formats which grammar formalisms and deductive inferential mechanisms can operate on. Notwithstanding the considerable advances in the development and theoretical testing of increasingly more complex systems, this kind of *cognitive* (or *knowledge-based*) language processing (based on monotone logics, symbolic representations, rule-based operations, serial processing, etc.) and the essential statics of their representational structures were challenged—although for differing reasons—by connectionistic and empirical approaches. These were particularly successful in simulating dynamic properties of processes of cognitive natural language processing (based on the theory of dynamic systems, sub-symbolic or distributed representation, numerically continuous operations, parallel processing, etc.) in ANN models [20], [22], [7]. And there were new insights gained into the wealth of structural patterns and functional relations which could be observed in large corpora of communicative natural language performance and specified by results from models of quantitative and statistical analyses (based on probability and possibility theory, stochastic and fuzzy modeling, numerical mathematics and non-monotone logics, strict hypothesizing and rigorous testing, etc.) [1] [17].

3.2 Common Grounds

Discussing connectionistic vs. rule-based approaches and models of natural language processing [44] [13] [45] [42], differences have in the majority of contributions been characterized from an epistemological position common for both directions in *cognitive linguistics*. Proponents of both sides seem to accept that the study of language *competence*, its principles, components, and their organization is the primary concern and basal objective for computational linguistics proper. Following the discussion so far, there is a predominant interest in the theoretical aspects of what the different, even *hybrid* model constructions would claim and may justifiably be said to explain [22] [6] [16] although the empirical and quantitative approaches in language research have hardly been involved yet. The reasons are manifold as the availability of masses of performative language data not only require the methodological mastery of a whole spectrum of tools and methods, new to most linguists, but also tend to imply some compensating shift from language competence towards language performance studies corresponding to a wider domain of research objects for linguistics proper [12] which many computational linguists would refrain from. However, in view of the formal complexity and applicational limitations which rule based cognitive models show on the one hand, and considering the surprisingly efficient practicability of stochastic parsers [10] [3] and statistical machine translation systems [5] [4] on the other, there is good reason to expect some revision of assumptions and basal hypotheses defining cognitive linguistics.

3.3 Language Reality

According to Chomsky [8] the cognitive study of natural language phenomena has to be concerned primarily with the principles underlying observable language phenomena, i.e. the structure and the organization of the human language faculty (*competence*) which may (theoretically) be analyzed and (formally) be characterized well w i t h o u t empirical exploration of observable language data as produced in situations of communicative interaction by real speakers/hearers (*performance*). Nowadays [9] the speaker's language knowledge or *competence* is named *internalized* (or I-) language whose set of entities (lexicon) aggregatable according to a set of rules (computational system) constitutes the proper domain (mental grammar) of linguistic inquiry; accordingly, the speakers' *performative* language use named *externalized* (or E-) language may be considered cognitively uninteresting. One of the results which the ongoing discussion may produce is the understanding [43] that the grounding of cognitive linguistic research so far might turn out to be based on a too principled abstraction of language reality as experienced in communicative interaction.

Taking into account some language regularities and structures which are empirically traceable but may not be identified within the categorial framework of established linguistic concepts[5], and in view also of tendencies in cognitive linguistics, computational linguistics, and AI research to come up with increasingly complex systems and/or narrowing scopes dealing with natural language structures and functions [24], the empirical study of performative language phenomena may provide valuable insights and explanations because of the domain's new research objects and methods complementary to those of competence centered linguistics. Moreover, it appears that empirical approaches allow for quantitative-statistical as well as fuzzy-theoretical model constructions which may allow for a more *semiotic* understanding of the functioning of language signs as used by interlocutors in communicative interaction.

3.4 Research Situation

The availability (and still increasing number) of very large text corpora[6], will facilitate to investigate a type of natural language properties whose categorial

[5]Phenomena like linear short-distance orderings (*Nah-Ordnung*) of performative language entities (e.g. co-occurences) whose regularities are deprived of rule-based notations but can easily be represented and processed as numerical expressions of correlation values with any precision, are an example in point here, as they appear to be the observable results of structuring principles which have been overlooked by competence theoretical investigations only because they do not comply with linear long-distance orderings (*Fern-Ordnung*) that are constitutive of linguistic categories as represented and processed by familiar grammar formalisms[14], [11], [15].

[6]The Trier *dpa*-Corpus comprises the complete textual material from the *basic news real service* of 1990–1993 (720.000 documents) which the Deutschen Presseagentur (*dpa*), Hamburg, deserves thanks to have the author provided with for research purposes. After cleaning of editing commands the *dpa*-Corpus consists of approx. 180 million ($18 \cdot 10^7$) running words (*tokens*) for which an automatic tagging and lemmatizing tool is under development. It is this corpus which provides the performative data of written language use for the current (and planned) *fuzzy*-linguistic projects at our department.

vagueness (uncertain, under-determined, fuzzily delimited, etc.) or whose limited to dubious observability (sparse data, uncertain information, etc.) had left them inaccessible and hence irrelevant to language research.

As the processing of very large language corpora (VLLC) has shown, categorial type concepts common in traditional linguistics have to be considered highly problematic when applied to classify finer grained structures which quantitative-numerical computations will easily identify operationally. Categorial type, symbol processing encounters increasing numbers of borderline cases, variations, and ambiguities which cannot be dealt with consistently. Such problems ought to—and can infact—be avoided from the very start as they emerge from mappings of structurally related data sets to inadequate categories. Therefore, classical categorial conceptions in linguistics have begun to be scrutinized and may possibly be substituted by *soft categories* [38] [27] before there is substantial hope to improve chances to understand and to explain *knowledge* as some form of (*world*) and/or *language*) structures emerging from rather than fed into information processing models that can truly be called *semiotic*.

3.5 SCIP Systems

Other than value attributing procedures that reorganize input data computationally according to predefined structures of intermediate representations (as hypothesized by *competence* theoretical linguistics and realized in cognitive CL models) *semiotic cognitive information processing* (SCIP) systems [37] may have to, and will indeed, be distinguished sharply as sets of procedures whose computations will transform structured input data according to its immanent regularities to yield new, structural representations emerging from that computation (as hypothesized by *performative* linguistics and realized in procedural models of computational semiotics). The elements of these new structures are value distributions or vectors of input entities that depict properties of their structural relatedness in a *granular* and multi-layered fashion, constituting multi-dimensional (metric) space structures (*semiotic spaces*). Their elements may also be interpreted as *fuzzy sets* allowing set theoretical operations be executed on these representations that do not require categorial type (*crisp*) definitions of symbol or concept formation. Computation of letter (*morphic*) vectors in *word space*, derived from n-grams of letters *graphemes*, as well as of word (*semic*) vectors in *semantic space*, derived from correlations of words [29], [25], [36] will serve to illustrate the operational flexibility and varying granularity of vector representations to identify regularities of language performance which traditional linguistic categories fail to represent.

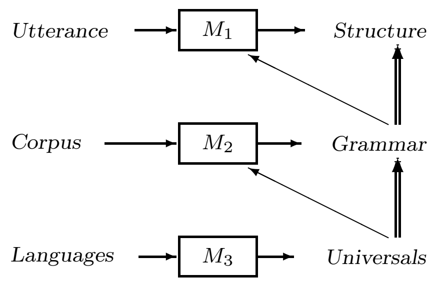

Figure 1

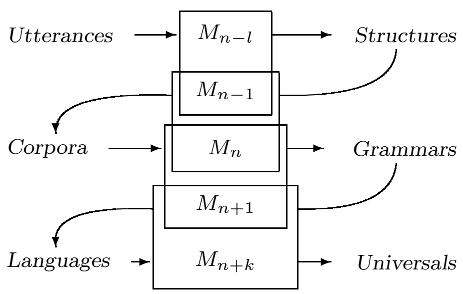

Figure 2

Schemata of model hierarchy of *cognitive linguistic* strata of mechanisms (Bierwisch) as compared to model tiling of *computational semiotic* coverage of procedures (Rieger) for the analysis and representation of (abstracted and observable) language phenomena.

3.6 Visualizing Vector Representations

Returning to the ecological systems theoretical view applied to information processing, we will focus on the problem of visualizing results of computational procedures developed to model and simulate semiotic processes whose numerical representations—by definition—do not have an immediate interpretation. Various techniques have been applied to analyze, scrutinize, and visualize the structuredness of vectoral representations and their results reported elsewhere [30] [31] [32] [33] [34]. As these have been able to demonstrate the definite non-contingency of *meaning points* z in the *semantic space*, a short introduction to its conception will suffice here. We will concentrate on the level of *semantic* meaning constitution as based upon the measurement of differences of usage regularities in VLLC of situated or *pragmatically homogeneous* texts.

For a vocabulary $V = \{x_n\}, n = 1, \ldots, i.j. \ldots, N$ of lexical items, their meanings $z_n \in \langle S, \zeta \rangle$ are re-constructed as a composite function $\tilde{\delta} \mid y_n \circ \tilde{\alpha} \mid x_n$ of the difference distributions

$$\tilde{\delta} \mid y_n : C \rightarrow S; \{z_n\} =: S \tag{1}$$

and the grounding usage regularity distributions

$$\tilde{\alpha} \mid x_n : V \rightarrow C; \{y_n\} =: C \tag{2}$$

The empirical measures employed to specify intensities of co-occurring lexical items are centered around a modified correlational coefficient

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

where $e_{it} = \frac{H_i}{L} l_t$ and $e_{jt} = \frac{H_j}{L} l_t$ denote (theoretical) estimate values, computed over a corpus $K = \{k_t\}; t = 1, \ldots, T$ of texts whose lengths summed up will define the overall length of the corpus $L = \sum_{t=1}^{T} l_t; 1 \leq l_t \leq L$ The length l_t of each text t in the corpus K is measured by the number of occurring word-tokens which form the basis of the vocabulary V of word-types whose frequencies are denoted by $H_i = \sum_{t=1}^{T} h_{it}; 0 \leq h_{it} \leq H_i$.
A second measure of similarity (or rather, dissimilarity) is applied to specify the α-value distributions' differences

$$\delta(y_i, y_j) = \left(\sum_{n=1}^{N}(\alpha(x_i, x_n) - \alpha(x_j, x_n))^2\right)^{\frac{1}{2}}; \tag{4}$$
$$0 \leq \delta(y_i, y_j) \leq 2\sqrt{n}$$

The consecutive application of (*Eqns.* 2) on input texts and (*Eqns.* 1) on its output data allows to derive for each word-type $x_n \in V$ an entity $z_n \in S$ which denotes a structural representation of differences of usage regularities detected and numerically specified by these coefficients. They allow to model the meanings of words as a two-level function of δ-values of *paradigmatic* selections (measured as differences of usage regularities) and of α-values of *syntagmatic* aggregations (computed as correlations of word-type pairs in texts of the corpus), as schematized in *Tab.* 1.

As a result of this two-stage *consecutive* mapping any meaning point's position in the *semantic space* $\langle S, \zeta \rangle$ is determined by all the differences (δ- or distance-values) of all regularities of usage (α- or correlation-values) each lexical item shows against all others in the discourse analyzed. Without recurring to any investigator's or his test-persons' word or world knowledge (*semantic*

$$\begin{array}{ccccc}
\alpha = V \times V & \alpha\text{-abstraction} & \delta = C \times C & \delta\text{-abstraction} & \zeta = S \times S \\
\begin{array}{c|c} \tilde{\alpha} & x_1 \ldots x_N \\ \hline x_1 & \alpha_{11} \ldots \alpha_{1N} \\ \vdots & \vdots \ddots \vdots \\ x_N & \alpha_{N1} \ldots \alpha_{NN} \end{array} & \begin{array}{c} \Downarrow \\ \xrightarrow{\tilde{\alpha} \mid x_i} \\ \Uparrow \\ \textit{Syntagmatic} \end{array} & \begin{array}{c|c} \tilde{\delta} & y_1 \ldots y_N \\ \hline y_1 & \delta_{11} \ldots \delta_{1N} \\ \vdots & \vdots \ddots \vdots \\ y_N & \delta_{N1} \ldots \delta_{NN} \end{array} & \begin{array}{c} \Downarrow \\ \xrightarrow{\tilde{\delta} \mid y_j} \\ \Uparrow \\ \textit{Paradigmatic} \end{array} & \begin{array}{c|c} \zeta & z_1 \ldots z_N \\ \hline z_1 & \zeta_{11} \ldots \zeta_{1N} \\ \vdots & \vdots \ddots \vdots \\ z_N & \zeta_{N1} \ldots \zeta_{NN} \end{array} \\
 & & \textit{Constraints} & &
\end{array}$$

Table 1: Formalizing (*syntagmatic/paradigmatic*) constraints by consecutive (α and δ) abstractions over usage regularities of items x_i, y_j respectively.

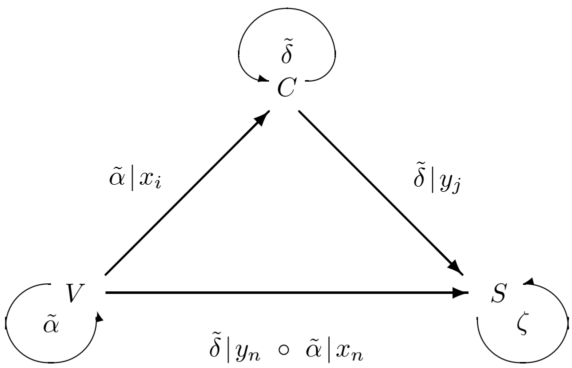

Figure 3: Fuzzy mapping relations $\tilde{\alpha}$ and $\tilde{\delta}$ between the structured sets of vocabulary items $x_n \in V$, of corpus points $y_n \in C$, and of meaning points $z_n \in S$.

competence), but solely on the basis of usage regularities of lexical items in discourse (*communicative performance*), some natural language understanding is modeled procedurally by computational processes which construct and identify emergent topological positions of any meaning point $z_i \in \langle S, \zeta \rangle$ corresponding to the vocabulary item's $x_i \in V$ meaning representation.

Thus, the application of *Eqns.* 2 and *Eqns.* 1 is tantamount to a double (α- and δ-) abstraction of identical elements in the ordered pairs as computed by the α-coefficient *Eqn.* 3 and the δ-coefficient *Eqn.* 4 respectively. This mapping may be considered a *semiotic* one due to the different ontological status which the α- and δ-representations of identical labels x_n acquire via this mapping process. Therefore, this *semiotic mapping* can formally also be stated as a fuzzy set theoretical *composition* of two restricted (fuzzy) relations $\tilde{\delta} \mid y$ and $\tilde{\alpha} \mid x$, or as a category of *semiotic morphisms* as in *Fig.* 3.

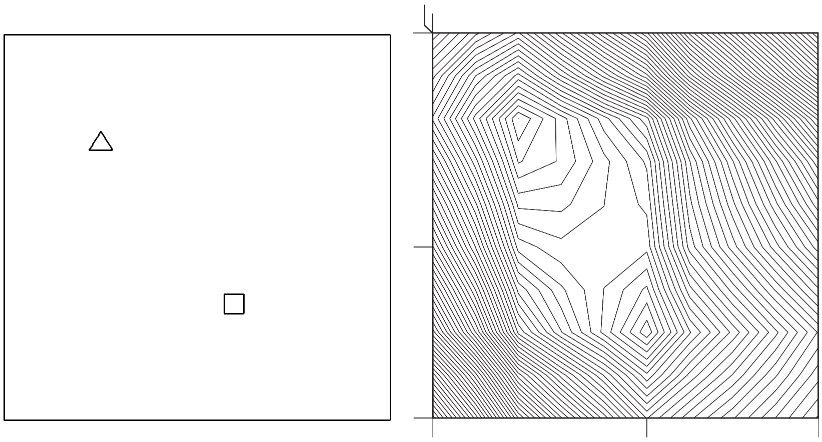

Figure 4: Reference plane with location of objects (△ and □) propositionally described by texts in the training corpus and 2-dim-image of SCIP system's *endo*-view showing regions of potential object locations (*isoreferentials*) under crisp hedge interpretation.

3.7 Reconstructing Reference

The *semantic space* structure $\langle S, \zeta \rangle$ may be viewed as the information processing system's internal representation (IR) of its external reality (ER). To be more precise: as an IR it depicts some of those ER structures which (firstly) are presented in the natural language discourse as its *meaning* a n d which (secondly) the system is able to detect according to its own structuredness and processing capabilities, constituting this *meaning* without knowing the i*semantics* underlying it. In order to let the internal (*endo*) picture which the system computes from discourse and represents in the form of the *semantic space* be evaluated against the external (*exo*) reality, this reality has to be mediated to the system as its environment in an intersubjectively controlled way a n d via those very propositional language structures whose non-propositional understanding by the system is to be tested. To facilitate this, the system is placed in an experimental environment which both are heavily restricted compared to our real world. The *immediate* environment consist solely of a text corpus of situated natural language discourse—aggregated from correct sentence expressions of true propositions—describing as *mediate* environment fixed object locations in a plane from changing system positions. The sentences were generated according to a (very simple) phrase structure grammar and a formal fuzzy referential semantics both unknown to the system. These formalisms (grammar and semantics) allowed to specify and interpret composite predicates of *cores* (like: *on the* *left, on the right* — *in front, behind*) and *hedges* (like: *extremely, very, rather* — *nearby, faraway*) in a consistent way as employed in sentences automatically generated to describe the object location and system position relations as the system's/environment's *structural coupling* in an intersubjectively controlled way. Submitted to the system's non-propositional, non-symbolic, numerical language processing capabilities, the generated corpus would reveal hidden structural informational constraints which the system's own structuredness (*attunement*) and internal processing or meaning constitution (*understanding*) would have to reflect in its representational structure (*knowledge*). The experimental setting and implemented tests which where reported in detail elsewhere [37] [38] [39] allowed to compare the external (*exo*) reality (*Fig.* 4 left)—as described by the texts and formally specified by the underlying syntax and semantics—with its two-dimensional transform of the system's internal multi-dimensional (*endo*) view of its discourse environment, demonstrating quite convincingly the computed structure's (at least partial) adequacy (*Fig.* 4 right).

Considering the structural properties, locational preferences, and adjacency relations of objects in the system's environment, the tests have produced under varying fuzzy interpretations of hedges very promising results which illustrate the *SCIP* system's miniature (cognitive) language understanding and meaning acquisition capacity w i t h o u t having any syntactic and/or semantic knowledge in whatever format made available to it prior to processing [39]. This is a case in point showing that the formats of syntax and semantics as employed in traditional linguistic and logic analysis and representation of natural language expressions is clearly one [but not the only) *façon de parler* or rather, way of submitting language structure to a representational framework which models prevailing constituents and dependencies.

3.8 Dispositional Dependencies

Following the semiotic understanding of meaning more as a constitutional process rather than an entity of invariable constancy or static representation, the present *semantic space* may be considered part of a word meaning/world knowledge representation system of a new kind. It is characterized by its two-stage representational process which separates the format of basic (stereotyped) meaning components (*meaning points*) from their latent (dependency) relational organization as meaning potential (*semantic dispositions*). Whereas the former is a static, topologically organized multi-dimensional memory structure, the latter can be characterized as a dynamic and flexible structuring process which reorganizes and thereby transforms the basic relatedness of the elements it operates on.

This is achieved by a recursively defined procedure that produces hierarchies of meaning points, reorganized under given aspects according to and in dependence of their co- and contextual relevancy. Taking up ideas from cognitive

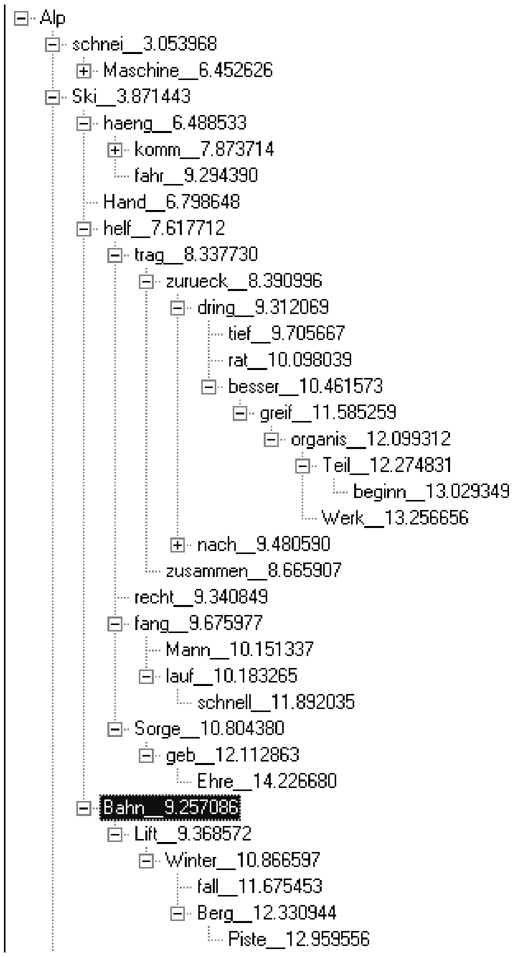

Figure 5: Fragment of DDS-tree of `Alpen` (root) as generated from *semantic space* data ($V = 345$ types, $H_i \geq 10$) of a German newspaper sample (Die Welt, 1964 Berlin edition).

theories of semantic memory, *priming*, and *spreading activation* [18], the *DDS*-algorithm was devised to operate on the semantic space data and to generate *dispositional dependency structures* (DDS) in the format of n-ary trees. Given one meaning point's position, the algorithm will

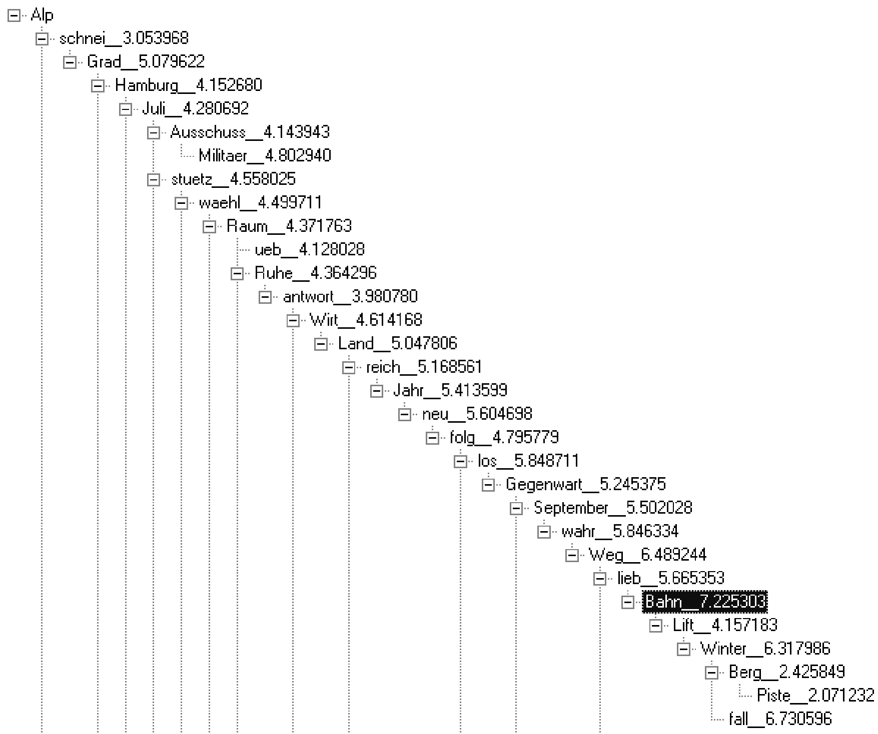

Figure 6: Fragment of MST-graph of `Alpen` (root) as generated from the same *semantic space* data.

1. take that meaning point's label as a start,
2. stack labels of all its neighboring points by decreasing distances,
3. open DDS-tree with starting point's label as primed head or root node. Then it will
4. take label on top of stack as daughter node,
   4.1 list labels of all its neighbors,
   4.2 intersect it with nodes in tree,
   4.3 determine from intersection the least distant one as mother node,
5. link it as daughter to identified mother node
6. and repeat 4. either
   6.1 until 2. is empty
   6.2 or other stop condition (given number of nodes, maximum distance, etc.) is reached
7. to end.

The tree structured graphs[7] may serve as a visualization of the dependencies that any labeled meaning point $z_n \in \langle S, \zeta \rangle$ chosen as root node will produce ac-

[7]The figures present subtrees of a semantic space which was computed from a sample of

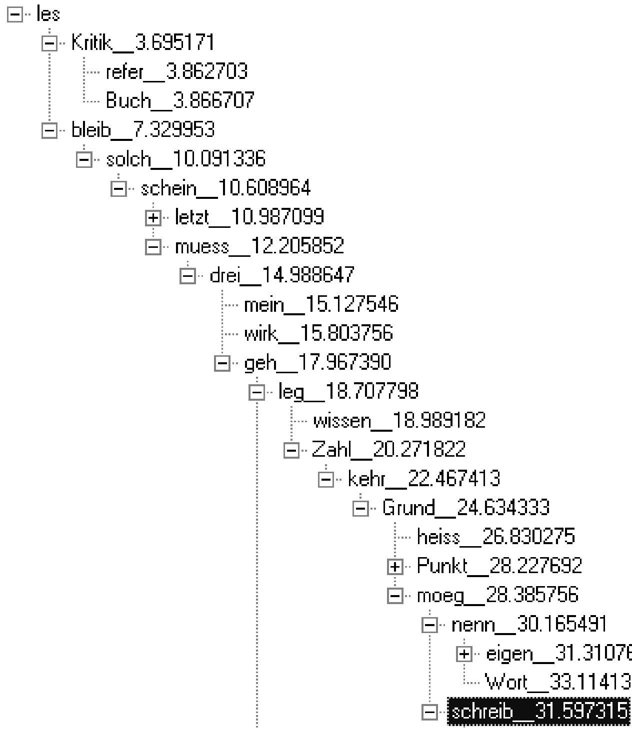

Figure 7: *Dependency path* of `lesen/`*to read* $\Longrightarrow$ `schreiben/`*to write* as traced in DDS-tree of `les`.

cording to the adjacencies of other points in the *semantic space* (*Fig.* 5). Their positions—being determined by and reconstructed operationally from the differences of usage regularities of word distributions in the texts analyzed—will thus guarantee that semantically related meanings will also be related in that tree structure, and that the direction of that relation (*dependency*) will vary contextually according to the semantic *aspect* (or starting node chosen) under which the system of structurally derived *meaning* representations is algorithmically reorganized. The tree has been named *dispositional* because of the potentiality of possible meaning relations and dependencies which it represents.

In order to illustrate the contextual sensitivity which distinguishes the DDS-algorithm from e.g. *minimal spanning trees* (MST) [23], the latter (*Fig.* 6) has been generated from the same data with the same starting node. Note, that the `Bahn`-node (*track, course, trail*) with identical subtrees is to be found on

---

texts from the German daily newspaper (Die Welt, 1964, Berlin edition); ⊕ marked nodes hide subtrees not expanded; the numerical values stated are direct ζ-distances from the root node.

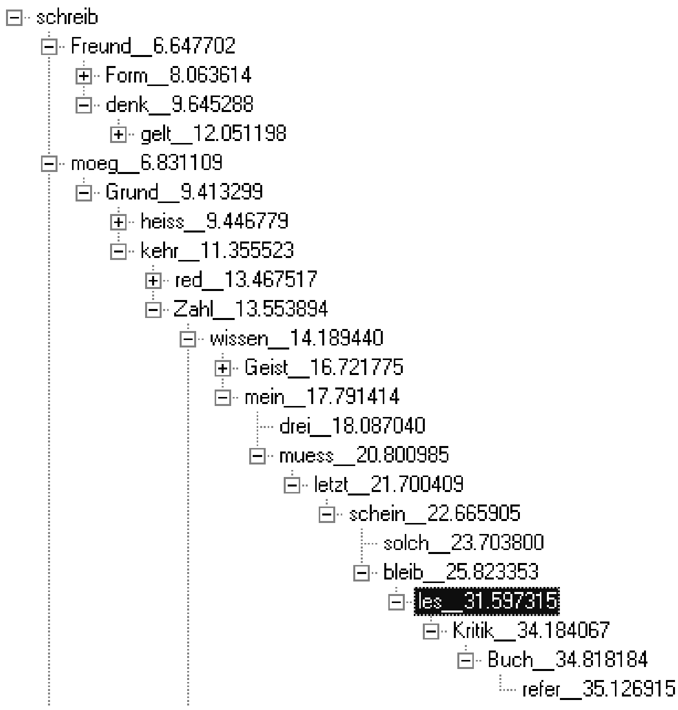

Figure 8: *Dependency path* of `schreiben`/*to write* $\Longrightarrow$ `lesen`/*to read* as traced in DDS-tree of `schreib`.

extremely different levels comparing DDS (level 3) and MST (level 23)[8].

Apparently, although the DDS-algorithm can simply be characterized as an encapsulated MST-procedure, this encapsulation serves a meaning constituting purpose. Where the MST is searching for shortest possible distance relations between points qualifying for tree node relatedness, the DDS is looking for highest *meaning similarities*, i.e. for shortest possible distance relations between points which are interpretable as *semiotically* derived representations. It is this property that allows the algorithm's search space to be *semantically* constrained as the starting point's or root node's topological environment (capsule), rendering it *aspect*-dependent and structurally *context sensitive*.

This has a number of consequences of which the following seem interesting enough to be listed:

- ▷ The procedural (semiotic) approach replaces the storage of fixed and ready set relations of (semantic) networks in AI by source- or aspect-oriented induction of relations among meaning points by the DDS algorithm;
- ▷ DDSs dependencies may be identified with an algorithmically induced *rele-*

[8] The numerical MST values given are direct ζ-distances between (mother and daughter) nodes.

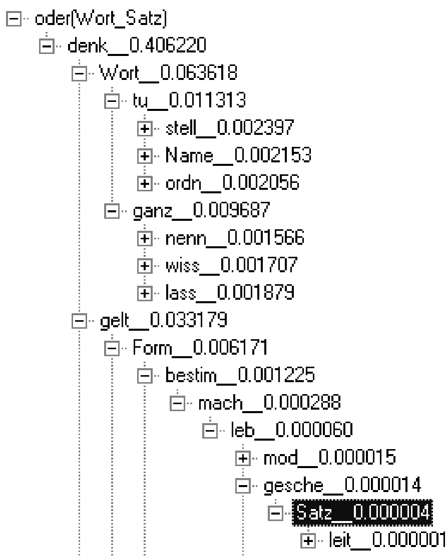

Figure 9: Fragment of DDS of `Wort`/*word* ∨ `Satz`/*sentence* (with *criteriality*-values) as generated from the new meaning point derived by the *fuzzy* `OR` operation.

vance relation which is reflexive, non-symmetric, and (weakly) transitive as illustrated by the *dependency paths*' listings of node transitions `les` (*to read*) $\Longrightarrow$ `schreib` (*to write*) and its (partial) inverse `schreib` $\Longrightarrow$ `les` (*Figs.* 7 and 8);

▷ the relevance relation gives rise to the notion of *criteriality* which allows to specify to what degree a meaning compound contributes to the *meaning potential* a root node's DDS is to represent. It may numerically be specified as a function of any node's level and ζ-distance by

$$Cr_i(d)_{\kappa+1} = Cr_i(m)_\kappa \cdot e^{-\frac{\zeta(d,i)}{\lambda+\zeta(d,m)}} \tag{5}$$

with i, m, d for root, mother, and daughter nodes respectively, and the counters κ for (left to right) nodes, and λ for (top down) levels in the tree;

▷ as the criteriality values are decreasing monotonously from 1.0 (root) they may be interpreted as membership values which reflect the relevance related *soft* structure of components (nodes) in the DDS as *fuzzy meaning potential*. Applying the fuzzy set theoretical extensions for logical operators (*and, or, non,* etc.) opens new possibilities to generate composite meaning points (`Wort`/*word* ∧ `Satz`/*sentence*) and `Wort`/*word* ∨ `Satz`/*sentence*)) without assuming a propositional structure a n d to get these composites' structural *meanings* determined by their DDSs as computed from the semantic space data (*Figs.* 9 and 10)[9];

[9]The numerical values given here are Cr-criterialities of daughter nodes as defined by *Eqn.*

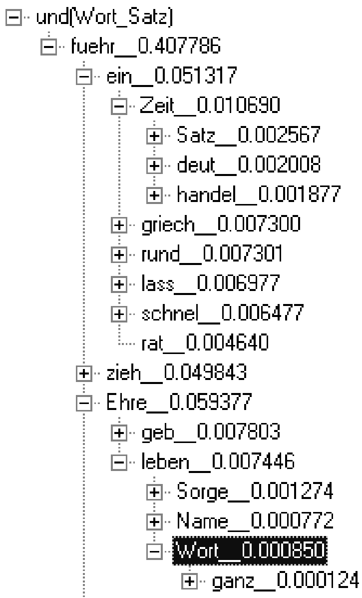

Figure 10: Fragment of DDS of `Wort`/*word* ∧ `Satz`/*sentence* (with *criteriality*-values) as generated from the new meaning point derived by the *fuzzy* `AND` operation.

▷ experiments are underway to employ DDSs as structural frame for semantic inferencing without the need to have the premises be stated in a predicative or propositional form prior to the concluding process. The DDS algorithm lends itself easily to the modeling of *analogical* reasoning processes by a procedure which takes two (or more) root nodes (as semantic premises), initiates two (or more) DDS processes each of which—in selecting their respective daughter nodes—will tag the corresponding meaning points in the semantic space. Stop condition for this process which may proceed *highest criteriality breadth first* through the respective DDSs could be the first meaning point found to be tagged when met by either (or any) of the processes active. This point would be considered the (first) candidate to be semantically inferred or concluded from the premises (with the option to extend the number candidates).

4 Conclusion

It is hoped that devising representational structures which result from procedures of systematic exploration of *syntagmatic* and *paradigmatic* constraints on different levels of natural language discourse will allow to come up some day with a new understanding of how entities and structures are formed which are

5.

semiotic, i.e. do not only emerge from processes as their results which have an objective (material) extension in space-time, but can above that and due to their (recursively defined) co- and context dependency be understood as having interpretable meaning.

In order to be able to interpret, we need to have structures, but we are about to experience that the model structures available so far do not serve the purpose we are looking for. When we have to deal with problems which might result from the lack of concepts, of structures, and of formats to describe or represent them adequately, we should not be too surprised to find these problems unsolvable. Procedural models and their computational enactment generating structures sensitive to situational embeddings appear to be good candidates for progress.

References

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

[2] J. Barwise/ J. Perry: *Situations and Attitudes.* MIT Press, Cambridge, MA, 1983.

[3] T. Briscoe/ J. Carroll: Generalized Probabilstic LR Parsing of Natural Language (Corpora) with Unification-Based Grammars. *Computational Linguistics*, 1993(1):25–59, 1993.

[4] P. Brown/ S. DellaPietra/ V. DellaPietra/ R. Mercer: The Mathematics of Statistical Machine Translation: Parameter Eestimation. *Computational Linguistics*, 1993(2):263–311.

[5] P. E. Brown: A statistical approach to language translation. In: D. Vargha (ed): *COLING-88 Budapest*, ICCL, J.v.Neumann Society of Computing Science. Budapest 1988, pp. 71–76.

[6] H. Bunke (ed): *Advances in Structural and Syntactic Pattern Recognition.* World Scientific, Singapore, 1992.

[7] G. A. Carpenter/ S. Grossberg (eds): *Pattern Recognition by Self-Organizing Neural Networks.* MIT Press, Cambridge, MA, 1991.

[8] N. Chomsky: *Language and Problems of Knowledge. The Managua Lectures.* (Current Studies in Linguistics 16). MIT Press, Cambridge, MA, 1988.

[9] N. Chomsky: *The Minimalist Program.* (Current Studies in Linguistics 28). MIT Press, Cambridge, MA, 1995.

[10] K. Church: A stochastic parts program and noun phrase parser for unrestricted text. In: *Second Conference on Applied Natural Language Processing*, ACL, Austin, TX, 1988, pp. 136–143.

[11] K. W. Church/ P. Hanks: Word association norms, mutual information, and lexicography. *Computational Linguistics*, 1990(1):22–29.

[12] K. W. Church/ R. L. Mercer: Introduction to the Special Issue on Computational Linguistics Using Large Corpora. *Computational Linguistics*, 1993(1):1–24.

[13] J. Fodor/ Z. Pylyshyn: Connectionism and cognitive architecture: A critical analysis. *Cognition*, 1988(28):3–71.

[14] W. Fucks: On Nahordnung and Fernordnung in samples of literary texts. *Biometrica*, 1954(2):116–132.

[15] D. Hindle/ M. Rooth: Structural Ambiguity and Lexical Relations. *Computational Linguistics*, 1993(1):103–120.

[16] G. Hinton (ed): *Connectionist Symbol Processing.* MIT Press, Cambridge, MA, 1992.

[17] R. Köhler/ B. Rieger (eds): *Contributions to Quantitative Linguistics. Proceedings 1st Quantitative Linguistics Conference – QUALICO-91.* Kluwer Academic Publishers, Dordrecht, 1993.

[18] R. Lorch: Priming and search processes in semantic memory: a test of three models of spreading activation. *Journal of Verbal Learning and Verbal Behavior*, 1982(21):468–492.

[19] A. Meystel: *Semiotic Modeling and Situation Analysis: an Introduction.* AdRem Inc, Bala Cynwyd, PA, 1995.

[20] L. Nadel/ L. Cooper/ P. Culicover/ R. Harnish (eds): *Neural Connections, Mental Computation.* MIT Press, Cambridge, MA, 1989.

[21] C. S. Peirce: Pragmatism in Retrospect: A Last Formulation. In: J. Buchler (ed): *The Philosophical Writings of Peirce*, Dover, New York, 1906, pp. 269–289.

[22] R. Pfeiffer/ Z. Schreter/ F. Fogelman-Soulié/ L. Steels (eds): *Connectionism in Perspective.* North Holland, Amsterdam, 1989.

[23] C. Prim: Shortest connection networks and some generalizations. *Bell Systems Technical Journal*, 1957(36):1389–1401.

[24] B. Rieger: Theoretische Grundlagen der Computerlinguistik: eine Paneldiskussion mit M. Bierwisch, C. Habel, B. Rieger, H. Uszkoreit und W. Wahlster. In: I. Bátori/ U. Hahn/ M. Pinkal/ W. Wahlster (eds): *Computerlinguistik und ihre theoretischen Grundlagen*, (KI–Informatik-Fachberichte Bd.195). Springer, Berlin/ Heidelberg/ NewYork, 1988, pp. 192–218.

[25] B. Rieger: Unscharfe Semantik: zur numerischen Modellierung vager Bedeutungen von Wörtern als fuzzy Mengen. In: H. Friemel/ G. Müller-Schönberger/ A. Schütt (eds): *Forum-90: Wissenschaft und Technik. Neue Anwendungen mit Hilfe aktueller Computertechnologien*, (Informatik-Fachberichte Bd. 259). Springer, Berlin/ Heidelberg/ NewYork, 1990, pp. 80–104.

[26] B. Rieger: Fuzzy Modellierung linguistischer Kategorien. In: W. Feldmann/ H. Hinrichs (eds): *Text und Lexikon*, (Lexicographica Series Maior 73). Niemeyer, Tübingen, 1996, pp. 155–169.

[27] B. Rieger: Warum Fuzzy Linguistik? Überlegungen und Ansätze einer computerlinguistischen Neuorientierung. In: D. Krallmann/ W. Schmitz (eds): *Internationales Gerold-Ungeheuer-Symposium*, Nodus Verlag, Münster, 1998, pp. 153–18.

[28] B. B. Rieger: Feasible Fuzzy Semantics. On some problems of how to handle word meaning empirically. In: H. Eikmeyer/ H. Rieser (eds): *Words, Worlds, and Contexts. New Approaches in Word Semantics*, (Research in Text Theory 6). de Gruyter, Berlin/ NewYork, 1981, pp. 193–209.

[29] B. B. Rieger: Fuzzy Representation Systems in Linguistic Semantics. In: R. Trappl/ N. Findler/ W. Horn (eds): *Progress in Cybernetics and Systems Research*, Vol. XI. McGraw-Hill Intern., Washington/NewYork/London, 1982, pp. 249–256

[30] B. B. Rieger: Clusters in Semantic Space. In: L. Delatte (ed): *Actes du Congrès International Informatique et Science Humaines*. Laboratoire d'Analyse Statistique des Langues Anciennes, LASLA, Liège, 1983, pp. 805–814.

[31] B. B. Rieger: Generating Dependency Structures of Fuzzy Word Meanings in Semantic Space. In: S. Hattori/ K. Iounu /eds): *Proceedings of the XIIIth International Congress of Linguists*. Comité International Permanent des Linguistes, CIPL.Tokyo, 1983, pp. 543–548.

[32] B. B. Rieger: Semantic Relevance and Aspect Dependancy in a Given Subject Domain. In: D. Walker (ed): *COLING-84 10th International Conference on Computational Linguistics*, International Committee of Computational Linguistics, ACL-ICCL. Stanford, 1984, pp. 298–301.

[33] B. B. Rieger: Inducing a Relevance Relation in a Distance-like Data Structure of Fuzzy Word Meaning Representations. In: R. Allen (ed): *Data Bases in the Humanities and Social Sciences. 4th International Conference on Data Bases in the Humanities and Social Sciences*, Paradigm Press, Osprey, FL, 1985, pp. 374–386.

[34] B. B. Rieger: Lexical Relevance and Semantic Disposition. On stereotype word meaning representation in procedural semantics. In: G. Hoppen-

brouwes/ P. Seuren/ T. Weijters (eds): *Meaning and the Lexicon.* Foris Publications, Dordrecht, 1985, pp. 387–400.

[35] B. B. Rieger: Distributed Semantic Representation of Word Meanings. In: J. D. Becker/ I. Eisele/ F. W. Mündemann (eds): *Parallelism, Learning, Evolution. Evolutionary Models and Strategies, WOPPLOT-89*, (Lecture Notes in Artificial Intelligence 565). Springer, Berlin/ Heidelberg/ New York, 1991, pp. 243–273.

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

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

[38] B. B. Rieger: Situations, Language Games, and SCIPS. Modeling semiotic cognitive information processing systems. In: A. Meystel/ N. Nerode (eds): *Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems*, 10th IEEE Symposium on Intelligent Control. AdRem, Bala Cynwyd, PA, 1995, pp. 130–138.

[39] B. B. Rieger: Situation Semantics and Computational Linguistics: towards Informational Ecology. A semiotic perspective for cognitive information processing systems. In: K. Kornwachs/ K. Jacoby (eds): *Information. New Questions to a Multidisciplinary Concept.* Akademie Verlag, Berlin, 1996, pp. 285–315.

[40] B. B. Rieger: Computational Semiotics and Fuzzy Linguistics. On meaning constitution and soft categories. In: A. Meystel (ed): *A Learning Perspective: Proceedings of the 1997 International Conference on Intelligent Systems and Semiotics (ISAS-97)*, (NIST Special Publ. 918). US Gov. Printing Office, Washington, DC, 1997, pp. 541–551.

[41] B. B. Rieger/ C. Thiopoulos: 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.* Kluwer Academic Publishers, Dordrecht, 1993, pp. 67–78.

[42] H. Schnelle: Die Herausforderung des Konnektionismus in der kognitiven Linguistik. In: H. Friemel/ G. Müller-Schönberger/ A. Schütt (eds): *Forum-90: Wissenschaft und Technik*, (Informatik-Fachberichte 259). Springer, Berlin/ Heidelberg/ NewYork, 1990, pp. 42–58.

[43] H. Schnelle: Welcher Sprachwissenschaft auf der Spur? Plädoyer für größere Klarheit und Toleranz. *Zeitschr.f.Sprachwissenschaft*, 1994(1):110–120.

[44] P. Smolensky: On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 1988(11):1–74.

[45] P. Smolensky: Connectionism and Constituent Structure. In: R. E. Pfeiffer (ed): *Connectionism in Perspective.* North Holland & Elsevier, Amsterdam/ NewYork, 1989, pp. 3–24.

[46] F. Varela/ E. Thompson/ E. Rosch: *The Embodied Mind. Cognitive Science and Human Experience.* MIT Press, Cambridge, MA, 1991.