Natural Grammar

Janos J. Sarbo

Radboud University
Faculty of Science, Dept. Computer Science
Toernooiveld 1
6525 ED Nijmegen
The Netherlands
janos@cs.ru.nl
tel. +31(0)24 3653049
fax +31(0)24 3653001

Jozsef I. Farkas

Radboud University
Faculty of Science, Dept. Computer Science
Toernooiveld 1
6525 ED Nijmegen,
The Netherlands
jfarkas@sci.ru.nl
tel. +31(0)24 3653049

Auke J.J. van Breemen

Van Breemen Onderwijs Advies,
Oude Graafseweg 52
6543 PS Nijmegen
The Netherlands
abreemen@semiosis.net
tel. +31(0)24 3786910

NATURAL GRAMMAR

Abstract

By taking as a starting point for our research the function of language to generate meaning, we endeavor in this chapter to derive a grammar of natural language from the more general Peircean theory of cognition. After a short analysis of cognitive activity, we introduce a model for sign (re)cognition and analyze it from a logical and semiotic perspective. Next, the model is instantiated for language signs from a syntactical point of view. The proposed representation is called natural insofar as it respects the steps the brain/mind is going through when it is engaged in cognitive processing. A promise of this approach lies in its potential for generating information by the computer, which the human user may directly recognize as knowledge in a natural, hence economic way.

INTRODUCTION

What is 'natural' in natural language? The answer to this question evidently depends on the interpretation of the word 'natural'. We interpret this word as expressing that the features we are looking for *represent* steps the mind or brain is going through when it is engaged in natural language use.[1] This is not to say that the representation stands for the process in all its detail or that it describes the actual operations the brain is going through, but only that whenever such a process occurs, that process can be represented by the steps discerned in our model. Since this interpretation of naturalness still leaves room for different routes of investigation we further narrow down our question to: Is it possible to develop a grammar that captures this naturalness formally?

The promise of this chapter is that indeed it is possible to develop a natural grammar. Additionally, our research has revealed that such a grammar is not restricted to natural language, but can be given a naive logical and a semiotic interpretation as well. The fact that a grammatical, a naive logical and a semiotic interpretation is possible indicates that the corresponding different domains of knowledge can have a uniform representation, a feat that makes our system very economic.

We maintain that an answer to 'what is natural in language' can be found if the function of language as a representation of meaning is the starting point of research. According to this view language can be seen as a kind of cognitive processing of signs. The word 'sign' is here used in the broad Peircean sense according to which anything which conveys any definite notion of an object is a sign (cf. CP 1.540).[2] This embeds our original question in a more general one: Can cognition be modeled formally as a sign recognition process? Clearly, if such a general model can be made, then, by restricting it to linguistic signs, which are

symbols, we may obtain a natural model of language, from which an underlying grammar can be derived easily. Since the proposed representation offers a basically computational account, while our meaningful sign use is not confined to computation, this chapter also investigates the complex relationship between computation and meaning.

Defining a computational, though natural representation can be difficult, as is illustrated by traditional language modeling. Humans process natural language in 'real-time', which is formally equivalent to linear complexity (Paul, 1984). This is opposed to the models of the traditional formal approach which are typically of higher complexity. A potential side-effect of relying on rules dictated by a formal ontology, instead of relying on the properties of a 'natural' process, can be the limited expressive power of such rules for a systematic specification of complex linguistic phenomena characterizing actual language use. We think the natural model of language, introduced in this chapter, does not suffer from such a limited expressive power and is on top of this efficient. While we can formally prove the last claim, the first one remains a conjecture, in line with our assumption concerning the dynamic nature of language. Some confidence in that conjecture, nevertheless, may be provided by the close relationship between our model and Peirce's theory of signs and with the conclusion of Peircean semiotic,[3] which roughly comes down to the statement that 'in order to be knowable something has to be of the nature of a sign'.

This conclusion about the all pervasive character of signs may shed some light on natural language and its conception as a (formal) grammar. For, by knowing the properties of the process underlying cognition, we may be able to answer the question how 'knowledge', as a cognitive process, could be defined *naturally* by means of (formal) rules.

Natural Grammar, as a formal grammar, bears similarity with the dependency based formalisms of cognitive linguistics like Word Grammar (Hudson, 1984). Word Grammar is a branch of Dependency Grammar in that it uses word-word dependencies as a determinant for linguistic structure. As a result it presents language as a network of knowledge that links concepts about words, like their meaning (e.g. grammatical function) to the form, the word class, etc. Such grammars do not rely on phrasal categories. Contrary to cognitive linguistics, which aims at incorporating the conceptual categories of language in rules which are dictated by a formal theory, the rules of Natural Grammar are derived on the basis of an analysis of the properties of cognition and the processing of signs. Natural Grammar is also remotely related to constituency based approaches, in that its types of rules define an induced triadic classification of language concepts which show some analogy to that of X-bar theory.[4] The view taken by the theories of sign processing, like the computational model of Gomes et al. (2003) or the cognitive theory of Deacon (1997), has been a philosophical one, dominantly. An approach, which tries to do justice to sign processing, as a cognitive as well as a semiotic phenomenon, has been first introduced in *A Logical Ontology* (Farkas & Sarbo, 2000) as far as we know. That theory forms the basis of the approach presented in this chapter.

A theory about cognition always involves assumptions about primary concepts. In the theory presented in this chapter qualia are assumed to exist. The quale results from the unifying operation of quale consciousness, it is upon these qualia that the intellect operates.

A quale is defined as '..every combination of sensations so far as it is really synthesized' (cf. CP 6.223). On the one hand this leads to the conclusion that a quale is not confined to seemingly simple sensations like in the perception of 'red', but extends to complex cases like when we perceive a work of art, a chair or this day. On the other hand it raises the question whether it is possible to extend our analysis in order to include the workings of our sensory apparatus and the neuronal fabric that lead up to qualia. This later question falls without our present scope. For the sake of completeness we mention the generative methodology due to Pribram (1971) and Prueitt (1995), which is related to the fundamental work by Gibson (1979). Their aim is to develop a generic mechanism for the definition of primary entities, but also of complex process compartments, in cognitive processing. Their analysis utilizes the concepts of measured perception and spectral properties in order to delve the gap between mind and brain (including the sensory apparatus) or, to put it in the perspective of our model, in order to give a physical/mental account of qualia. An analysis of the relations between that theory and ours is beyond our scope.

TOWARDS A MODEL OF COGNITION

Cognition is concerned with the interpretation of phenomena according to their (possible) meaning. In this chapter we will attempt to deal with a small part of this far too complex problem only, we focus on the restricted domain of *perceptual judgments* (CP 4.540). Inasmuch as we may know about phenomena, by means of observations, the problem of sign recognition can be reduced to a question about the nature of observations. We will consider this issue from two points of view. The first is the *cognitive theoretical* one, according to which, phenomena appear as stimuli and the question is: How can stimuli finally be recognized in a meaningful reaction? The second is the *semiotic* one, according to which, we may know about phenomena by means of signs and the question is: How can signs develop to their meaning?

These two views are interrelated and their dependency is expressed in this chapter by postulating that *stimulus* potentially is a *sign*, and that *reaction* is an interpreting proposition by which a stimulus or percept is transformed into a *fact of immediate perception*. For example, if we observe smoke, which appears as a visual stimulus or percept that can be a sign, then shouting 'fire!' may be our meaningful reaction.

The problem of sign recognition gets further complicated by asking for a computational solution, which tacitly requires a computational model for cognition that satisfies the link between the cognitive and the semiotic viewpoints. This complication can be further developed along two different lines.

The first line of argument is related with Searle's thought experiment in which he develops the Chinese Room Argument. By stating that knowledge emerges from a natural, human process, the rules of which have to be, on that account, natural too, the following question can be raised: Which computational rules and interpretation can satisfy this

condition in such a way that the result of the application of the rules is naturally meaningful too? According to Searle here we are facing a fundamental problem, due to the limitations of the computer. For, contrary to man, the computer is unable to intensionally connect a sign with its object.

The second line of argument departs from the distinction Peirce makes between mechanical- and purposive action. Mechanical action is characterized as a blind, compulsory process that leaves little room for variation; just causes and effects following each other in sequences. Purposive action on the other hand involves the mediation of a goal or purpose that interferes with the course of events. Purposive action aims at the removal of stimulation (cf. CP 5.563), but is quite open ended regarding the means (mechanical processes) used to achieve this goal. It, in short, is learning and introduces abstractions by its reliance on abductive inferences for the generation of satisfactory solutions. This kind of action may be hard to formalize. Our current model does not capture learning, it aims at capturing habits of thought or habits of thought like action.

COGNITION AS A PROCESS

Since we regard cognition as a process, a word about our understanding of processes is in order at this point. A process will be considered to be any sequence of events such, that (1) one event initiates the sequence and another terminates it, (2) every event that contributes to the sequences yielding the terminating event is regarded as part of the process and (3) the terminating event governs the decision of which events make up the sequence. Although the events making up the sequence generate the terminating event (efficient causation), the whole process is governed by its goal (teleological causation). An event will be considered as whatever makes a difference (Debrock, Farkas and Sarbo, 1999).

The input of cognitive processing is the stimulus, which is recognized by the mind/brain. This view of cognition is compatible with the assumption laid down by the Peircean theory of perceptual judgments that the 'real' world is forced upon us in percepts (CP 2.142), from which, perceptual judgments are obtained through interpretation (CP 5.54), by means of a process which is utterly beyond our control (CP 5.115).

A fundamental property of all systems, including biological ones, is their potential for generating an answer (*re-action*) to a stimulus (*action*). The `goal' of this process is the generation of an *adequate* reaction on the stimulus, regarded as an external effect. An important element of response generation is the *interaction* between the external effect (stimulus), on the one hand, and the interpreting system, on the other. From the assumption that the source of all reaction or meaning is an interaction and knowledge is a *re*-presentation of such interactions, it follows that knowledge too must be inherently dynamic: hence a *process*.

The external *effect* (stimulus) is affecting the interpreting system, which occurs at the moment of affectation as a *state*. As anything appearing as an effect can as well appear as a state there must be something common in both. We call this a *quality*, after Peirce. Because

state and effect are in principle independent, all phenomena are considered to be interactions between independent qualities. Let us emphasize that there may be any number of qualities involved in an interaction, but, according to the theory of this chapter, those qualities are always distinguished by cognition in two collections (state and effect), which are treated as single entities. The potential for considering a collection of qualities as a single entity (cf. 'chunking') is an assumption shared by the theory of perceptual judgments as well (CP 7.530).

Processing schema

Phenomena are an *interaction* appearing via the mediation of a *change*, as an *event* (cf. reaction). Following the received theory of cognition (Harnad, 1987) the re-presentation of phenomena by cognition can be modeled as follows.

By virtue of the appearing change, the sensory signal is sampled in a percept. In a single operation, the brain compares the current percept with the previous one, and this enables it to distinguish between two sorts of input qualities (in short, *input*): one, which was there and remained there, which can be called a 'state'; and another, which, though it was not there, is there now, which can be called an 'effect'.[5] In cognitive theory, qualities as perceived are called *qualia*.

The change, signifying an interaction in the 'real' world, can be explained as follows. During input processing the stimulus may change, meaning that its current value and the value stored in the last percept can be different. That difference can be interpreted by the brain, as a change, which mediates the actual value of the stimulus to its current meaning.

The reaction of an interpreting system is determined by the system's 'knowledge' of the properties of the external stimulus, including its experience with the results of earlier response strategies (habit). Such knowledge is an expression of the system's potential for interpreting or *combining* with a type of input effect, depending on the system's state. Such properties shall be called the 'combinatory' properties of the input qualia or the *context* of the observation.

In complex biological systems, 'knowledge' is concentrated in functional units like the sensory, central and motor sub-systems. The most important of these is the central system, which includes the memory. The 'translation' from external stimuli to internal representation (qualia) is due to the sensory sub-system, which itself is an interpreting system, generating 'brute reactions' (translations). For the goal of this paper, the role of the motor sub-system is secondary, therefore omitted.

The primary task of cognitive processing is the interpretation of the external stimuli, by making use of the latter's combinatory properties. Since the input is assumed to consist of two types of qualia (state and effect), together appearing as a 'primordial soup' ($[q_1\ q_2]$), the stages of processing can be defined as follows (see also fig. 1). Square brackets are used to indicate that an entity is not yet interpreted as a sign; no bracketing or the usual bracket symbols indicate that some interpretation is already available.

(1) The sorting out of the two types of qualia in the 'primordial soup'
   as state and effect respectively.
   Sorting: $[q_1]$, $[q_2]$
(2) The separation of the collections of the two types of qualia.
   Abstraction: q_1, q_2
(3) The linking of the qualia with their combinatory properties ([C]).
   Complementation: (q_1,C), (q_2,C)
(4) The establishment of a relation between the completed qualia.
   Predication:[6] (q_1,C) (q_2,C)

$$(q_1,C) - (q_2,C)$$

(q_1,C) $\qquad$ (q_2,C) $\qquad$ *(4) predication*

$\qquad$ *(3) complementation*

q_1 $\qquad$ $[C]$ $\qquad$ q_2 $\qquad$ *(2) abstraction*

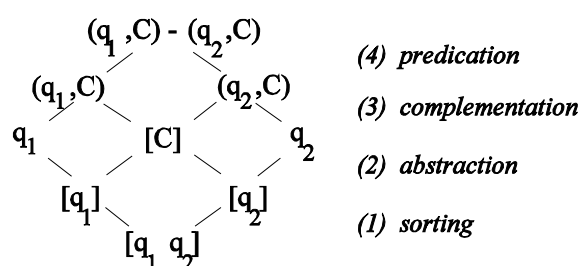$[q_1]$ $\qquad$ $[q_2]$ $\qquad$ *(1) sorting*

$$[q_1 \ q_2]$$

Figure 1: The schematic diagram of cognitive processing

PERCEPTION AND COGNITION

In an earlier version of the model of this paper *A Peircean Ontology of Semantics* (Farkas & Sarbo, 2002) we introduced two levels of cognitive processing, which we called perception and cognition. The 'goal' of *perception*, as a process, is the establishment of a relation between the input qualia and the memory information (the importance of the relation between the input qualia is secondary in this process). As a result, perception obtains the meaning of the qualia in *themselves*. In accordance with perception's goal, the memory response or the context ([C]) contains information about the properties of the input qualia independently from their actual relations. This information is 'iconic' (cf. lexical meaning).

The state and effect types of input qualia are indicated by *a* and *b*, respectively; those of the memory by *a'* and *b'*. All four signs may as well refer to a type as to a collection of qualia.

Among the representations obtained by perception, only the final one (step 4) is of interest for the rest of this paper. We assume that the *a'*(*b'*) memory response arises by means of the *a*(*b*) input qualia, triggering the memory. Although the two types of memory response signs are independent, they contain reference to a *common* meaning. This is due to the existence of interaction between the input qualia.

Depending on the activation of the memory, there may be qualia in the memory response ([C]) having an intensity above (i) or below (ii) threshold, respectively referring to an input (meaning) which is in the brain's *focus*, and which is only *complementary*. The distribution of

the roles in any given case depends on the actual activation of the memory which defines the state of the mind/brain.

A high intensity response of type (i) signifies the recognition of the input as an agreement between the input and the memory response: $a(b)$ is recognized or 'known' as $a'(b')$. A low intensity response of type (ii) refers to input recognition as a possibility only: the input $a(b)$ is not recognized or 'not known' as a'(b') as a consequence of which the memory response only represents a secondary or even less important aspect of the input qualia.

By indicating the first type of relationship between input and memory response by a '$*$' symbol, and the second type by a '+', the signs of perception can be represented as: $a*a'$, $a+a'$, $b*b'$, $b+b'$. For example, $a*a'$ signifies the event of positive identification of a by a' (type (i)); as opposed to $a+a'$ (type (ii)) which refers to the identification of a possible meaning of a by a' (hence, to a *denial* of a positive identification).

In perception as an actual process the four signs are presented as a single sign. The *recognition* of the difference between the four types of relations is beyond its scope.

Cognition

The process of cognition is an exact copy of the process of perception except that the 'goal' of cognition is the interpretation of the relation between the input qualia which are in the focus, in the light of the qualia which are complementary. (Now it is the relation between input and context which is secondary.) In accordance with cognition's 'goal', the context ([C]) contains relational, complementary information about the input qualia, which involves indexicality.

In the process of cognition the input appears as a 'primordial soup' too, which is defined by the synonymous signs of perception. In fact, the difference between the four meaning elements functions as a ground for the process of cognition. This is acknowledged in our model by the introduction of an initial re-presentation of the four relations that function in perception: $a*a'$ as A, $a+a'$ as ¬A, $b*b'$ as B and $b+b'$ as ¬B. The presence or absence of a '¬' symbol in an expression indicate whether the qualia signified, are or are not in focus. Hence '¬' can be interpreted as a 'relative difference' with respect to the collection of a type of qualia (state or effect), represented as a set. How the processing schema of sect. 4 can be instantiated for cognition, is depicted in fig. 2.
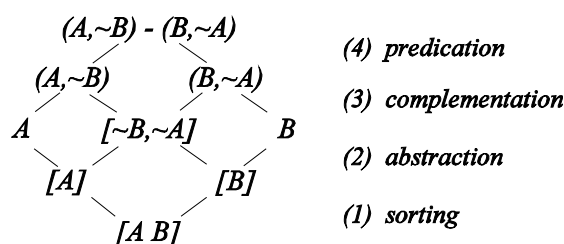

(A,~B) - (B,~A)   (4) *predication*

(A,~B)   (B,~A)   (3) *complementation*

A   [~B,~A]   B   (2) *abstraction*

[A]   [B]   (1) *sorting*

[A B]

8

Figure 2: The schematic diagram of cognition, as a process[7]

We especially want to point to step 3, in which the link between input qualia and context is established. This is done in accordance with the 'goal' of cognition and the duality of phenomena alike. This explains why there can be a relation between A and $\neg B$, and $\neg A$ and B, and why there is no relation between A and $\neg A$, or B and $\neg B$.[8] Finally, in step 4, the cognition process is completed by establishing the relation between A and B.

The three relations, which correspond to the three types of interactions between the input qualia, can be characterized by means of the meaning of their constituents (from a computational stance this interaction is a relation that will be indicated by a ' ' symbol):

(1) $A \quad \neg B$:
   A is 'known', but B is 'not known';
   the complementation of the input state ('actualization').
(2) $B \quad \neg A$:
   B is 'known', but A is 'not known';
   the complementation of the input effect ('refinement').
(3) $(A, \neg B) \quad (B, \neg A)$:
   both A and B are 'known';
   the assertion of the relation between A and B ('proposition').

If neither A nor B is `known', interpretation terminates before reaching its goal, meaning that cognition does not occur. The reader may have noticed the mediative function of the context signs which is operative in step 3. Indeed, through the correspondence between the two signs, $\neg A$ and $\neg B$, which are triggered by the same input, the context *implicitly* determines the actual relation between A and B. That relation can be called a 'proposition' resulting from a hypothetic inference, but only if we acknowledge, in accordance with the Peircean view of a perceptual judgment, that the percept's "truth consists in the fact that it is impossible to correct it, and in the fact that it only professes one aspect of the percept" (CP 5.568).

LOGICAL ANALYSIS

The above interpretation of cognition already illustrates, to some extent, the completeness of that process, but this becomes even more clear from a logical analysis of the processing schema. This section is an attempt to elaborate such an analysis, on the basis of the model of cognition introduced above. It is good to bear in mind that the results are directly applicable to the model of perception as well. The hidden agenda of this section is the tacit introduction of logical operations in the process model of cognition. What makes the use of such concepts valuable is that they have a well studied, precise meaning.

An essential element of a logical approach to cognition is the abstraction of a common

meaning for the different types of qualia, which is the concept of a *logical variable*. In virtue of the duality of the input, the logical interpretation of the process model of cognition requires the introduction of two variables, which are denoted by A and B. The difference between the qualia which are in the focus and which are only complementary, is represented by the difference of their expression by means of a logical variable which is stated positively or negatively. Perceived state and effect qualia which are in the focus are indicated by A and B respectively; those which are complementary by $\neg A$ and $\neg B$. Here, '$\neg$' denotes logical negation, that is, relative difference with respect to the collection of a type of qualia, represented as a set. For example, the complementary subsets of the set of A-type qualia are denoted by A and $\neg A$ (hence the label A is used ambiguously).

Conform the above mapping, the logical meaning of the cognitive relations can be defined in the following way. The relational operators introduced in the instantiation of the processing schema for perception ('+' for possibility and '$*$' for agreement) are inherited by the cognitive model and its logical interpretation as logical 'or' ('+') and 'and' ('$*$').

$[q_1]= A+B$, $[q_2]= A*B$:

> expresses the simultaneous presence of the input qualia which are in focus as a simple, possible co-existence ($A+B$); and in the sense of agreement, as a meaningful co-occurrence ($A*B$), respectively.

$q_1= A*\neg B$, $\neg A*B$:

> expresses the abstract meaning of the input qualia which are in the focus as constituents, irrespective of the actually co-occurring other type of qualia. It is *this* perspective that makes the two signs synonymous (the "," in the definition of q_1 directly above is a meta-level expression of this equivalence).

$q_2= A*\neg B +\neg A*B$:

> expresses the input as an abstract co-occurrence in terms of a compatibility relation (a possible co-existence) of the two types of abstract constituents of the input (which are now considered as being different). In *A Logical Ontology* (Farkas & Sarbo, 2000) we have proved that the logical expression of q_1 and q_2 can be formally defined as the relative difference of $[q_1]$ and $[q_2]$. The context ($[C]$) is defined by the complementary qualia represented as a co-existence ($\neg A+\neg B$) and as a co-occurrence relation ($\neg A*\neg B$). The synonymous representation of these signs is an expression of the complementary (secondary) meaning of the qualia, but also of the common property referred to by the simultaneously present $\neg A$ and $\neg B$ type of qualia comprising the context.

$(q_1,C)= A+\neg B$, $\neg A+B$:

> expresses the abstract constituent (q_1) completed with the meaning of the context ($[C]$) or, alternatively, the 'actual' meaning of the input qualia as constituents. For example, the actual meaning of A as a constituent, is defined by A itself and by $\neg B$, the complementary property linking A with B, implicitly (as the relation between A and B is not yet

10

established, the *B* type qualia cannot contribute to the actual meaning of *A*, as a constituent). Alternatively, the meaning of $\neg A *B$ in context, is defined by the qualia completing this abstract meaning, which are: *A* and $\neg B$. As the two interpretations of *A* as an actual constituent are related to each other by the relation of co-existence, the logical meaning of (q_1,C) can be expressed by $A+\neg B$. For the same reason, as in q_1, the two representations of (q_1,C) are interpreted in the model as synonymous.

$(q_2,C)= A*B +\neg A*\neg B$:

expresses the abstract compatibility relation in context, thus interpreting the input as a characteristic property which appears as an event. That such an event occurs between *A* and *B* or, alternatively, between $\neg A$ and $\neg B$, represents the interaction which is in the focus, respectively, positively and negatively. Again, we refer to *A Logical Ontology* (Farkas & Sarbo, 2000), in which we have proved that the logical expressions of (q_1,C) and (q_2,C) can be formally defined as the logical complements of, respectively, q_1 and q_2, by means of interpreting the interaction with the context ([C]) as a logical negation operation ('$\neg$').

(q_1,C) $(q_2,C)= $ *A* is *B*:

expresses the logical relation between the input qualia which are in the focus, represented as a proposition.

The logical expressions describing the process of cognition are summarized in fig. 3. The logical signs, '*0*' and '*1*', which are omitted, can be defined as representations of a 'not-valid' and a 'valid' input, respectively. Notice in fig. 3 the presence of *all* Boolean relations on two variables, reinforcing our conjecture concerning the completeness of the underlying cognitive process. The results of the above analysis show that logical signs (hence also the concepts of cognition, as a process) can be defined as a relation (interaction) between *neighboring* signs that is in need of settlement. In fig. 3, such signs are connected with a horizontal line.
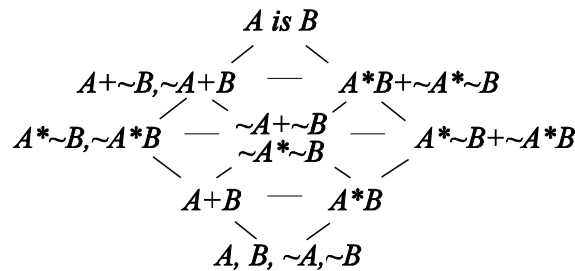
$$
\begin{array}{c}
\textit{A is B} \\[2pt]
\diagup \qquad \diagdown \\[2pt]
A+\sim B, \sim A+B \quad - \quad A*B+\sim A*\sim B \\[2pt]
\diagup \qquad \diagdown \quad \diagup \qquad \diagdown \\[2pt]
A*\sim B, \sim A*B \quad - \quad \begin{array}{c}\sim A+\sim B\\ \sim A*\sim B\end{array} \quad - \quad A*\sim B+\sim A*B \\[2pt]
\diagdown \qquad \diagup \qquad \diagdown \qquad \diagup \\[2pt]
A+B \quad - \quad A*B \\[2pt]
\diagdown \qquad \diagup \\[2pt]
A,\ B,\ \sim A, \sim B
\end{array}
$$

Figure 3: The logical signs of cognition, as a process

The figure can be traversed by the application of the operation of relative difference to the connected pairs. We make a distinction between three types of this operation. *Sorting* is relative difference with respect to qualia in themselves. The input contains two types of qualia which are in the focus, *A* and *B*, that we represent from the point of view of co-existence $(A+B)$ and co-occurrence $(A*B)$,[9] that is to say as sorted qualia. *Abstraction* is

relative difference of sorted qualia with respect to each other. An example is $\neg A_*B+A_*\neg B$. The reader may check this by computing *(A+B)\(A*B)*. *Complementation* is relative difference of an abstracted quality with respect to the input as a whole. An example is $A_*B+\neg A_*\neg B$. The reader may check this by computing *1\(A*¬B+¬A*B)*.

SEMIOTIC ANALOGY

That the formal computational and the intuitive interpretation of a sign are tightly related to each other must be clear from the above explanation of the logical relations of cognition. This dependency forms the basis for the semiotic interpretation of those 9 types of relations, which can be explained as follows.

[q_1]: represents that the constituents are trivially part of their collection, as a whole. Hence they are similar to it. So, the representation of the input, as a constituency relation, expresses *likeness* with respect to the input, which is represented as 'primordial soup'.

[q_2]: represents that the aspect of *simultaneity* is a primary element of the input, as an appearance (event) that happens now.

q_1: represents that the abstract conception of the input is an expression of its being as a *qualitative possibility*.

q_2: represents that the compatibility of the abstract meaning of the input qualia is expressive of a *rule*-like relation.

(q_1,C): represents the meaning of the abstract constituents in context. It is a definition of the *actual* meaning of the input qualia, as something *existent*.

(q_2,C): represents the interpretation of the abstract compatibility relation in context as a characteristic property; it presupposes the existence of a *consensus*.

(q_1,C) (q_2,C): represents that the assertion of a relation between the input qualia involves the formation of a *proposition* which is a hypothesis.

From this semiotic interpretation of the logical relations, the analogy with the Peircean nonadic sign classification follows trivially. A serious treatment of this classification would demand a chapter of its own; here we will have to do with some introductory remarks.[10]

   Throughout his philosophical career Peirce was occupied with attempts to classify signs in a systematic way. The roots of the system reside in his phenomenological work on the Doctrine of Categories, which, in the spirit of Kant, has the task to "…unravel the tangled skin of all that in any sense appears and wind it into distinct forms;.."(CP 1.280). This work led him to believe that there are three basic categories: monadic Firstness (the possible), which appears in consciousness as feeling or the consciousness of quality without recognition or analysis; dyadic Secondness (the actual), which appears as a consciousness of interruption in the field of consciousness or as the brute intrusion of another quality; triadic Thirdness (the lawful), which synthesizes the content of consciousness or the mediation by thought of the

different feelings spread out in time (cf. CP 1.377).

The work on sign classification proceeds by repeatedly applying the three basic categorical distinctions to signs. It starts with the definition of a sign as "…something, A, which denotes some fact or object, B, to some interpretant thought, C." (CP 1.346). This definition yields three ways in which sign may be considered. First, a sign may be considered in itself. If we do so, we neglect the relations a sign may have with its object and interpretant and we only regard the sign as a possible sign. Second, we may regard the sign in its relation with its object only and neglect the relation it has with its interpretant. If we do so, we regard the sign as an existing sign, but still without any effect. And third, we may look how the sign addresses its interpretant. If we do so, we regard the sign as a real or effectual sign. In this last case we try to unravel the *full meaning* or import a sign *may* have by figuring out how the sign manages to relate the interpretant of the sign with its object. If we concentrate on a sign-interpretant sequence in some concrete situation, we study embedded signs.

In a second round Peirce applies the categorical distinctions to the sign relations just discerned. The first tenable result, the *nonadic classification*, is summarized at the left hand side of figure 4, the bottom right diagonal gives the relational aspects pertaining to the sign in itself, the intermediate gives the aspects pertaining to the way the sign may relate to the object and the top right diagonal gives the ways in which the sign may address its interpretant. On the right hand side of figure 4 the technical terms that give the meaning aspects are stated in more mundane terms, which are also used in the semiotic interpretation of the 9 types of relations above.

We get at a typology of signs by selecting a term from each diagonal. The least developed existing sign type is a rhematic, iconic sinsign (for instance a term), involving qualisigns. The most developed an argumentative, symbolic legisign in which all less developed signs are involved. This model of nine sign aspects yields ten sign types as a result of the constraint that the categorical value of a term on a higher right diagonal cannot be higher then the value of a term on a lower right diagonal. By taking the process of cognition, which always is an argument, as our focus and by assuming that all less developed meaning aspects are involved, we turn the static nonadic classification into a dynamical process model.
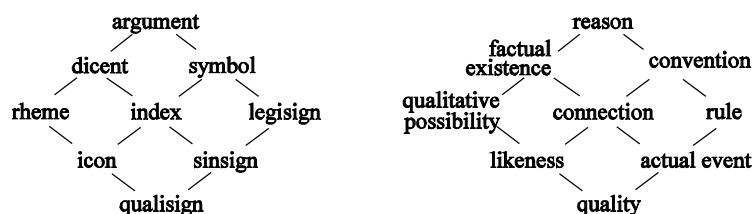

Figure 4: The Peircean types of signs and their aspects of meaning

It is important to note that we interpret the relational aspects of the nonadic classification as the parameters of (full) meaning. The *isomorphism* between the cognitive process and Peirce's nonadic classification is a consequence of the isomorphism between the induced order of cognitive processing, on the one hand, and the interdependency of the Peircean sign aspects, based on categorical distinctions, on the other. But this is all there is! The model

introduced in this paper is suited for a *computational* interpretation. But, although the above mapping establishes a link between the Peircean signs and the Boolean logical relations, hence define a computational level of authentic semiosis, the full meaning of the Peircean *sign types* is *qualitatively* more than such logical relations, since the relations always exist between two sign aspects, while a sign type irreducibly contains three aspects. Finally, let us mention that the process view of signs can be introduced also from the Peircean theory itself (see Debrock, Farkas & Sarbo, 1999).

By assuming that the full meaning of a sign emerges, through *embedding* in real life interaction with the world, from the relations 'generated' by cognitive processing, the 9 sign aspects can be *hypothetically* considered to be a link between the computational and the semiotic level of meaning. Inasmuch as those aspects are enclosed in a process, which finally results in something that can be characterized as truly meaningful, it is probably best to consider the 9 aspects as unfinished 'meaning elements', that is, as *signs* which are in a process of becoming signs (Breemen & Sarbo, 2005). Such signs are called in this paper, pre- or *proto*-signs[11] (Sarbo, 2005).

The different characterizations of knowledge - a combinatory process of qualia (fig. 2); a representation of logical relations (fig. 3); a hierarchy of increasingly more complex 'meaning elements' (fig. 4); but also its other possible interpretations - are interrelated, and it is their collection that approaches full meaning. The conjecture of this research is that, if a uniform representation can be proved to exist, this can be the key for an efficient (computational) merging of knowledge obtained in different domains, into a single representation. Experimental evidence pointing in this direction is found in Hagoort et al. (2004).

COMBINATORY RELATIONS AND PROPERTIES

We may observe an entity, as a state, only by virtue of an appearing effect, but the occurrence of an effect always entails the existence of a state. This asymmetry between state and effect is the ground for a semiotic interpretation of the differences between the three types of relations recognized by cognitive processing.

(1) $A \quad \neg B$:
   A is a potential meaning, which is actualized by ¬B.
(2) $B \quad \neg A$:
   B, which is in principle self-sufficient, receives its
   full meaning from its association with ¬*A*.
(3) $(A, \neg B) \quad (B, \neg A)$:
   A and *B*, which are self-sufficient both,
   together generate a new meaning.[12]

The interpretation of these (cognitive) relations, as syntactic signs, has been introduced in

Syntax from a Peircean perspective (Debrock, Farkas & Sarbo 1999). Syntactic signs, which are a representation of the three types of a *nexus* between syntactic symbols, correspond to the Peircean categories of Firstness, Secondness and Thirdness, for example, a primary syntactic entity (a word), a syntactic modification (of a noun by an adjective) and syntactic predication (subject and predicate forming a sentence) respectively. The important consequence of this transitive relation between cognition and the categories is the existence of a necessary and sufficient condition for ontological specifications, which are typically syntactic relational too, in particular those which are also meant to be used in computer applications. The analysis of the meaning of the constituents, in the three types of relations, proves that the specification of the (combinatory) properties of qualia can be restricted to three cases. The specification of

(1) the qualia in themselves;
(2) with respect to other qualia which are complementing it
    or, which they are complementing;
(3) with respect to other qualia, together with which, they can
    generate a new meaning.[13]

We call such a specification a trichotomic specification or, briefly, a *trichotomy*. In virtue of the dependency between the Peircean categories, the meaning of a more developed class contains the meaning of a less developed one, in a trichotomy. By assuming that the three types of meanings can be defined in each trichotomic class, recursively, we have in front of us the hierarchical schema of ontological specification suggested by the theory of this research.

A framework which is remotely related to the approach presented here, is the theory of Nonagons (Guerri, 2000), which has been introduced originally for supporting the completeness of a design, for example, an architectural design. Nonagons are also based on Peirce's sign aspects and a recursive expansion of his nonadic classification. There is however an important difference between the two approaches in regard to the interpretation of a sign, either as an entity which emphasizes its character as a single unit (the Nonagon approach) or, as an entity which stresses the inherent duality implied by truly triadic semeiosis (the view maintained by this work).[14] A practical advantage of the latter view lies in its capacity for defining the 9 classes as a product of (dual) trichotomies, thereby simplifying the specification task, potentially. An example, illustrating the benefits of recursive specification of the properties of qualia, in text summarization, is presented in (Farkas & Sarbo, 2005) and (Sarbo & Farkas, 2004).

In sum, there are 3 types of relations between signs, in accordance with the 3 categories of phenomena. Qualia, which are the constituents of a syntactic sign interaction, can be analogously characterized, recursively as:

(1) a quality, which is a potential existence;
(2) a state, which appears by virtue of an effect (*a* or *A*);

(3) an effect, which implicates the existence of a state (b or B).

The three categories are not independent from each other. Though thirdness is the most developed, nevertheless it requires secondness and firstness (the latter via the mediation of secondness). Analogous with the categorical relations, an effect can be said to contain a state and, transitively so, a potential existence. By means of the induced ordering of the dependency between the categories ('$<$'), as a polymorphic operation, the relation between the cognitive types can be abstracted as follows: $a<b$ and $A<B$. For example, the meaning of a quale which is an effect, implies the existence of its meaning as a state.

Example

This section contains an example, illustrating the recognition of the 'real' world phenomenon, smoke, as the sign of danger. Assume, we are watching for some time the smoke rising from the chimney of a roof, and suddenly we 'see' fire to blaze up. The input qualia of perception can be defined as follows (boldface symbols are used to indicate input qualia):

   $a=$ **smoke**, $b=$ **fire**

and the final signs of perception, re-presented as the initial signs of cognition (italic is used to indicate memory signs):

   A= **smoke**∗*smoke*,   ¬A= **smoke**+*roof-in-burning*
   B= **fire**∗*fire*,   ¬B= **fire**+*burning*

The recognition of these signs yields the following relations (in a reference to an interpretation which is dealt with as 'not known', the input is omitted):

   $(q_1,C)=$ (**smoke**∗*smoke*, *burning*)
   $(q_2,C)=$ (**fire**∗*fire*, *roof-in-burning*)

which together generate the proposition, through the mediation of the *burning* of the *roof*:

   (q_1,C) $(q_2,C)=$ (**smoke**...) IS (**fire**...)

This sign can be re-presented, eventually, by shouting "*Fire!*" or, simply, by "*running away*", as our interpretant reaction, assuming there is a real need.

In order to enable the recognition of the input as such a meaningful relation, the input qualia have to be adequately specified by means of trichotomies. Let us exemplify this with the specification of *smoke* (hence implicitly also of **smoke)**, which can be regarded:

(1) in itself: as an entity which is a quale, having properties underlying its combinatory potential, like color, density, etc.
(2) in relation to another quale which is complementing it: e.g. *rising-from-the-chimney*; or, which it is complementing: e.g. *blowing* (as a smoke producing effect).
(3)as a self sufficient sign: like the subject of *any-burning*.

Notice that the qualia of (1) function as the ground for the connections of (2), which in turn underlie the meaningful relations of (3).

NATURAL LANGUAGE

The process model of cognition can be applied to natural language easily. In natural language the input qualia can be a morpheme or a word (a morpho-syntactically finished symbol), in a morphological and a syntactical analysis, respectively. The perception process, linking the input with memory information, corresponds to lexical analysis. The 'single stimulus' or percept view of the cognitive model of this chapter requires that, conceptually, the entire input is present in a single observation. To this end, the order of appearence of the input symbols can be represented as a quale,[15] which eventually leads to a sequential model of cognitive processing. In the model of language, as syntactic signs, the second process ('cognition') corresponds to parsing, interpreting the input symbols, as (morpho-)syntactic *relational needs* (combinatory properties), which may combine (*bind*). In the rest of this chapter the focus will be on syntactic symbols, the process of morpho-syntactic recognition is omitted.

Step 1 (cf. fig. 1) corresponds to the classification of the input symbols as nominals ($[q_1]$) and non-nominals ($[q_2]$). Which one of the symbols is in the focus, and which is only complementary, follows from the syntactic rules and the order of appearance of the input symbols. Step 2 amounts to the identification of noun (q_1) and verb phrases (q_2) but also of modifiers and complements ($[C]$), step 3 is devoted to modification ((q_1,C)) and complementation ((q_2,C)), and step 4 to syntactic predication.

The classification of syntactic concepts is depicted in fig. 5, on the left-hand side. On the right-hand side you may find an illustration of the syntactic concepts, by the analysis of a simple utterance. The goal of this example is twofold. The first is an analysis of syntactically meaningful concepts and their dependency; the second is an illustration of the sequential order of the input symbols (cf. surface structure) as potentially meaningful.
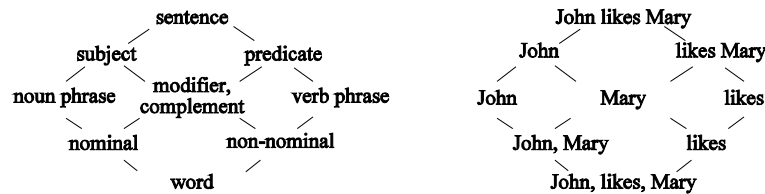
Figure 5: *The classification of syntactic concepts and the concepts of the sample utterance `John likes Mary'*

Sequential processing

Contrary to the earlier example of smoke-and-fire, in the language phenomenon presented in fig. 5, "John likes Mary" (in short *J l M*), we have more than two qualia, but these too can be distinguished in two collections: [*JM*] and [*l*]. These collections are interrelated: '[*l*] happens to [*JM*]'. The linguistic interpretation is more refined, however. The relation between *l* and *J* is different from the one between *l* and *M*. This difference is signified, in English, by means of the order of the input symbols. This ordering, as a quale, is recognized by the language user, together with the other meaning(s) of the involved symbols.

This line of thinking has led us to a sequential version of our model of language, in which the effect of the order of the input symbols is defined by means of the nine Peircean sign aspects (Sarbo & Farkas, 2001). On surface level, the input symbols appear one after the other. Because each symbol may contribute to the meaning of the entire sequence (the sentence) only as a proto-sign, the recognition of the individual input entities may overlap. In general, the processing of subsequent sentences may overlap as well. In the utterance of fig. 5, the input symbols appear as qualisigns.[16] As the input qualia are, in principle, independent, but also partake in the same sentence (phenomenon), the appearance of *l* forces us to reconsider the earlier interpretation of *J* (qualisign). A possible solution of this can be the *re*-presentation of *J*, as a constituent (icon) of the entire input, in accordance with the principle of economy, characterizing language recognition, which states that a less developed representation of a phenomenon has to be generated before a more developed one.

The appearance of *M* has similar consequences on *l* and, transitively so, on *J*. The latter is due to our assumption that the signs yielded by 'sorting' (cf. sect. 4) signify the qualia of a single phenomenon, as potential co-existence (icon) and co-occurence (sinsign). As *J* and *l* together do not mediate such a meaning, it follows that *J* has to be re-presented again, but this time as a possible abstraction of the entire input (rheme). Notice that any re-presentation of a symbol must contain the earlier meaning states of the same symbol. The subsequently appearing dot symbols (the role of which will be explained in the next section) trigger a chain of representations, of which we explain only the re-presentation of the nominal meaning of *M* (icon). We start with applying the same strategy, like in the case of *J* (icon), but then parsing will eventually fail, as we cannot represent the entire input in a single sign. So, we assume the parser backtracks until the first point that provides an alternative[17] which

represents M as a complement of *l*. This example also illustrates how our model may discover whether a symbol is part of the complementary context, if that property is not indicated otherwise. The parsing of the example of fig. 5 is displayed in fig. 6.
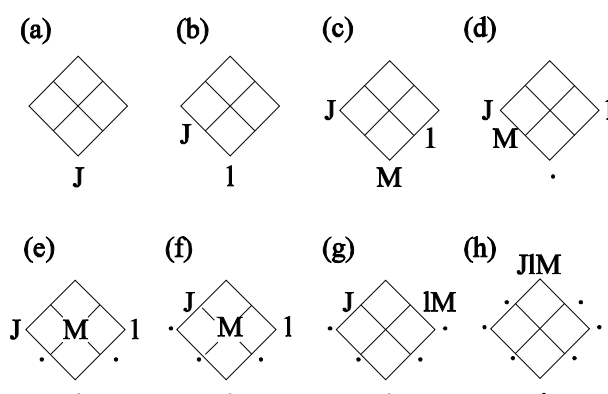

Figure 6: *Syntactic analysis of 'John likes Mary'*

Due to the sequential processing of the input we have to consider two new cases of an interaction (relation) between symbols: the first is *accumulation*, which is a 'binding' between signs having an identical status and compatible information; the second is *coercion*, which is an interaction that does not actually happen. What does happen in such an interaction is that an existing sign is forced to be re-presented by or is 'coerced to' a more meaningful interpretation. Coercion and accumulation are degenerate versions of (normal) binding.

The specification of the qualia with respect to other qualia which are complementing it or, which they are complementing (cf. sect. 'Combinatory relations …'), corresponds to syntactic modification and complementation, respectively. This indicates that the two types of phenomena can be treated *uniformly*, in our model.

Towards a formal model

The syntactic interpretation of the classification of qualia (cf. sect. 'Combinatory relations …'), as the constituents of a binding defines the three types of syntactic relational needs acknowledged by our model: (1) neutral, (2) passive and (3) active (in short, *n-*, *p-* and *a-*need). The 'goal' of language cognition, as a syntactic process, is syntactic well-formedness. This is formally interpreted by requiring that the final representation of the (entire) input may not have any unsatisfied relational needs; it must be neutral, syntactically. The different types of binding can be characterized, from this point of view, as follows. A coercion 'satisfies' an *n-*need; in an accumulation two relational needs of the same type are merged to a single need; a (normal) binding satisfies a pair of *a-* and *p-*needs.

Fig. 7 summarizes the potential syntactic relational needs for the types of speech. The input qualia are defined as follows: *A*=noun; *B*=verb, adjective, adverb,[18] prep(-compl), where 'compl' can be a noun, verb, adjective or adverb. The trichotomic specification of syntactic symbols is such that, through re-presentation, *A*-type symbols can have a relational

potential, as an (1) icon, (2) rheme or index, and (3) a dicent type of sign; and, *B*-type symbols, as a (1) sinsign, (2) legisign or index, and (3) a symbol type of sign. In English, the category related dependency between the different interpretations of a symbol, as an *A*- or a *B*-type quale, is used, amongst others, for the modeling of modification phenomena like '**runs quickly'**. In this example, the *a*-need of **quickly** (index) satisfies the *p*-need of **runs** (legisign), indicating that the effect (**run**) is considered in this interaction, as a state (**run**). Also, object complementation phenomena like '**painted black**' can be modeled isomorphically.

We define a 'dot' symbol as an *A and B* type sign, which cannot bind except with its own type. Dot symbols can be used to force the 'realization' of pending interactions. We assume that the entire input is closed by a constant number of dots. From the specification of fig. 7 and the earlier introduced properties of sequential processing, a formal definition of a Natural Grammar of English can be derived easily[19] (Sarbo & Farkas, 2002). In figure 7 optional *n*-needs are omitted.
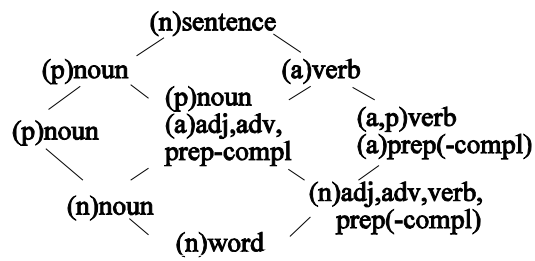

Figure 7: Classification of syntactic symbols

The *kernel* of an algorithm for the parsing of coordination has been presented in (Sarbo & Farkas, 2004). The essential point of this algorithm is the merging of signs having an identical sign type, to a synonymous meaning. Multiple modification and complementation, but also embedded clauses can be modeled by means of recursive incarnations of the parsing 'machinery'. In (Sarbo & Farkas, 2002) we formally proved that the complexity of cognitive processing in our model is *linear* in the number of input qualia, and operations on them.

Example

This section is an attempt to illustrate the potential of our approach for modeling complex linguistic phenomena. The presentation of the analysis is simplified by introducing a tabular form for the sign 'matrix' as used in fig. 6, in which, a column corresponds to a sign aspect, used as an indicator of the processing status, and a row to re-presentation act(s), arising due to the application of *rules*, indicated in last column. The following abbreviations are used: input (i), accumulation (a), coercion (c), binding (b). Accumulated signs are separated by a "/" sign. The names of some of the sign aspects are abbreviated. The parsing of dot symbols is omitted. Predication is not displayed due to lack of space.

Table 1: Syntactic analysis of *A man entered who was covered with mud*

nr.	qual	icon	sins	rhem	index	legis	dicent	symbol	rule
0	a man (a_m)								i
1	entered(e)	a_m							i,c
2	who(w_h)		e	a_m					i,c,c
	recursion								
3	was covered (w_{cd})	w_h							i,c
4	with mud (w_m)		w_{cd}	w_h					i,c,c
5			w_m	w_h		w_{cd}			c, c
6				w_h	w_m	w_{cd}			c
7					w_m	w_{cd}	w_h		c
8						w_h		w_{cd}-w_m	b
	return								
9	who-..-mud(w_{cm})		e	a_m					i
10			w_{cm}	a_m		e			c,c
11				a_m	w_{cm}	e			c
12						e	a_{hm}-w_{cm}		b
13							a_m-w_{cm}	e	c

The first example illustrates discontinuous modification. Let us take as our starting point the morpho-syntactic analysis of the sample utterance: (A man) (entered) (who) (was covered) (with mud). In the syntactic parsing (cf. table 1) we make use of recursion in the analysis of the segment initiated by 'who' and closed by the sentence ending dot. In the recursively analyzed part, 'who' functions as a (dummy) subject. The final sign of this segment is represented degenerately, by a single quale (w_{cm}), having an adjective-like relational need due to the unsatisfied relational need of 'who'.

The second example (cf. table 2) illustrates coordination. We assume the morpho-syntactic analysis: (Mary) (is) (a democrat) (and) (proud) (of it). In the syntactic analysis, in step 8, 'proud' and 'of it', are accumulated in a single sign. The coordination of 'proud of it' with 'a democrat' is possible, as both symbols can be 'is'-complements, syntactically.

Table 2: Syntactic analysis of *Mary is a democrat and proud of it*

nr.	qual	icon	sins	rhem	index	legis	dicent	symbol	rule
0	Mary(M)								i
1	is(i)	M							i,c
2	a democrat (ad)		i	M					i,c,c
3	and(&)	ad		M		i			i,c,c
4					ad	i	M		c,c
	save								
5	proud(p)								i
6	of it(oi)		p						i,c
7			oi		p				c,c
8					p/oi				c,a
	coordination								
9					ad&p/oi				
	restore								
10					ad&p/oi	i	M		

SUMMARY

An advantage of our model of knowledge representation presented in this chapter lies in its potential for generating information by the computer, that the human user may directly process *naturally*. The essence of such processing can be explained by means of the metaphor of apparent motion perception. If the snapshots of a series are presented correctly, we may easily experience the meaning of the series, as a whole. If the presentation is not correct, as for example when the snapshots are in the wrong order or the difference between the consecutive pictures is too large, an adequate interpretation may still be possible, but it can be difficult.

The idea behind our theory is that an analogous "correct" presentation of information in knowledge modelling, may entail an immediate 'natural' interpretation of the computations yielded by the cognitive model (cf. frames), as different representations of a *single* interaction between some state and effect. More specifically, the hypothesis of this work is that information processing that does respect the 9 types of relations of cognitive processing and their ordering, may *enhance* the interpretation of the full meaning of such computations as a whole (through the mediation of proto-signs).

This relation between the cognitive and semiotic concepts of meaning production is the key to the *natural* definition of the combinatory properties of qualia in language as a *habit*. Additionally, the processual interpretation of the Peircean classification is the key to the understanding of sign recognition, generating increasingly better *approximations* of the final meaning of the observed phenomenon. It can be shown that there exists a correspondence between the types of relations generated by cognitive processing on the one hand, and the interactions between sign aspects that are each other's neighbors,[20] according to Peirce's classification, on the other. From the isomorphism of the representation of knowledge in different domains it follows that different interpretational viewpoints (like syntactic, logical, etc.) can be merged to a single whole through coordination (in a broad sense).

What is `natural' in natural language? In our view the natural aspect of language involves the *types of distinctions* that can be made cognitively; the organization of such events in a *process*; and the appearance or 'feeling' of such a process as knowledge (the last lies beyond the scope of our model). This is the basis, in our opinion, for the understanding of *computational semiotics* as a cognitively based semiotic and, *therefore*, as natural computation.

Bibliography

Goldstone, R.L. & Barsalou, L.W. (1998). Reuniting perception and conception. *Cognition*, 65, 231-262.

Breemen, A.J.J. van & Sarbo, J.J. (in press). Surviving in the Bermuda Triangle of Semeiosis. *International Journal of Computing Anticipatory Systems*.

Deacon, T.W. (1997). The symbolic species: the co-evolution of language and the brain. New York: W.W. Norton.

Debrock, G.Y., Farkas, J.I. & Sarbo, J.J. (1999). Syntax from a Peircean perspective. In P. Sandrini, (Ed.), *5th International Congress on Terminology and Knowledge Engineering*, (pp. 180-189). Innsbruck (Austria).

Farkas, J.I. & Sarbo, J.J. (2000). A Logical Ontology. In G. Stumme (Ed.), *Working with Conceptual Structure*s: Contributions to ICCS'2000 (pp. 138-151). Darmstadt (Germany): Shaker Verlag.

Farkas, J.I. & Sarbo, J.J. (2002). A Peircean ontology of semantics. In U.Priss and D. Corbett (Eds.). *Conceptual Structures: Integration and Interfaces* (ICCS'2002), volume 2193 of LNAI, (pp. 177-190). Borovets (Bulgaria): Springer-Verlag.

J.J. Gibson, J.J. (1997). *The ecological approach to visual perception*. Boston: Houghton Miffin.

Gomes, A., Gudwin, R. & J. Queiroz. (2003). Towards meaning process in computer from Peircean semiotics. *S.E.E.D. Journal (Semiotics, Evolution, Energy, and Development), 3(2)*, 69-79.

Gruber, T.R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition, 5*, 199-220.

Guerri, C. (2000). Gebaute zeichen: Die Semiotik der Architektur. In Uwe Wirth (Ed.), *Die Welt als Zeichen und Hypothese. Perspectiven des semiotischen Pragmatismus von Charles S. Peirce* (pp. 375-389). Frankfurt: Suhrkamp.

Hagoort, P., Hald, L., Bastiaansen, M. & Petersson, K-M. (2004).Integration of word meaning and world knowledge in language comprehension. *Science, 304*, 438-441.

Harnad, S. (1987). *Categorical perception: the groundwork of cognition*. Cambridge: Cambridge University Press.

Hudson, R.A. (1984). *Word Grammar*. Cambridge: Basil Blackwell Inc.

Lizska, J.J. (1996). A General Introduction to the Semeiotic of Charles Sander Peirce. Bloomington: Indiana University Press.

Paul, W. (1984). On heads versus tapes. *Theoretical Computer Science, 28*, 1-12.

Peirce, C.S. (1931). *Collected Papers of Charles Sanders Peirce*. Cambridge: Harvard University Press.

Pribram, K.H. (1971). Languages of the Brain, experimental paradoxes and principles in neuropsychology. New York: Wadsworth.

Prueitt, P.S. (1995). A Theory of Process Compartments in Biological and Ecological Systems. *Proceedings of IEEE Workshop on Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems*. Monterey, Ca, USA. Organizers: J. Albus et.al.

Sarbo, J.J. (in press). Peircean proto-signs. AIP Conference Proceedings of *Computing Anticipatory System.*

Sarbo, J.J. & Farkas, J.I. (2001). A Peircean ontology of language. In H. Delugach and G. Stumme (Eds.), *ICCS'2001, volume 2120 of LNAI* (pp. 1-14). Stanford (CA): Springer-Verlag.

Sarbo, J.J. & Farkas, J.I. (2002). A linearly complex model for knowledge representation. In U.Priss and D. Corbett (Eds.), *Conceptual Structures: Integration and Interfaces ICCS'2002), volume 2193 of LNAI* (pp. 20-33). Borovets (Bulgaria): Springer-Verlag.

Sarbo, J.J. & Farkas, J.I. (2004). Towards a theory of meaning extraction. In H. D. Pfeiffer, H. Delugach, and K. E. Wolff (Eds.), *Proceedings of ICCS'04* (pp. 55-68). Hunstville (Alabama, USA): Shaker-Verlag.

[1] We acknowledge that a distinction can be made between natural and designed languages like Esperanto. The origin of the language however is of no consequence for our present concerns.

[2] A reference to Peirce (1931) is given by volume and paragraph, separated by a point.

[3] Peircean scholars also use the term *semeiotic*.

[4] The three concepts of X-bar theory, which are denoted by X, X' and X", are a representation of a lexical category, a relation and a phrase, respectively.

[5] The importance of similarity (comparison) in cognitive theory is also emphasized by Goldstone & Barsalou, (1998).

[6] This whole process is beyond our direct conscious control.

[7] Here and in later diagrams a '¬' is denoted by a '~' symbol.

[8] A and $\neg A$ (but also B and $\neg B$) arise due to the same input trigger, indicating that the two signs are *not* independent.

[9] As A and B are commonly considered as logical variables, the separate representation of the input qualia contains both variables. The difference between their meaning as co-existent and co-occurrent is expressed by means of the '+' and '$*$' operators respectively.

[10] For a detailed treatment see Liszka 1996.

[11] We gladly acknowledge that the term proto-sign has been suggested by Gary Richmond.

[12] Conform our assumption that all interaction is between state and effect, the constituents of a type (3) relation show an analogous difference.

[13] Notice that (1) allows only a single interpretation, (2) provides two and (3) can be expanded in three meanings, that differ from each other in the question which one of the qualia has a dominant function in the relation (either the one, or the other, or both).

[14] Both models depart from a triadic definition of signs, but the difference in goals served puts a different emphasis on the properties of signs. We maintain that full understanding of semiosis is only possible when the different perspectives are combined.

[15] Cf. the phenomenon of embedding in sect. 'Towards a formal model'.

[16] Please note that in this analysis the Peircean terms are used as pointers to the status of a language symbol in the process of parsing.

[17] We assume non-determinism is implemented by backtracking.

[18] In virtue of the sequential character of language processing, also adjectives, adverbs, etc. are considered as appearing effects and treated as *B*-type qualia.

[19] Essentially such a grammar consists of rules, which are instances of the syntactic relations - *sorting*, *abstraction*, *complementation* and *predication* - as rule schemas. Schema instantiation is lexically driven by the defined combinatory properties of the syntactic qualia.

[20] The neighborhood relation is indicated by horizontal lines in fig. 3, sect. 'Logical analysis'.