Consciousness at the Cutting Edge between World and Brain

The Interaction between Perception, Imagination, Thought, and Language

Renate Bartsch

Preface

This book consists of three parts, which explore consciousness from different angles: First, from the philosophical point of view, showing the central tasks of consciousness for representation, intentionality, and for designing and following rules, as they have to be performed in order for language, understanding, and especially denotational semantics to be possible; second, from the perspective of connectionist and neuro-cognitive, especially neuro-dynamical approaches, which model at least partially how, under a large-scale view, the brain might work in achieving the central tasks of consciousness; and third, from a view which combines the philosophical with the neuro-dynamical approach in order to find and determine a standpoint about some disputed questions in consciousness research.

The first part consists of chapter 1, which deals with the function consciousness has, mainly with respect to language, thinking, intentionality, and the possibility to devise rules and norms and to learn and to follow these. The central task performed by consciousness is entertaining representations such that we can evaluate these under different points of view. An important strain of argumentation in this chapter is the difference between causal semantics and denotational semantics, and what role consciousness plays there in order to make intentionality of semantics possible. A further point of discussion is the nature of rules and norms and how they function in language and in the linguistic reconstruction of language. Furthermore two standpoints will be argued for, namely that consciousness is possible without language, and that free will is possible because of the special contribution of consciousness, which consists in providing representations of situations and actions such that comparison and judgement is possible.

The second part consists of chapter 2 and 3, which explore the contributions dynamic conceptual systems and connectionist models can make to explaining the nature and function of representations. Chapter 2 elaborates the possibilities of dynamic conceptual semantics and the possibilities and the shortcomings of plain connectionist models consisting just of input, output, and one hidden map of units. The shortcomings are evident with respect to modelling contiguity relationships, classification, and the understanding of basic sentences. Chapter 3 is the central chapter of the book, which shows how an architecture of connectionist maps with circuits of activation in principle works as a model for perception, imagination, and for understanding situations and basic sentences. The representations are percepts or imaginations of situations and of basic sentence inscriptions, i.e. of sentence utterances or written instances. Central is the notion of an episodic map on which activation circuits involving conceptual maps and sensorial fields get expressed in the form of the representations which constitute consciousness. The material basis of episodic maps are the primary sensorial and pre-motoric fields, emotional fields, and also proprioceptic fields, which express feelings concerning our own body. Sensorial fields of the different modalities, or sensorial maps, receive input from the respective sensors. Pre-motoric maps form gestures for realising the motoric output. These maps have connections to higher order maps, conceptual maps, in which groups of neurones indicate when the system has classified and ordered input with respect to previous input under similarity or contiguity relationships. From the higher maps activation is sent back to the primary fields. Such circuits of activation, namely those that hit the primary sensorial, emotional, and pre-motoric fields, receive an expression in consciousness. They have some short term stability in that the firing of the neurones involved is co-ordinated in a certain oscillation. By hitting the primary fields in this way the phenomenal qualities and forms are brought about which constitute representations of situations, objects, and linguistic utterances. These episodes are the conscious expressions of the primary fields in their interaction with the conceptual maps. The role of episodic and conceptual maps in understanding situations and sentences is explicated in the architecture of maps, in which smaller and larger circuits form the constituent structure of our perception and imagination of situations and utterances of basic sentences. Special attention is given to the possibility of integrating an episodic memory into the architecture, such that current remembering consists in creating episodes on the episodic maps, organised from out the episodic memory map in interaction with conceptual maps. In chapter 4 the ability to evaluate representations, and herewith the relationship between understanding and interpretation (in the modeltheoretic sense), is elaborated on as a special capacity of consciousness. Thinking is seen as manipulating representations on episodic maps, with constraints given by the control through evaluation. Thinking and judging is discussed with respect to the question of whether all thinking and judging requires consciousness, and whether all thinking has to be in language.

The third part of the book consists of 4 discussions of controversial questions in consciousness research. The issues are whether consciousness is an internal monitoring device of brain states, or rather a monitoring of the external, whether all conscious states involve thought or judgement, whether there are different kinds o consciousness, and whether there is a one-one correspondence between (a certain kind of) brain states and conscious states. The different arguments are discussed and a standpoint is taken, derived from the evaluation of the arguments and the position on consciousness developed in this book, namely that consciousness is a product of the episodic maps, i.e. of primary sensorial fields in their interaction with conceptual maps, and that therefore all consciousness consists of episodes in the form of representations, with or without their evaluations.

Table of contents

Chapter 1

Consciousness and intentionality in perception, semantics, representations, and rules 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 1

Consciousness and intentionality in perception, semantics, representations, and rules

This chapter discusses the role of consciousness in the distinctions between reception and perception, between a purely causal and a referential or denotational semantics, and between linguistic ability and linguistic analysis, including representations and rules. The first two topics are treated by designing several thought experiments based on the phenomenon of blindsight. It is argued that reception, causal semantics, and linguistic ability do not require consciousness, while a denotational semantics, a notion of truth and reality, linguistic analysis, forming representations and rules, and following these require consciousness and imagination, like any design activity does. They presuppose a linguistic or a picturing medium in which they are formed. The medium is interpreted, i.e. connected to the world, via a neural network background established in training as our linguistic ability, which does not contain symbolic or picturing representations. Rather it functions as a system of dispositions in the ability to interpret pictorial and linguistic representations and rules. Dispositions towards the production of certain neural activation patterns are implicit in the connectivity between neurones. They are not language-like or picture-like representations. It is further pointed out that rules and representations can only indirectly function in changing or forming linguistic ability by serving in consciously constructing series of examples, which in learning processes can be a basis, a training set, on which linguistic abilities then are formed or reformed.

1. Introduction: The case of blindsight and the episodic character of consciousness

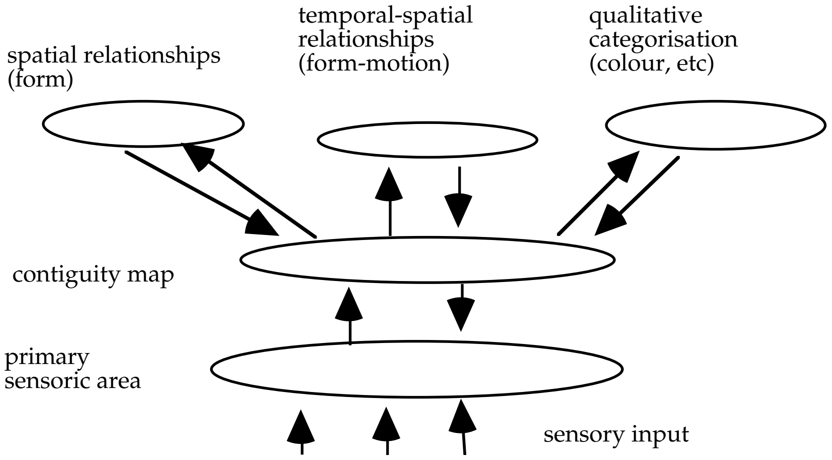
It is generally agreed upon that experiences have an underlying neuro-physiological basis, though it remains a problem what kind of relationship can possibly exist between two ontologically so different realms as, on the one hand, neuro-physiological states and, on the other, phenomena of consciousness. Physiological states and processes in the brain are necessary for experience to occur, but it is not the case that every such state or process has a conscious expression, since many such states are not accompanied by experiential awareness in consciousness. What kinds of features of such states or processes or what relationships between these result in consciousness is not yet known. With respect to states and processes that usually go together with awareness, there can be, due to an accident, a missing link in the causal relationship between reception of data submitted via the senses and the brain states or activities that express themselves in perceptual experience, such that we are not consciously aware of the fact that our body correctly measures, for example, the colour differences which are present before us. Without awareness we still might react bodily to colour and visual movement and form registrations in our brain. Dennett (1991) and Zeki (1992) refer to research about blindsight by Weiskrantz (1986), where the V1-region of the visual cortex does not function properly due to a lesion. Still, patients with this kind of lesions, who report seeing nothing in the blindsight field, registrate differences of light frequencies and motion in this field, though they do not see the colours or motion. Their ability to correctly assent or dissent to judgements of others about what colours, forms, and movements are presented before them is not perfect, but they make significantly better than random "guesses" about these, when these guesses are elicited by presenting to them in their blind field a choice, say between two forms X and O. Zeki (1992: 48) characterises their situation as "Blindsight patients are people who see, but do not understand. Because they are unaware about what they have seen, they have not acquired any knowledge." He suggests that the V1-area might play a role in pre-processing incoming visual data, transferring them through V2, which probably preserves the contiguity order of the input and sends the form, colour, and motion aspects through to the specialised categorising areas V3, V4, V5, and receives them back in an analysed form, recurrently after having been processed in these specialised areas of the visual cortex, which deal with form, colour, and motion, whereby V4 in majority contains cells involved in processing light wave length but also cells that respond to aspects of form.

The phenomenon of blindsight, radicalised in extent and form in philosophical thought experiments, has been used in philosophical argumentation about the nature of consciousness. Vision (1998) criticises Dennett's (1991) claim that blindsight shows that phenomenal consciousness is not central to seeing and that by training and accepting the available reactions as information about the surroundings that fall within the blindsight area, consciousness about the colour, form and motion properties of these surroundings could be attested, although the phenomenal content of these is not available. Without phenomenal content the person might thus have access to visual information and be conscious of this information. This position, which Vision (1998) calls 'doxasticism', denies that vision need be phenomenal. Of course, this position implies the claim that a robot with uptake of visual or other information about its surroundings can be claimed to be conscious of its surroundings. The criticism of the claim that the blindsight person sees or has visual experience about the blindsight area stresses the fact that this person has consciousness merely about experiencing certain non-visual reactions about which others explain to him that they must be caused by what can be seen in the area that falls within his blind field. From the information given by others he can infer by induction that these kind of reactions are indicative of certain colours, forms, or motions. Having this information involves consciousness about other kinds of perceptual input and about certain steps of inference via linguistically supplied information, but it is not seeing that there is a certain object in the blindsight field. Vision (1998) also criticises McGinn (1991), who takes blindsight to indicate that there must be deeper layers in consciousness that are not accessible to introspection. He argues that such a conclusion does not follow from the blindsight phenomenon. Also animals with a blind field can evade obstacles that are located within the blind field. There is no intervening unconscious belief in the blindsight person that he has seen that there is, for example, a car in front of him. The motoric reactions do not require such a belief but can directly be linked to certain input without forming such a belief. The blindsight person has no disposition to say that there is a car in front of him. We can add that from Weiskrantz (1986) case studies we might expect that probably he will be prompted to answer 'Yes' when asked whether there is a car in front of him, but he does not know why he gives that answer. To take this as evidence of a deeper layer of consciousness, which includes hidden beliefs, goes too far, because also here stabilised connections between brain states from visual input processing to motoric dispositions and behaviour cause the right linguistic reaction, although the blindsight person does not know why he reacts in this way. No consciousness about the reason for answering 'Yes' to the question is involved. If we call this a phenomenon of deeper or hidden consciousness we must take all conditioned reflexes between input into the body and reactions to them as evidence of deeper consciousness. This would inflate the notion of consciousness such that it becomes totally useless for discriminating mental phenomena from purely bodily phenomena.

An auditory case of unconscious awareness is that we do not hear sounds below the normal bass-range though we react towards these. We feel insecure and anxious without knowing that this is due to our body experiencing low frequency sounds above certain decibels. Based on these facts of unconscious awareness, let us engage in the following thought experiments, which have the purpose of getting some clearer idea about the role of consciousness in linguistic usage and thinking. We cannot say what consciousness is in reductionist terms. Rather we experience it as the basic fact of life, and we can try to understand it by way of clarifying its role in language use and thought.

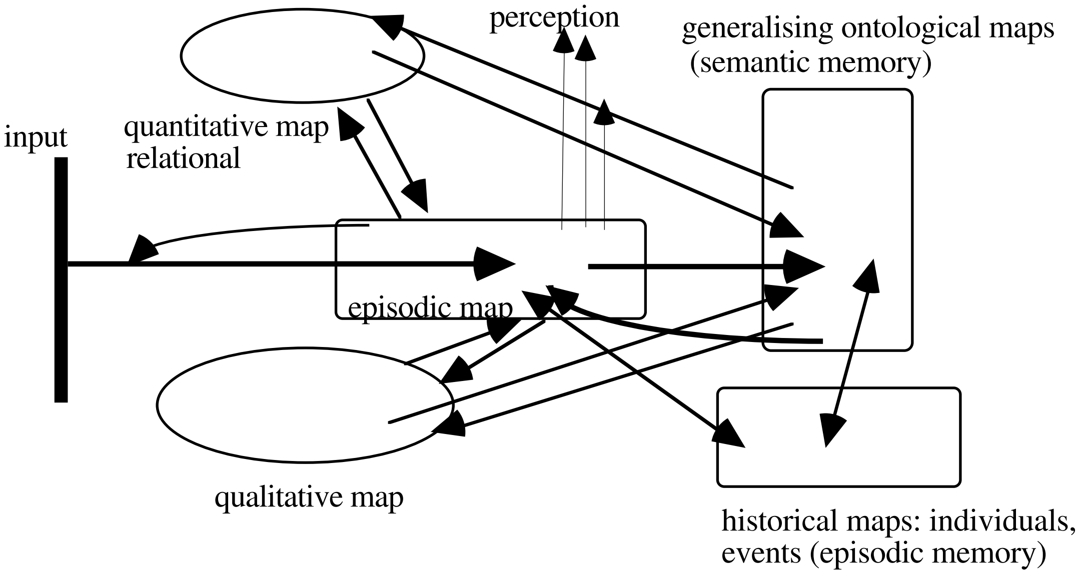
It seems plausible that consciousness about data requires some basic categorisation and ordering of the data under several aspects, i.e. linking them to other similar data by partial identity of basic features and relationships under certain perspectives, and connecting them in contiguity relationships. Awareness of the data in perception, then, can be understood as a selection from the categorised previous input under relevance and attention, and hereby getting directed towards certain aspects. Episodic memory may consist in integrating the variously categorised data, percepts, into world- or situation-maps under space-time, causal, means-ends, and other relationships, which relate objects to each other or to actions and events, or to interests of the perceiving subject. These are kinds of contiguity relationships. Over these contiguity-maps a second layer of categorisations, namely generalisations by feature similarity and relational similarity could be organised by mapping episodes analysed on the first level onto maps in which generalisations take place. Processes of this kind, involving association by contiguity and categorisation by similarity, can be modelled in connectionist models. These second level maps organised by similarity under certain perspectives, furtheron serve for categorisations of incoming data and are thus used for filtering and transforming the incoming data such that they, in the forms so processed, can be sent back onto the pairs of contiguity preserving and primary sensoric maps, which hereby become episodic maps. From the episodic maps episodes emerge in consciousness. These recurrent ways in the whole construction of the neuro-physiological background of our conscious data require circuits of fibres, such that both kinds of maps, the contiguity preserving maps and the similarity preserving maps, i.e. the categorising and ontologically generalising maps, are linked sufficiently to each other and to the primary sensoric maps, which are hit first by the effects of the sensory data.

The architecture of visual episodes



The episodic maps seem to be strongly related to consciousness, while the generalising ones, the concept representing ones, are related to consciousness merely indirectly via the episodic maps. We can, for example, not imagine a concept. Rather we only can perceive or imagine series of situations that exemplify the concept, which may originate from a stimulus to the implicit episodic memory and are constructed in imagination analogously to perceived situations of the episodic maps. Somehow, the binding of the different aspects of analysis, the contiguity and the similarity aspects, should occur in the episodic maps, where the characterisation and identification of an episode, situation or object take place. Characterisation by categorisation through connection to generalising maps, and identification by space-time and other contiguity relationships are united in the representation of an episode, situation or individual. Episodic maps have to register temporal and other relational properties.

A connectionist architecture of the following kind seems to be appropriate for explaining why consciousness is mainly episodic:



Sensoric and motoric input is being processed in maps specialised for qualitative and quantitative categorising these kinds of input. It is received there after going through the base of the episodic map, consisting of a primary sensoric map and a contiguity preserving map, which provide for the qualitative and the quantitative phenomenal material represented in space and time relationships. The current input vector from the sensory units is concatenated by recurrent fibres with a vector from the episodic map representing the previous input on the primary sensoric map, which forms the base of the episodic map. This repeated concatenation establishes the temporal ordering of incoming data on the episodic map by temporarily re-activating connections, and possibly causing a geometrical closeness of the received input on the episodic map. The episode thus gets a marked place in a contiguity net for ordering episodes. The episodic activation is the result of the sensory input and the analysis of the input on the qualitatively and quantitatively categorising maps. This parallel episodic contiguity and categorising activation is forwarded to the generalising maps for being further categorised in the coarse-grainedness we need for orientation in our surroundings. This amounts to an analysis of the input situation in our established ontology. From there the activation goes back by recurrent fibres to the primary modality specific episodic maps, which hereby, together in parallel and possibly also in interaction, constitute the episodic map of situations, which are understood in terms of our historical entities and our normal ontological categorisations.

Quantitative maps process object-internal relationships amounting to form, and process object-external relationships representing relative object positions and movement. These maps are connected to the episodic maps and to the generalising ontological map by circuits of fibres. On the episodic map we then get activation patterns which for us have their expression in terms of objects and situations, events and actions, as we experience them in consciousness, and especially in perception. They are presented, remembered, or imagined.

Such a model in principle explains why we can represent our concepts merely by series of episodes, i.e. real or imagined situations and objects. And, on the other hand, it explains why we can remember something consciously only when it has been categorised. Only then we can understand a situation or an object as different from others, only then we can remember it as a certain situation or object, and only then we can locate a situation or an object with respect to others, i.e. identify it and understand it as being in the world. There cannot be an episodic memory without some degree of categorisation. And as far as categorisation is a classification of individuals and events, and not just of situational impressions, the normal ontology of general concepts presupposes historical concepts, namely concepts of individuals and historical events, which have become part of the episodic memory. In chapter 3, the connectionist architecture of episodic and conceptual maps will be elaborated in more detail. In 3.4, the nature of the episodic memory will be elaborated as consisting of historically individuating contiguity patterns of indices with circuit connections to the characterising conceptual, including relational, categorisations.

Consciousness is restricted to perception, imagination, and episodic recollection. Even in so-called abstract thinking, sentence inscriptions or utterances or imaginations of these are the 'episodes' we deal with consciously. Episodes are perceived or imagined and they are characterised via the connections of the episodic maps to the categorising maps. They are identified by being represented on an episodic map by indicating their place in space and time and in other contiguity relationships with other episodes or situations and individuals that participate in them. In blindsight cases the primary sensoric area of vision is partly damaged. Since it is the basis of the visual episodic map, no representation and especially no percept can be formed there, which would have a conscious expression as a perceived situation. Representation on episodic maps makes use of the sensor related capacities of their sensoric base areas, as it also makes use of the motor related capacities of pre-motoric maps. The neurones dedicated to certain effects caused from the sensors must, at the same time, provide the phenomenal potential of the sensoric areas. In cases of blindsight, this phenomenal capacity is missing.

In order to shed some light on the role of consciousness in language use I now want to clarify the difference between a merely causal semantics and a denotational, i.e. referential-intentional semantics, and show what role consciousness plays for the latter kind of semantics. Causal semantics is part of denotational semantics in the sense that it is the causal background of our semantic knowledge. This background is causally established in learning and consists of stimulation by sensory input, neuro-physiological reactions on the input, and possible motoric output. In what follows I shall show the difference between denotational and causal semantics with respect to the role consciousness plays for semantics. I shall do this by designing thought experiments based on the blindsight cases.

2. Consciousness is necessary for denotational semantics and following rules.

In the case of blindsight, a person reporting not to see anything can react to things located in his blind visual field. With high probability of success he is able to fetch things before him, when asked to do so, and he can make good guesses about what was placed in front of him, when he is confronted with making a choice between suggestions about what objects are possibly there. However, he has no visual experience about these objects. The conclusion is that neural visual registration must have become directly connected, in a previous process of learning, to the neural registration of linguistic expressions as well as to the motor capacities. Missing is, in these cases, the visual experience, i.e. actual consciousness or the awareness of seeing something. This means that part of the mind is missing, if we understand a mind as the whole of an individuals current beliefs, feelings, and attitudes towards surroundings and actions, including the operations by which the individual makes connections between these. The following thought experiments, which radicalise the blindsight case, make no claim of being realistic; they merely serve the purpose of philosophical argumentation for the distinction between causal and intentional denotative semantics, between having dispositions and learning and following rules, and generally for the distinction between unconscious processes in the brain and their conscious results, which constitute our mind.

The first thought experiment:

Imagine that somebody has a defect causing complete blindsight from birth on. In such a case he cannot have instances of colour concepts and visual form concepts in his mind. They cannot play a role in his cognitive capacities, though he might be trained to show certain bodily responses to visual input. For example, lifting his left arm might, in a series of training examples, get indirectly associated to showing him instances of a certain colour. This might be done by telling him to raise his left arm, when we see the colour presented in front of him. Raising his left arm then not only gets linked to our telling him so, but also to the accompanying colour input he cannot consciously experience, but which nevertheless becomes associated by neural connection strength with the effects of the colour reception in his brain. He cannot see the colours, i.e. does not experience their effects on him, except that he somehow suddenly feels compelled to do certain things, namely to lift his left arm, or to ask his trainer "Why don't you tell me to lift my left arm? I feel like doing this." These are the occasions when we show him certain colour samples. But he does not know that we do so. Though his behaviour has become conditioned to these colours, we cannot say that he has learned a rule to the effect that he has to react in certain ways to certain colours. He cannot explain his behaviour as following a rule, because he never can state the rule; and he also can never follow that rule because he cannot perceive whether the conditions are fulfilled which require the reactions. He just feels at certain moments, beyond his control, the disposition towards these reactions. If there are no reasons against, he simply will react in the way he is trained to do. But we cannot say, he will react in the way he has learnt in the sense that he has learnt that he ought to behave in a certain way when a certain condition pertains. He also cannot experience that he shows a regular reaction towards certain colours because he cannot experience one of the two things that are connected in the associative training. Thus we conclude that following rules requires awareness of perceptual input and consciousness of a regularity, though a robot without consciousness can produce behaviour that can be described by rules.

An interesting question now is whether we can say that our blindsight person, in some sense, knows the colours which he has been trained to react to.

Second thought experiment:

He might learn to say that a certain colour or form is in front of him if he feels the bodily reaction which the experimenter induced in association with the colour in the training set. Instead of learning the name as the name of the colour he learns the name as a name of the type of his motoric reaction. He can be told that when he feels an urge to realise the satisfaction conditions for the name, something we call colour X is in front of him. When we then tell him that at a certain place somewhere away there is colour X, he understands this in his way: At that place he will feel that he has to lift his left arm. He nevertheless does not feel, i.e. experience, anything like colours. He can learn, to some extent, to make and to verify colour statements, and form statements if he also would be trained for these in the way described. He can by himself determine truth values for statements about colours and forms in his surrounding. In order to find out whether a certain thing which he can touch has a certain colour, or whether something else behind it or next to it has the colour, he has to perform experiments in the way of removing it and putting it at other places in order to see whether he still feels disposed to reaction X. This is rather troublesome, but in principle he would be able to learn to make the right statements about the colours of things, and also to state what forms things have without touching the things. For him these would not be colour- or form statements, but they would work pretty well in the social co-ordination of judgements. We could say that he knows in some way what colours things have, and that he knows in another way what forms things have because he can also classify them by touching, besides classifying them via his typical causal reactions to the visual input he was trained to, and of which he is unaware. He cannot see the real thing but he can classify it by symptoms, which are his kinds of reactions to received and partly processed, but unperceived, visual stimuli. We also can think of our own colour perceptions as symptoms of certain light frequencies, although they are based on innate dispositions, while the arm movement is learned.

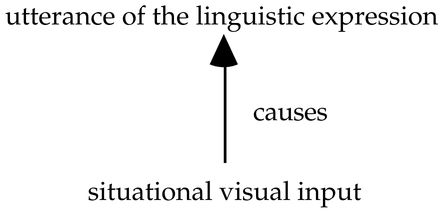
Our question now is whether such a totally blindsighted person has a semantics for colour statements or form statements. He certainly has some causal semantics, since he assents or dissents to colour statements, according to the training he has gone through. This is a causal semantics we also find in neural networks trained towards perception patterns. These nets can learn to connect these patterns with expressions by building up associations between activation patterns of the object perceptions and the expression perceptions. They thus can learn to classify input by attaching names to it. As long as a person is not aware of any bodily sensations that occur to him in cases were we would see colours, he just would be caused to endorse statements about things having certain colours as we do, but he would not be aware of what causes him to assent. He would never be able to learn a rule that connects certain colour perceptions with certain colour labels, though he would learn to apply colour labels correctly. He would be able to judge about correctness in the sense that he could detect a difference between cases of correct and of incorrect applications of the terms by testing them against his dispositions towards certain arm movements in these cases. He remembers approved situations of correct application of terms in opposition to experienced or imagined situations of wrong application, in which he did not feel compelled to lift his arm. The denotations of our colour expressions for him would then be certain arm movements. Such a person would have a criterion of correctness of term application, but for him it would be correctness of applying movement terms. He would have a denotational arm movement semantics and not a denotational colour semantics. Rather we can attest him a merely causal semantics for colour terms. A denotational semantics, namely, requires consciousness about criteria of correctness, the application conditions of the labels, and in our example, the conscious awareness of colours in perception, as that to which the ascription of the labels is conditioned to. For our blindsighted person the denotational semantics of our colour terms is rather a denotational semantics of making certain arm movements connected with certain terms, and about this he could make correctness statements.

Third thought experiment:

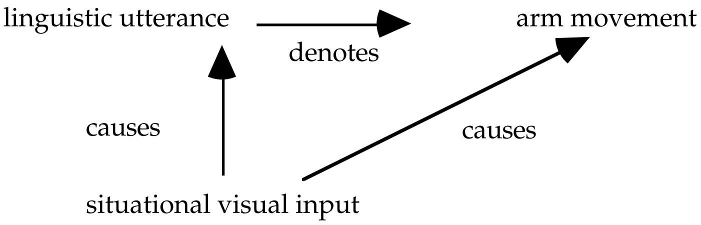
If the completely blindsighted person had not learnt to make certain arm movements in satisfaction situations for certain colour terms, but would have learned the colour terms previous to his total blindsightedness, his semantics for our colour terms would now simply be a causal one, especially if he cannot anymore remember colours in imagination. - And this is very likely because recollection and imagination of qualitative experiences probably will make use of the same neuro-physiological processes that give sensory reception and processing an expression as a certain perception in consciousness. - Whenever he would try to imagine what the difference is between a situation of correct application and of incorrect application of a certain colour term, he would just register the difference between a situation in which he feels compelled and one in which he does not feel compelled to utter the respective term. But he could not correlate his feeling compelled with a certain aspect of the situation at hand, as we can, when we relate our utterance of *red* to something red in a situation. For us, having a concept which is expressed by a term, is not just feeling compelled to utter the term in some situations judged publicly as appropriate. Rather having a linguistically expressed concept implies to relate the term to a certain aspect of the situations in which we feel compelled to use it. To distinguish aspects in situations and to realise how situations are built up from such aspects requires perceptual awareness, i.e. consciousness about the situations, and not just feeling compelled to use certain terms. Consciousness about the relationship of reference and satisfaction, which is established between aspects of a situation and the linguistic expression used for these aspects, is typical for having a denotational semantics. We intend the relationship to hold for the expression to be true or satisfied.

Denotational semantics includes a conscious realisation of the correspondence between the conscious results of two causal processes, the image or utterance of the linguistic expression caused by the situational, here visual, input, and the percept or image of the visual input caused by the situational input. The completely blindsighted person trained to make or select the appropriate linguistic utterance, the completely blindsighted person trained to make the appropriate arm movement and to make or select the appropriate utterance, and the normal person, who has a denotational colour semantics can be represented and contrasted by the following three schemata:

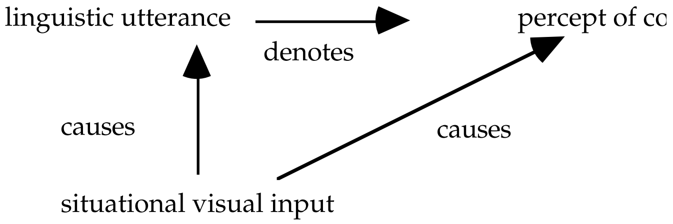
1. The completely blindsighted person without arm movement
(after the training process):



2. The completely blindsighted person with arm movement
(after the training process):



3. The normal person (after the training process):



Against a current version of Wittgenstein’s anti-private-language argument, I argue that an individual, all by himself, can make correctness judgements, if he is conscious about applicability conditions of a term in the sense that he simply remembers, imagines, or refers to aspects of situations of application of a term. These then are present to him in consciousness. Such acts of remembering or imagining aspects of situations of use for a term, or referring to some such situations outside, are conscious acts of control with regard to coherence and regularity of the use of a term, additional to just feeling compelled to use the term. The several cases of feeling compelled to use a term are compared to each other with respect to a regularity under the perspective or frame provided by a certain kind of practice. Hereby, the use of a term is placed within sequences of actual, remembered, and imagined situations, which have to show some convergence towards a regularity, i.e. an internal similarity or aspect of the situations that approaches what would be a set of the applicability conditions, if they could be formulated. In such a procedure of comparing situations of use, each of us makes correctness judgements, of course, without ever being able to exclude the possibility that one was wrong in the judgement by failing memory. And this possibility cannot be excluded even for the language community, which provides the criterion of correctness for public language. We only can say that we normally are quite reliable in these judgements, measured against the likewise, normally, quite reliable correctness judgements of other speakers. Finally, the relevant reliability is measured not merely by previous and actual consensus with other speakers, but by the undisturbedness of our worldly practices that rely on the regular use of the respective terms. Because of this aim of undisturbedness, regularities in the world have to be taken into account by appropriate regularities in the use of our terms. This undisturbedness is reason for a single language user, all by himself, to be confident in his use of linguistic means, even if no other language users are around to support or correct him. The same holds for a whole linguistic community. The last criterion of correctness of our linguistic usage is the undisturbedness of our practices and the experience of regularities in our surroundings in correspondence with regularities in our linguistic utterances.

A blindsighted person may also feel assured about the reliability of his use of our expressions of visual perception and about his correctness judgements, if he has been trained to the additional arm movements, or something similar. But without that training for some associated movements or other behaviour he can be consciously aware of, he would not have any means to understand what correctness of his use of the visual perception terms amounts to, other than that his judgements agree with those of others. He would, though, realise undisturbed practices. When he feels compelled to utter a certain expression, say *a tree before me*, he would at the same time feel compelled to take a detour and would be pretty safe in not bumping into the object located on his path. But all that does not provide him with the direct awareness that he uses the term in accordance with certain conditions of application of which he could become aware by the regularity in imagining situations. When he is sitting he would not feel compelled to evade the tree, only when he is walking, and it is on his path. And when he is just sitting there, he might not feel compelled to utter the expression he has been trained to, because he is occupied by something else, or has no urge to communicate. Thus the situations of application of the term he can be aware of, namely taking a detour or feeling compelled to a certain utterance, are much more dependent on situational factors and they do not show a common regularity, though they sometimes factually overlap temporally, without being similar to each other in any way: the situation of feeling compelled to take a detour, and the situation of feeling compelled to utter a certain expression overlap in time but are totally unrelated to each other. For a person with visual awareness there would be one situation which he can see, and with it certain actions are connected that might be appropriate, such as uttering a descriptive term and evading the obstacle seen ahead. A similar case is the following: When we ride our car while thinking about quite different things, we still are disposed to evade obstacles without being consciously aware of them. We have not seen them, though we have got visual input and automatically processed it such that the appropriate connections to the motoric reactions are made sufficiently.

We can compare the situation of the totally blindsighted person having merely a causal semantics for colour terms with Churchland's (1986) model of the connection between the crab's vision and its arm movements in catching a prey. There is a direct transformation between the sensory topographic map and the motor topographic map. The visual input causes a pattern on the sensory map, which is transduced to the motor map, whereby it gets transformed such that a motoric pattern is activated which causes the arms of the crab to reach for the prey it has seen in the sense that it has received visual input. We do not know whether the crab has really experienced seeing the prey. - That would, according to the present theory (cf. chap. 3), mean that the crab would experience episodes, and this would require that there are conceptual maps, which are in circuit interaction with the primary sensory areas. - The whole crab situation suggests a simple causal semantics, whereby no conscious awareness in the sense of having a conscious representation or picture of the prey is required. The crab's correlation between visual reception and motoric activity might just function like the correlation between unconscious colour effects in the brain and the use certain colour or form terms by a totally blindsight person, except that the person consciously perceives the terms he uses, or at least feels being compelled to use the terms.

What is the additional benefit of having available a representation which one can be conscious of? Consciousness is a prerequisite for intentionality and for learning and applying rules, instead of just showing regular behaviour achieved by training. The observer can see that the behaviour is regular with respect to colours shown, but the blindsight person himself cannot know whether his behaviour or his feeling of being compelled to do X, utter a certain expression or move his arm, is regular. He misses the comparison with the associated colours. He only can take into account his behaviour or disposition, but he cannot relate to what it is correlated with in a regular way. He cannot understand his behaviour towards colours as regular and thus he has no rule for doing X, for example using certain terms or making certain arm movements. Having available a rule implies understanding our learned behaviour as regular. It also makes it possible to sometimes abstain from the behaviour, since a rule or regular behaviour can be deliberately broken by intentional inhibition of the regular response.

Consciousness further makes it possible to imagine situations and plan behaviour accordingly. Planning requires imagination and intentionality, and imagination presupposes previous perceptional experience of situations, which is conscious, i.e. real perception and not just reception of input with some processing. -- Perception, as the term is used here, includes conscious presentation of what is seen, while reception merely includes visual input and part of its processing without coming to conscious awareness. Intentionality relates the imagined or represented situation to the world in a direction of fit which can be world-to-mind in belief and perception, or mind-to-world in desire and planning (Searle 1983).

In this context, representations should not be taken as syntactic constructions out of basic concepts in the way they are thought of in a mentalistic semantics, like the one of Fodor's (1976) language of thought. Of course, there are linguistic representations in our natural or designed formal languages, but they are not necessarily involved in propositional attitudes like belief and desire, or in perception and action. They only play a role in representations of these in reports. The only representations in our mind are the perceptual presentations and the recollected or construed images of linguistic utterances and inscriptions, of real or imagined situations, and of the relationship of approval between an expression and a situation in which the expression is used correctly. Our linguistically labelled public concepts are equivalence classes of stabilising sequences of growing sets of such satisfaction situations for an expression. Representatives of these equivalence classes are available to every language user. Each representative of such a concept or equivalence class is a sequence or set of situations, subsets of which we can call up in our mind as sequences of examples for the concept. We can come up consciously with such sequences of imagined examples, and we can refer to such sequences of examples in the world. These sequences of examples of our concepts consist of presentations and representations of situations or images. We conclude that our totally blindsighted person has no colour concepts, since he cannot imagine or design examples. Of course, he must have assemblies of neurones activated with the reception of certain colours and forms, which connect his arm movements or the endorsed colour statements to actual colours and forms he does not perceive. Maybe that in our completely blindsighted person some received input can go directly to V2 and the categorising visual areas, bypassing the damaged primary sensoric area V1.

Although our blindsighted person, like a robot, naturally merely has a causal semantics of colour terms, I have pointed out that a denotational semantics can be achieved if the blindsighted person is trained to both, making certain movements for different colours and also using our colour terms. Then he can combine in correctness judgements both external facts he is conscious of, the movements and the use of terms, and achieve a denotational semantics by connecting the terms with the respective movements, although he is not able to use any colour experience as a missing link. He now can imagine situations of false use of a certain term by just imagining a contrasting movement that does not belong to this but to another colour term. But note that these terms in his denotational semantics are in fact not colour terms but movement terms; he is conscious of the fact that he feels compelled to use a certain term, our colour term, in case he feels compelled to a certain movement. The term, as well as the movement, is located in a network of contrasting terms and movements, respectively. The term is correctly or incorrectly used on the background of the sets of contrasts within each network, which correspond to each other. By virtue of these contrasts a term can be judged to be correctly associated with a movement. The point of this second thought experiment was that the person is conscious of his use of expressions in connection with certain movements, of which he is also conscious, and he feels compelled to use a certain term in case he feels compelled to make a certain movement. He had been trained to connect two external things he was conscious about, expressions and movements in situations in which others applied these terms to colours, of which he is not conscious, but to which his movements and use of expressions have been causally connected by a training process. For the causal semantics between expressions and unconscious colour stimulation, and for the causal semantics between movements and unconscious colour stimulation, no consciousness of the visual uptake is necessary, and even using the expressions and making the movements can be unconscious, as we assume it to be in neural nets and robots generally. To have a causal semantics does not require consciousness. Therefore to point out that robots have a sensory uptake which is connected to a motoric output by complex processes of information processing is no argument for assuming consciousness in intelligently programmed computers and robots, as Chalmers (1998) is inclined to do.

On the other hand, for a denotational semantics, which is a semantics based on correctness, especially truth judgements, two consciously available media are necessary, which have to be related in a structure preserving manner. This relationship is intended by the language user. To intend this relationship of reference, satisfaction , and truth means to understand language as language, i.e. as a system of symbols, which stand for something. To understand a language as a means of representation requires to be consciously aware of the signs and of what they represent.

Denotational semantics has been learned by training on external examples one has been conscious of. The structures are defined by identities and oppositions between the cases of expression use on the one side, and by the identities and oppositions between the instances of the corresponding partial extensions or denotations, on the other. In their structural fit to each other, the individual's correctness judgements are founded. In this way he could have a private language with correctness judgements as far as the fit of the structures goes. In fact, such a private language is learnable and teachable via the external examples on which the structures are erected, and thus is in principle a public language. In order to learn a public language, the elements of the structures have to be related to external things and practices about which we are conscious, and to which we refer consciously in our use of expressions. This relation between the two structures and their external aspects is not one between two separate realms, rather the external denotations are the data on which the structures are built by operations of ordering data according to factual contiguity and to similarity under certain perspectives. The relationship between both structured realms of data is induced in establishing the structures on them in association between expressions used in satisfaction situations and situational data present in these (cf. Bartsch 1998). Thus the relationship between both structures is an internal correspondence rather than an external one between two independent settings. In the process of learning and understanding language and world, linguistic expressions and facts satisfying them are necessarily taken to be partially identical in structure. The linguistic expression selects from the set of orderings into which the situation fits, conceptual orderings depending on how far the expression is analysed in the available language. This selection is what is highlighted in a situation when we realise that it supports a certain fact. The situation is then represented under this respect as a situation described by the expression.

What now are the major shortcomings of getting information in a setting of a merely causal semantics, compared to our conscious manner of knowing and perceiving, in which our denotational and intentional semantics is rooted? In order to give an answer to this question let us push our thought experiments somewhat further and compare our imagined blindsighted person with normal perceivers on the one side and robots on the other.

Our totally blindsighted person behaves rationally with respect to colour and form judgements, i.e. draws the same implications we do. But he does not have any visual aesthetic judgements, since he has no colour experiences and visual form experiences, and therefore no visual imagination of situations. The main difference with us is that he has no visual imagination. He may have motoric aesthetic judgements, or aesthetic judgements about certain sentences he feels compelled to generate by some cause unknown to him. The aesthetic judgements may or may not coincide with our judgements about visual experiences, and he might learn that it is culturally appropriate to call certain of his motoric aesthetic experiences beautiful, sublime, or ugly, following our judgements about, for example, a colour ballet. And he might learn the same about his linguistic combinations.

Let us suppose again that our completely blindsighted person had not been conditioned to making motoric reactions to colours, forms, and movements, but merely had been trained to utter and accept certain statements in situations where he receives non-perceived visual stimuli, as described in the third thought experiment. He then would have no visual imagination and also no motoric aesthetic substitute. Because he lacks any such imagination, he never could design anything visual and he could not anticipate situations, except that he could form sets of statements which he knows are coherent in the sense that he has experienced that they belong together somehow because they occurred together before. He thus can only imagine feeling compelled to make certain linguistic combinations. He and others, together, somehow have used them and could use them as descriptions of situations he could not see. His only reactions towards these situations, he is otherwise unaware of, is linguistic, i.e. a disposition to utter certain sentences. The coherence of these sentences can not be felt in any way, except in a structural similarity or analogy to previous discourse. He can, of course, acquire the notion of a situation only via his other senses, as certain units of smell, touch and sound he is conscious of. On this basis he could make designs, using his imagination of previously experienced situations and modifying them. But this is not at issue here.

We can imagine the extreme case in which he can only perceive linguistic utterances and other sounds. A zombie or robot is a further radicalisation of this case, where no perception is available but merely reception of sensory input, processing and behavioural output. A robot has no sense experience at all in which he could consciously register situations. The robot also cannot feel compelled to make certain linguistic expressions, and therefore he could not imagine situations of feeling compelled to produce such expressions. He gets expressions as input but does not perceive them. He just is caused to give expressions as output triggered by a certain input. In fact, he cannot imagine any situation. Rather he just could come up randomly with new linguistic combinations, possibly guided or restricted by previous probability assignments to certain expressions or structures, closeness of expressions in previous examples of discourse, or he could produce expressions by being set to execute rules programmed or hard wired in. Yet a robot with sensors, but without perceptual consciousness, could be constructed to memorise and generalise input and output patterns, for example linguistic output associated to sensory input. Some correlations could be kept in the memory of a controlling device. In this way he even could mimic correctness judgements and conscious comparison between good and bad cases of term application. Such a monitoring device could replace our conscious monitoring and be developed into an ability to give automatic correctness judgements, including truth judgements. This device then could be used to teach others, without knowing what it does.

The correctness judgements of such a controlling and monitoring device are unconscious and unintentional. In fact, they are not real judgements for the monitor, because for him they are not about anything. The controlling and monitoring computer does not know what correctness or truth is, though he 'judges' correctness and truth rightly. But he could not intentionally break his own procedure and accept an incorrect example, knowing that it is incorrect by the standards. He could not decide voluntarily to follow or not to follow his own rules. He does not even know that there are any rules. Still this computer might be an effective teacher for a certain task. Even though he does not have a denotational semantics, he still can produce correctness statements. But they are correctness or truth statements only in the eyes of the program designer and user. The monitoring and controlling device itself does not know what correctness of a task performance amounts to. He blindly registers and checks the task performance of his pupil against his own learned task performance, or against the rules and representations programmed in. A teacher who understands what he does could ask himself whether he himself was quite correct in performing the task. A monitoring and controlling robot could not do that, though he could be programmed to, now and then, repeat a training session for himself on an old or new training set. But he would not know that he trains himself in order to uphold his own ability to task performance and correctness judgements on others' performances. This taking in of a training set can be seen as the causal semantics of the term 'learning task X' or 'retraining task X', like his performance of the control and correction task is the causal semantics for the term 'giving correctness judgements about performance of task X'. Thus he has a second order causal semantics, but this does not make his first order semantics denotational. Our advantage as conscious beings is that we are always ahead of the robot. He only can take the next step of meta-performance if we first program him accordingly. The point is that he does not understand his own behaviour. He does not know what he does. He just performs.

What is missing in a robot is the development of the semantic ontology, by which input stimuli can be synthesised by the power of imagination, as Kant has emphasised. The synthesis of series of situational impressions in space and time gives rise to historical concepts, such as concepts of individual objects, events, and acts, and it gives rise to general concepts. They are comparable to Kant's empirical schemata of classes of objects, events, acts, and abstract objects. In these schemata concepts are not merely sets of features, but these features are aspects of things or situations ordered in forms of space and time by the power of imagination. These aspects are not all explicated like features would be in traditional definitions of concepts. This same power of imagination is part of every act of perception and experience generally. Transposing these insights into today's approaches to cognition we may assume that a bias towards our coarse grained ontology is built into our neural net organisation in the course of evolution. The network then will be able to synthesise incoming data into the evolutionary selected ontological frame in a way which is causally efficient with respect to our behaviour within our surroundings. This all does not really require consciousness. But still, using historical concepts and general concepts in understanding and in recognising objects and categorising them requires consciousness of the unity of an historical object and of the unity of a class as a general concept. We have to recognise parts of space-time slices of an individual while assuming, in space and time, the continuous whole, i.e. the total individual, of which they are parts, and we have to perceive the unity of the class through the examples of the concept. It is the assumption of totality or completeness, although we factually only realise partiality, which is at the root of our notion of reality and truth. These are not part of our experience, but regulative ideas of reason, as Kant and Husserl have characterised them. However, it seems possible that such regulative ideas could be built into a robot as procedures towards coherence and completion of partial general and individual concepts. Such procedures would correspond to our intention towards truth for statements and reality for referents of our partial situational concepts and individual concepts.

Denotational semantics is built on the semantic notion of truth. The following argument shows that self-consciousness is a prerequisite for having such an objective notion of truth. In order to understand sentences about the past, *PAST(p)*, one must be able to think "If I had been a suitable observer in that past time I would have observed that p". In a causal semantics, such as a robot can have, a link can be laid between past experiences and the present in the following way: 'Yesterday it rained' is true if and only if in the memory on the location labelled as 'one day before now' a perceptually caused situational representation connected to the phrase 'it is raining' can be found. But note that this is not sufficient for a realistic notion of truth. It is a purely phenomenological or procedural relationship depending on the perceptive history and the memory of the robot. Rather a counterfactual relationship has to be added, namely "If I would have been a suitable observer the day before the day now, I would have observed that it was raining then", i.e. I would have had a situational perception fitting to the phrase 'it is raining'. Such a counterfactual statement requires consciousness and even self-consciousness. This statement can be generalised over the whole speech community as: If there would have been a suitable observer the day before now, he would have had a situational perception fitting to the phrase 'it is raining'. To have and to handle a notion of reality above mere phenomenal registrations presupposes self-consciousness. The realistic notion of truth requires consciousness about one’s own and others' possible and counterfactual positions as verifiers of a statement. The same is true about the realistic notion of an individual; to have the concept of a real individual, surpassing the knowledge of partial individual concepts, requires consciousness with respect to one’s own and others' possible and counterfactual positions as perceivers of the individual. Understanding simple statements about the past requires self-consciousness and consciousness about oneself as a possible observer among others. The same is true with respect to statements about the present, if they are about some situation at a distant spatial location.

3. Consciousness is necessary for having general knowledge and structural semantics

Structural semantics is a set of semantic networks in which a term is explicated by its place in the network. The semantic network of a term consists of relationships between this term and other terms, which are interpreted as semantic relationships, i.e. as holding between the extensions of the terms in real and possible worlds. These relationships make up the theoretical concept expressed by the term in question. Alternatively a theoretical concept can be expressed by a set of general sentences in which these relationships are formulated as holding between entities that belong to the extensions of the terms in question. A taxonomy, for example, is a special kind of semantic network which expresses the relationships 'is a', i.e. inclusion or subordination, hyponomy, and the relationship of exclusion, i.e. opposition between concepts. Such relationships can also be formulated symbolically in traditional lexical rules, meaning postulates. On the other hand, they can also be found, in a more flexible way, in structures on the sets of activation patterns in connectionist networks, where they are, so to speak, directly contained in the conceptual 'contents' as they appear as activation patterns on the nets of units which can be understood as conceptual maps (cf. chap. 2 and 3). The activation patterns can be taken as 'contents', or better as conceptual indicators, by virtue of their relationships in which they are causally bound to the perceptual input in learning processes. The activation patterns are due to input examples and the adjustment of connection strengths between neural units in learning, i.e. in generalising the examples presented as input. They are directly addressed, i.e. activated, by new perceptual input of similar kind. The causal semantics of activation patterns, encoded implicitly in the connection strengths between units on the neural net and the input units, is established in the learning process and reactivated for each new occurrence of similar input. The activation patterns therefore function as indicators of the concept or category of the instances.

Our structural semantics, established publicly in endorsing general statements and theory building generally, has on the connectionist ‘content’-level a counterpart which is built upon causal semantics while building up this ‘content’ level in learning. This basic structure of inclusion, overlapping, union, and exclusion of potential activation patterns is causally established in learning and does, by itself, not require consciousness. However, the symbolic formulation of relationships between concepts does. It would be an uninterpreted network between terms if the terms were not causally linked with the objects and situations in the world about which the relationships of structural semantics hold. This causal linking has been built up in a regular way as a stabilisation in the distribution of neuronal connection weights during learning processes. The current activation patterns are the results of our dispositions, implicit in the network architecture and the connection strengths, to activate the respective patterns in reaction to certain input. These activation patterns then cause a certain linguistic or other behavioural output. Activation patterns for the perception of linguistic expressions and activation patterns for situational perception are connected through the process of language learning, and in this way we connect the use of a term to situations.

The interpretation of expressions as a relationship between expressions and the world is a conscious one in intentional denotational semantics, which delineates objects and situations which are coarse enough to be the units or chunks in the world we have to reckon with, and which are publicly referred to and communicated about. Thus consciousness and intentionality in perception and action make us select the right size objects and situations to deal with, and our denotational semantics is directed toward these objects, which constitute our world, the world we know. Structural semantics is our conceptual ordering on this level of intentionality, and in this way is a derivative of denotational semantics, which constitutes the interpretation of its symbolic formulations. Only on our intentional level of coarseness of objects, actions, events and situations generally, we have knowledge, and according to this level of coarseness the activation patterns on the connectionist neural level get delineated and partitioned by causal relationships to action routines and to public language use. Causal relationships with these two aspects select and thereby delineate structures of activation patterns by enhancing them in joint stimulation. This happens in a consistent and precisifying way by increasing the connections' strengths built up in the association between conceptual activation patterns, on the one hand, due to perception of our factual surroundings and our actions in these, and, on the other hand, activation patterns due to the perception of linguistic expressions. Linguistic expressions, and also action routines, serve by means of these connections to emphasise, i.e. strengthen the parts of activation patterns which correspond regularly to the use of these expressions or to the performance of the respective actions. Hereby the patterns of publicly relevant objects and situations, as encoded in our public language, get distinguished on the level of 'contents' or conceptual indicators as they appear on the neural networks. These considerations show how our public denotational and intentional semantics and causal semantics depend on each other. The first can only be handled by means of the second, and the second in its specific causal structure depends on the first, in the sense that it is induced by the first.

I shall now show that for acquiring the semantics of certain terms, consciousness of a specific quality of perception is not required. However, there must be some perception of discriminating qualities we are conscious of, in order for a denotational and structural semantics to be possible. We have to be able to be aware of identities and differences necessary for building up structural semantics, or general knowledge. In order to see this, we now change our thought experiment in such a way that certain reactions to unperceived, though well received, visual stimuli are not brought about by training but rather are assumed to be innate.

Fourth thought experiment:

Imagine that someone has experiences connected with colours which are quite different from our experiences. For example, instead of having a colour experience like we have, he feels a certain kind of itching at a certain area of his body if a certain colour is present before him. In certain context types his concepts of certain colours would be structurally equivalent with the ones of other speakers.

But this would not be so in all contexts, especially not all our oppositions between the colours could be made. If, what is very unlikely, there would be a total equivalence in colour judgements between him and us in all satisfaction situations, then we have to say that his concepts are the same as ours. Suppose that with every colour term he has learned to associate a certain kind of itching. This serves for using the colour terms correctly with respect to satisfaction situations, except in cases where the itching is caused by something else than lightwave frequencies. He can report to others about colours in objective situations correctly. Still, the difference of his colour experience from those of others is statable. The subjective situation will be different; the others will say that for them colours do not cause itching anywhere. How far does the equivalence extend? Full equivalence would require that the itching would be located in the visual space. That is not possible. Only if a space can be formed in which different kinds of itching are related according to their location in space, his use of colour terms could be similar to ours. Otherwise he cannot locate colours at the place where they are. He only would know whether some colour is around. To locate them as belonging to certain objects would require the same complicated procedure of moving objects around, which has been described above in the context of the second thought experiment.

In our example, we have identical structural relationships in the conceptual space of colours and of itchings, but the external location, a property of the visual field, is not represented. The itching would have to be located outside the body in space. We would need something like a pain space outside us, in opposition to pain in and on our body. The itchings, though due to stimuli on our skin, have to be noticed by us as if projected outside our body such that we perceive them as properties of itchy objects outside. Instead of objects of different colours we would have objects of different itchiness. If objects outside would itch us by the light they reflect, and by certain qualities of the itchings, their intensity, and a kind of feeling of the angle under which the light reflections hit the skin, we would also get information about the place in space, the distance and direction in which we find the object which itches us by the light waves it emanates. This, of course, requires correlated motoric experiences about reaching out towards a place in space. Note that an auditive space is similar to this kind of construct.

Suppose now that somebody would merely see a certain black-white pattern instead of a colour, and for different colours there would be different patterns. They are placed in visual space corresponding to the colours we find there. Nevertheless, his notion of colour would differ from ours, and this difference can be described. His colour perception would be part of his form perception. He could not distinguish whether indeed there is the form pattern, or whether it is the colour that for him has become this pattern. He therefore would not have the same colour concepts we have. The difference is due to the fact that the colour oppositions and identities would be integrated into the form oppositions and identities and hereby get indistinguishable from these. This would also be so in the example of itchings, if he also would experience other itchings than those caused by colours.

The off-shot of these thought experiments is that perception of other qualities, replacing our visual perceptions, can result in partially identical semantic networks, i.e. in partially identical structural semantics, or alternatively expressed, in partially identical general knowledge. In principle the specific quality of sensory perception does not matter as long as it serves to make the publicly endorsed identities and differences. This is the old structuralist principle which holds not only for language but also for knowledge generally. All our knowledge, as far as it is publicly shared, is structural. Substance or quality matters only as far as it functions in stating identities and differences, and this can be done couched in different kinds of substances and qualities, or different kinds of phenomenal material, in an equivalent way.

Fifth thought experiment:

Let us change our thought experiment and suppose that, for example, somebody would see red as green, blue as yellow, etc. Then he would have the same notion of colour we have, and also his colour concepts could not be distinguished from ours. We would not even come to know that he perceives colours differently. We could not but say that he has the same colour concepts we have. This is due to the fact that all oppositions and identities are the same as ours, and we can communicate about these. However, he could connect colours differently to feelings from what we experience as normal. Suppose that our red is for him identical with our green feeling; then what he calls 'red', would calm him, and our green, which he experiences as our red feeling, would make him nervous or stressed. This we could interpret as a sign that he has a different colour perception from ours. However, we also could explain the difference by supposing that his colour perception is the same as ours, but that the connection to other feelings is different for him, which makes him nervous in sight of a green meadow and makes him relaxed when seeing a room decorated red all over.

The result of this last thought experiment is that by an interchange of colour qualities the structural semantics of colour terms would just remain the same, as, alternatively formulated, would our general knowledge about colours. We would not even detect whether there is such a difference in colour perception between people, and by this very reason logical empiricists have used the statement that different people might have a qualitatively different colour perception as an example of an (empirically) meaningless statement.

4. Consciousness is necessary for an intentional creation and improvement of language

There are many physiological registrations of internal and external processes, reactions on surroundings, and inside and outside effects on our body and its nervous system which we are not consciously aware of because we do not experience them consciously. For these bodily registrations also neural registrations and classifications could be built up, very much comparable to the ones of which we have experience. Of course orientation of our body can be achieved by conditioning to these influences, similar to the orientation of the man who reacted on what he "saw", without seeing it consciously. This leaves us with a lot of regular and irregular behaviour, unaccounted for in terms of rationality and intentionality, because in explanation and understanding we can only take recourse to causes and motives we are conscious of.

We might react regularly to sounds of frequencies beyond those we can be aware of in hearing. This we could detect if we would monitor our behaviour while carrying with us a frequency meter, which we routinely check. An example of unconscious training would be a robot who has learnt to react to his surroundings, or Churchland's (1986) crab, if we assume that such low animals do not have consciousness, and that their reactions are not fully hardwired in to begin with, but require some learning. There is implicit in the trained neural maps a rich structural semantics in terms of inclusion, overlap, union, and exclusion of concept indicating activation patterns, which mostly may not evolve into consciousness, and therefore cannot be expressed in a language which, also by the language user, is understood as a language, i.e. as a symbolic system.

We cannot by ourselves invent symbols for registering unconscious processes that are reactions to properties of our surroundings, but we could be trained to feel compelled to use certain expressions whenever the unconscious stimulation of our body by these external properties occurs. The training would have to be performed by some external teacher who is consciously aware of these external properties as an observer. He can consciously perceive our neural reactions and has devised the terms for these properties. He would call out the term the moment he registers the activated neuronal pattern. After such a training we would be able to use expressions like a robot does who has been trained to certain input- output correlations. We just would not now why we suddenly use these expressions. They therefore would not really be language for us, though they would be part of the observer's or teacher's language.

In order to create, develop, and improve a language we need consciousness of that about which the language is going to be used, and this is something a robot without consciousness cannot do. To design or develop a language presupposes some knowledge that certain things exist and events take place, such that one can make statements about these, and these things can fulfil the statements, i.e. can make them true. It presupposes the intention to use certain objects as signs, namely as symbols for something else which one is also conscious of. This capacity to develop a language which oneself understands as being a language or a symbol system, might be an essential distinction between human and artificial intelligence. In consequence of being trained within an unconscious state, a robot cannot have any knowledge which is a 'know that', but merely has a 'know how'. Even if he would utter expressions of statements, these would be statements, i.e. expressions of knowledge, only from the standpoint of the designer or user of the robot. An expression can only express knowledge if it is understood as being about something, which hereby is intended as that which makes the statement true or false.

We can imagine a computer run by a program in which a procedure has been programmed that randomly selects a term out of a set of possible terms in correspondence with a certain internal activation pattern, which has come up regularly in connection with a sensory input. If similar inputs occur the same term association is made. Another machine can adapt itself to this association by a learning process. For the outside observer it looks as if these machines create a language all by themselves. In fact, these artificial devices for language creation do not know that here a language is created. They use signs for outside objects via the causal semantics developed internally according to the program and they seem to communicate with each other. However, they do not know that they use a language and communicate by means of this language. Only the designer and observer can understand and describe the situation as one of language use and communication. In doing this he sees this situation in line with situations of human language use and communication, even though the machines miss the intentionality and consciousness about what is at hand.

Non-experienced registrations, i.e. reception, can give rise to a non-perceptive and non-cognitive adaptation to the world in terms of internal states and processes. Such adaptation is not knowledge that this or that is the case, or must be done. These registrations, and the connections among them, are part of our 'know how', abilities, patterns of reaction to natural and social conditions. Non-perceptive and non-cognitive orientation based on merely sensory reception and its effects on neuronal connections, which collectively embody dispositions towards certain activations, are not part of our consciousness. We are only aware of its consequences, namely that we, for example, walk pretty safely along a road, and that we can bicycle (cf. Searle 1983 and 1992 about the Background). There is also no reason to assume that this mostly non-perceptively but merely receptively acquired non-cognitive background, or ability to perform certain procedures, can be made conscious by just somehow lifting it to a conscious level. It can be reconstructed scientifically by a neuro-physiological theory about how a system adjusts to its surrounding. Searle, rightly, points out that our background, although being unconscious, is necessary for using any rules and representations, such as linguistic expressions and pictures, because only via the background these can be interpreted, i.e. be related to the world. This happens in an automatic causal way. Our causal semantics is such a know how, an ability which makes understanding and interpretation possible.

Causal semantics is a pre-condition for employing denotational semantics. It has to be acquired in training. As Wittgenstein has argued, internal rules and representations by themselves can not serve as a medium for relating public rules and representations, for example linguistic expressions, to their cases of application. Such internal mental objects would need interpretation themselves in order to be utilisable in interpretation of public expressions. This means that the mental objects must already be understood themselves as examples of the rule or as examples of application of the expression. Likewise consciously comparing new cases of use with recollections of old cases of use of the expression cannot be the solution for the problem of regulating the correct use of the expression. The recollections of the old cases would have to be judged as correct examples beforehand and for this we would already require the criterion of correct use which we try to construct as similarity with these recollected examples. Rather a causal, unconscious relationship established in learning from cases of publicly approved language use, establishes for each of us the link between previous cases of use of an expression and new cases. Causal semantics for each of us is the condition for the possibility of having a denotational semantics. The origin and the reality of denotational semantics is a public affair. It is in fact social, but it can in principle also be public in the weaker sense that it is simply outside us, such that we can be conscious in perception of the two sides of the relationship of interpretation, language and world. Consciousness is not a condition for causal semantics per se, but it is a condition for the specific denotational structures imprinted on it during the process of learning the public language.

Our ‘know how’ capabilities are not processes governed by rules, though the regularities in their results may be described, in idealisation, by rules. Firstly, learning rules as rules and conscious application of rules require data one is aware of. Secondly, representing these processes by context free rules, taking into account the kind of input and output in form of descriptions, proves to be incomplete because symbolisation of input and output cannot take into account all the specifics of these data and their surroundings which might have causal influences on reception and neuronal effects. Therefore rules constructed by the linguistic researcher could at most be idealisations of the relationships between the results of the processes that in reality take place without being governed by rules. There also will always be situations where no action is possible in accordance to the rules constructed. A small deviance in the preconditions of an operation is already sufficient for the break down of an assumed linguistic algorithm. Our behaviour is rather caused by sub-cognitive processes, which, as far as they remain unconscious, can not be understood in terms of rules, a point made by Searle (1992). Thinking of the background as providing basic orientation in our surroundings and as not being a set of rules, makes neuronal networks to be the only presently available models of acquiring basic orientation in a non-conscious way.

5. Linguistic ability is not a generative system of rules

We now address the question whether, and in how far, language is part of the 'know how' background of knowledge and action, and whether, and in what sense, it is part of cognition and knowledge as that what is conscious and thus part of our mind. I shall argue that both is the case: Our language ability is part of the subcognitive background, while language as a set of rules and representations is the result of linguistic analysis and does not describe our language ability in kind, but merely is about structural properties of the results produced by our language faculty. Hereby a system of rules and representations indirectly suggests some structural properties of our language ability as it must be encoded in our language faculty. However, these structural properties are not the results of hidden rules that would guide our language ability. Rather they are idealising descriptions of globally emerging conscious phenomena that come about by cumulative effects of local micro-processes of quite different kind.

We speak correctly without being able to cite the rules which might have regulated our speech; we use terms with meaning, i.e. concepts, without being able to define these concepts. We, apparently, do not have representations in our mind of external meanings such as concepts in the traditional sense, still advocated in mentalistic positions. We are merely able to present, identify, and imagine examples or cases of use of respective terms, or, even without using terms, we are able to arrange situations in classes, in alignments according to similarity, and in orders of contiguity. This is an ability, a know how, as is rightly stressed by Wittgenstein (1960), Putnam (1981), and Peacocke (1992). A concept is a certain ability, according to these philosophers, not a representation of something common to all the cases of application, like the mentalistic notion of a concept would have it. There are no concepts or representations of meaning to inspect, when we attempt to analyse a concept and define it. Rather we can only scan through what we are consciously aware of, namely real and imagined situations of language use, which are the external proofs or products of our linguistic ability, and which, for us, constitute what linguistic ability means. In our scientific endeavour of linguistic analysis we can only inductively and hypothetically conclude from the examples to some definition of their kind, which is a certain partial representation of a concept in our public language, or in some formal language devised publicly for scientific purposes.

Linguistic ability as a faculty within the language using individual is a brain state (not a mind-state) which stands in a causal relationship to the sets of input by which it has been formed and stabilised in learning processes, and it stands in a potential causal relationship to new data which activate those groups of neurones and structures of groups strongest which have been formed by previous input which was most similar to the new data under the perspectives of measurement available. Language as the individual's language faculty or ability is part of the background of our mind, and not part of our mind, and as such ability our language does not consist of rules and representations in our mind. The background can be thought of as a state of our neural networks, possibly with some innate biases in its architecture, in which the connections between the neurones have got and still get adjusted such that the neural nets come up with certain activation patterns under externally or internally caused input. The potential patterns are directly addressed, i.e. activated, by certain perceptual input or indirectly via other patterns, that got activated first. Input, hereby, is classified and associated with possible context types in which it has been learned. This association happens by activating directly and indirectly the respective patterns of units. Certain results of these subsymbolic processes appear as conscious perception and cognition in the mind. In chapter 4 and 5 it will be discussed what kinds of results have an expression in consciousness.

On the other hand, linguistic research comes up with rules, and conceptual analysis provides at least partial definitions of concepts we use, or rather explicates partially our conceptual ability. In linguistic analysis we find systematisations or structures in the sets of linguistic data, by applying structuralist methods such as syntactic-semantic substitution tests, elimination tests, movement tests, and paraphrase tests. Data are pairs of linguistic utterances with their situations of use, especially the satisfaction situations of the utterances. We have a normative awareness of what counts as correct or incorrect linguistic forms and correct or incorrect linguistic usage. This awareness of correctness is called 'linguistic intuition'. We can also improve our language use by consciously monitoring it in comparing it to normative models, and we can even improve and develop our language further by institutional language planning performed in academies and schools.

These quite different facts, the unconscious workings of linguistic ability on the one hand, and the conscious handling of language on the other, can be reconciled by realising that our linguistic ability belongs to the non-cognitive and unconscious background, and merely our linguistic products and their understanding and interpretation in the world are part of our mind. These are real and imagined situations of the use of expressions, and they are the data of our analysis, rather than some unknown internal conceptual representation, which, after all, does not exist in our mind. The concept indicating activation patterns, encoded in the weights of the connections between neurones as the result of learning, are not conscious and, therefore, are not part of our mind. Only what, at a certain time in our life history, is or can be experienced, and this implies that we can be consciously aware of it, is or can be part of our mind and can be treated rationally in contexts of knowledge, rules, and explicit preferences.

It is unnecessary to assume an unconscious part of our mind in which hidden theories and rules can have a place and where they could be found by introspection or reasoning. No introspection ever has lead to such a part of mind and no convincing arguments have been given to postulate it nevertheless, neither by McGinn 1991 in his argumentation for a hidden consciousness based on the blindsight case, which has been discussed above, nor by Chomsky who thinks that a hidden mind would be the best explanation of our linguistic intuitions. Our theories about the world and our knowledge about language is not something hidden in our mind. To have knowledge means that we deal with coherent sets of general sentences held true, related to each other and to experiences of situations by inferences and interpretation (Quine 1974). We experience, for example, what peace is in the use of the word *peace* in linguistic and practical contexts. In our mind we only can recur to such experiences, not to any hidden rules of application of terms. There is no rule or definition used in the application of the word *peace*. This inaccessibility of that which makes us use the expressions of our language in a correct manner, points to the fact that our linguistic ability belongs to a sub-cognitive background. This implies that language belongs to the mind only as far as we perceive utterances and speech situations, have memories of these, and imagine new ones. All these perceptions and images of linguistic usage are the data we are aware of and which constitute the results or products of our language faculty and what linguistic ability means for us. We can analyse them in comparison with each other, and we can compose constructions out of the parts gained in analysis. Linguistic ability is grounded in the sub-cognitive background of our mind as the semantic, syntactic and phonological faculty, which is a set of dispositions encoded distributively in the connection strengths between neural units. Our conscious awareness of our linguistic ability is nothing but a growing set of data of acceptable language use.

In mentalistic linguistics and philosophy, exemplified in Chomsky's and Fodor's work, it is assumed that there is an unconscious part of our mind in which linguistic rules preside and from where they secretly guide our behaviour. Such a notion of an unconscious mind does not make sense and moreover is totally superfluous for explaining linguistic behaviour, since connectionist network models of important structural properties of the neuronal connectedness in the brain explain the causal individual aspects of language use, while public rules and representations describe the normative aspects. There is no hidden mind in between these two sides of language use, as I shall argue for in the next paragraphs.

A notion of unconscious mind is used in Psychoanalysis, according to which we have hidden beliefs and desires, which can be brought to consciousness by therapeutic communication. However, these and other non explicit beliefs are just dispositions of the brain to get activated under certain kinds of stimulations. In contrast to the hidden linguistic rules and representations of the assumed unconscious mind, the unconscious beliefs and preferences assumed in Psychoanalysis are potential activation patterns which can have conscious expression in the therapeutic dialogue. If the activation is effective enough and inhibition sufficiently diminished, we can experience that we have these beliefs and desires in imagining cases in which they clearly play a role, and in the therapeutic sessions they become finally represented in our mind in natural language formulations. But the same is not true for linguistic rules. We never experience these rules in our data. Rather we only can detect them by ordering the data in systematic linguistic analysis, and we can recall them afterwards, if they have been formulated before, like we can recall any representation, which we have experienced, whether it is a picture or a sentence in some public language.

After having been explicitly formulated, rules can be part of our mind like any other belief or knowledge we have acquired. As such they actually come to our mind in the formulations of linguistic science. What we then have in our mind as knowledge are only representations of rules we find as observable regularities in our data, but these are not rules on which our linguistic ability thrives. Even if rules regulating our ability were formulated in some hidden unknown symbolic medium and would exist in this form, they could not be part of our mind, also not of our potential mind, because our mind can only comprise what we can be aware of. And furthermore, since we do not need such hidden rules for explaining our language ability, there is no need for assuming a hidden mind in which we would have stored such rules in some internal code of representation. In the following paragraphs, I shall elaborate the arguments against the assumption of rules and representations as internally guiding and organising linguistic ability.

The cognitive tradition following Fodor (1976, 1987) assumes a hidden language, Mentalese, somehow similar to logical languages. The arguments have been that learning applicability conditions for lexical items, and learning other linguistic rules presupposes such a mental language in which the rules get stated. It has also been argued that propositional attitudes require a medium of representation for their propositional content, which such mental language provides. The first argument is weakened since the growing success of connectionist and probability based models of language learning, which do not require that linguistic rules are learned. It is not anymore used in Fodor's (1998) position of informational atomism, where concepts are supposed to be atoms no longer defined in a language of thought. The second argument is weak, since implicit propositional attitudes have contents which can be thought of in terms of dispositions to come up with neuronal activation patterns, causally connected via the previous learning processes to objects, events and actions, while explicit current propositional contents are represented by the internalised public language or situational images. Neither for learning language, nor for propositional attitudes an innate mental language is necessary as a means of representation (cf. chap. 4). Therefore such a symbolism is useless, and in consequence there cannot have been any evolutionary pressure towards its development. Rather we can assume some evolutionary selected neural biases for forming circuits of activation patterns (cf. chap. 3) conforming to objects and events that are enhanced and delineated by their occurrence in activity and action routines in which the learning individuals are involved. But this is not a mental or innate language. Any conceptual system and any public language developed in human history will be constructed in parallel with such basic action, event, and object patterns. Piaget, for example, based the logical abilities of individuals on a process of internalising external operations on actions, objects, and situations (Piaget 1953). The capability to use such operations is partly determined by innately programmed neural growth, which enables the young child to learn certain operations at certain developmental stages.

6. Linguistic theory formation is rational reconstruction

Linguistic ability has been reconstructed on a logico-philosophical and on a linguistic level. The dominant brands of theoretical linguistics, the mentalistic Chomsky school and the non-mentalistic logical grammarians following Montague, have reconstructed linguistic ability as an ability to form linguistic expressions according to rules. However, logical and linguistic rule systems do not fit very well to real utterances, and even break down, when they are applied in the analysis of these. For a long time it has been assumed, for example by Chomsky and his followers, that a great many utterances in daily communication are corrupted, because they suffer from intervening performance factors, and that the competence of a speaker, the rule system by itself, would produce perfectly correct utterances if no disturbances would intervene.

More recently, the claim that competence means rule following has been weakened (Matthews 1991): Utterances are not built by following the rules, rather in accordance with the rules. This position leaves open how utterances are built. It merely says that they are constrained by the rules in some (unknown) manner. But still recently Chomsky writes that someone "follows the rules of T", the language, "because that is the way he is constructed" (Chomsky 1995: 36). What rule following means remains unclear, like the status of the "inner language" remains very ambiguous, mainly due to the ambiguity of Chomsky's mind/brain notion, a combination of two ontologically totally different notions that is not understandable. Schnelle (1996: 111), among others, points out that the working of the language faculty in the brain is a property of the implicit memory embodied in the dynamic neural connectivities, and not a property of declarative memory, and only for the latter a symbolically computational model would be adequate. This means that to view the language faculty as adequately described by algorithms of rules is far away from what it really is. Also Smolensky (1988) argues in this line when he points out that cultural knowledge and conscious rule interpretation, which have the properties of public access, reliability, and formality, is the domain of symbolic processes, i.e. is part of the conceptual level, while individual knowledge, skill and intuition are the domains of sub-symbolic or sub-conceptual connectionist processes. He views the level of connectionist modelling as a theoretical construct between the neural and the conceptual or symbolic level. It certainly is sub-cognitive and sub-symbolic, and represents on a very abstract mathematical level some relevant properties of the neuronal network structure of the brain. I do not agree with Smolensky's use of the term 'individual knowledge', because I take individual knowledge to be representational too, and thus part of the cognitive level, the conscious mind. Rather I would replace his term 'individual knowledge' by the term 'ability' or 'know how'. It comprises adaptation to surroundings and language use of others. The connections on the neuronal level, trained in learning processes, implicitly contain dispositions or tendencies towards regular linguistic behaviour, which gets idealised in the form of rules in linguistic analysis and theory formation. Here rule formulations are set up and linguistic behaviour is understood in a rationalising fashion as being guided by these rules, although it is merely conform to the rules to some degree.

According to Chomsky (1995), the mind is a computational state of the brain and the language faculty is part of the brain and brings about a cognitive state, i.e. a mind state, which is the I-language, our internal language. In this model it must be the language faculty, i.e. a brain part, which follows the rules according to a programme that is installed in it, basically innate and adjusted in learning. The installed program then must be part of the mind, as well as all states of execution it can be in. If we take the computer metaphor seriously, the program must be formulated symbolically and the states of its execution as such must be likewise formulated symbolically. And when the program is installed and runs it determines the states of the computer. Unfortunately, here the metaphor breaks down: the mind, or the installed I-language does not run the brain, it rather is a state of the brain and thus the program would have to be hardwired within the computer. Apparently in the hidden mind, brain states and hereby parts of the program appear like they can appear on a monitoring computer screen. However, a screen can be observed and the program can be read. A program is public and not in principle hidden. But the problem for the I-language is that we are unconscious of all this, and cannot become aware of our internal language. And therefore it cannot be in the mind.

If we know anything, then it is something that can come to our mind. It is conscious or it can be made conscious, which simply means that there is a brain disposition, a potential towards activation patterns, that can result in conscious awareness of knowing something, for example, knowing sentences of the public language as products of our language ability. There cannot be any, in principle, unconscious knowledge in the mind. However, like the workings of the language faculty, there are many processes and states in the brain which we are never conscious of. But that is no reason to claim an additional hidden part of the mind for these. They are just in the brain, and that is enough. There is no need to declare certain global brain states and architectures as showing up in the hidden mind or as being the hidden mind. A cognitive state, a mind state which would be our internal language, is totally superfluous and thus does not make sense in a scientific linguistic theory. We are never conscious of any brain state or process itself, as such a brain state or process. Rather we are conscious of some of its products, of which we then say rightly that we have them in our mind.

Giving another twist to the computer metaphor, which fits the facts better, the mind can be seen as a kind of interface between brain and world. On the level of this interface we deal with coarse-grained representations of parts of world, and all our mind phenomena are of the same kind, similar to our perceptions of parts of world, playing a perceivable role on this level. If we stress the computer metaphor somewhat further in this direction, we might be tempted to say that the mind is something like the interface running on a screen or monitor, where the input into the computer and its output are presented to us in the same format and can be worked with together in relation to each other. Mind phenomena are all of this kind: Perceptions and images of situations and of sentences, or other phenomena that are used as signs. They are all of the type of perceptions and motoric procedures applied to these phenomena, of which we are consciously aware. We can use, in our mind, images of sentences and other phenomena to operate on by procedures such as substitution, insertion, movement of forms to other places. These operations happen freely or according to rules. In this way we can consciously control the results of the unconscious brain processes that result into thought products in the form of representations such as series of images of situations or sentences.

On the logical and linguistic level we manipulate sentences according to publicly endorsed schemata, and on the level of experiential concepts we manipulate concepts by addressing sequences of examples which represent the concepts. We integrate new data into these representations of concepts salva stabilitate of their internal similarity degree or the internal coherence by contiguity. In rationalist philosophy and cognitive science on the other hand, it is assumed that input and output get related by means of a rationalising construct of rules which are taken to be applied to the input and result in the output. The steps in between are in principle of the same syntactic format as the format assigned to input and output. In such a model it makes some sense to talk of a hidden mind as an extension of the open mind which contains the same kind of syntactic objects and operations on these. The problem with this view is that it is just a construct of rationalisation which is useful for designing some AI-products, but which is not a model of how our brain works.

I basically agree with Searle's connection thesis (Searle 1992: 132) which says that only what can in principle be conscious can be part of the mind, whereby the mind phenomena are somehow caused by the brain. I think that the term 'caused' is problematic in this context, because it presupposes that we have here a relationship within a single ontological realm. The term 'cause' is here used metaphorically and only expresses that consciousness does not exist independent of the respective brain states that, so to speak, carry it. It might be a better metaphor to say that consciousness is an 'expression' of certain brain states or is 'brought about' by these brain states. The mind is at any moment what is conscious and it can be thought of being extended by the potential mind which consists of the results of brain processes and states that can come into the mind under stimulating conditions. The 'potential mind' is merely a theoretical construct which comprises all which can come into the actual mind. It is not a space in which mind products are located and mental operations take place, since it does not really exist. It is just a logical space of possibility of the products of brain activity that can come up into consciousness. The possibility of these mind products is implicitly contained in the connection strengths between neurones of the brain.

Instead of talking about the brain or the mind as following rules, we can better understand the impact of linguistic rules when we view the system of rules as the scientific representation of public constraints on the linguistic output by way of controlling and possibly correcting the much more generous linguistic products our language faculty comes up with. This view takes rules to be norms of the forms and the use of linguistic products which secure easy and co-ordinated uptake of utterances in communication, as opposed to viewing them as an internal algorithm. Nevertheless, that our linguistic norms are largely so regular that it makes sense to formulate linguistic rules comprising these norms in systematic ways is, besides being induced by the semantic function of language and thus bound to regularities in the world, partly caused by a certain architecture of the neural nets in our brain, which have evolved during evolution. It is also partly caused by a tendency towards global regularity which results from accumulating locally steered processes. The network architecture makes certain operations possible which result in substitution of equivalent expressions in contexts, by which we get categories of syntactic or semantic kind. Likewise substitution makes it possible to understand that parts of an object or situation can be substituted by other parts, and that certain parts can be extended in certain ways, without effecting their role in the whole in which they function. This general ability for substitution, insertion, and hereby also recursion, requires certain architectures of recurrent neural networks. This is not specific to the language faculty but serves also in perception, imagination, and action. Architecture and adjustments of connections to the stimuli from the outside world in situational perception and language learning gives rise to regularities in the output, and these structurings on sets of linguistic output get more systematised and idealised by analysis such that they can be described in terms of linguistic rules. However, rules which register regularities, and are idealisations of these, are merely found in our products, not in a hidden mind.

Our brain does not follow these rules in speaking or interpreting. It does not operate according to these rules. Rather the rules are idealisations of regularities that result globally on the basis of local micro-processes and of regularities in input data. -- Later in this chapter, I shall discuss how normative rules, which are external to the language faculty, can perform the function of constraining linguistic output. Idealised descriptions of linguistic output can be emphasised and become a norm, such that only products that conform to them are accepted as correct. By such normativity the creativity of our language faculty gets constrained in processes of learning from accepted examples.

In the light of new developments in cognitive science, especially in the field of connectionist en probability based models (discussed a.o. in Scha 1990, Bartsch 1996, developed for syntactic-semantic parsing in Bod 1995 and 1998), the competence-performance distinction has to be reconsidered (Scha 1990). The gulf between deficient utterance forms and the products of rule systems has to be understood in a different way. The rule systems are idealisations of the output, not models of our competence of language use, and especially not of our production of utterances; our brain works differently, much more flexible and far less restrictive than the rule systems could do which have been assumed to model competence. Utterances, whether conform or not conform to the rules, are rather the products of our subcognitive linguistic competence. Linguistic capacity is organised such that also so-called mistakes are output, which are proper products of the neural network activation patterns which constitute our competence at a certain time and which can change over time due to new input. On the other hand, our linguistic intuitions formulated in correctness judgements, and the standards presupposed by them, are not primarily provided by our individual internal language competence, as Chomsky (1986) has thought; rather they are idealisations and rationalisations of the linguistic products we find in public normatively endorsed language use. They are normative stipulations as far as speakers are norm authorities, members of subgroups that provide models for other members of the language community (Bartsch 1987). These standards, as standards, are linguistic products of these speakers, which can be consciously controlled by comparing them with products of other norm providing speakers. The standards are reproduced by those that have adapted to the norms, and they have to be reproduced by those who still have to adjust to these speakers in learning. The internal language faculty of speakers is not a separate module or organ; rather it is interspersed with, or even consists of perceptual, motoric, emotional, and cognitive functions. It will be more or less adjusted to the products that conform to the public linguistic norms. In this way speakers have built up a competence which is, by itself, not a set of rules, but still gives the output required within broad margins.

The normatively effective output of subgroups of speakers can be formulated in codified norms, and in more general and deeper ways in terms of rules, in which we can realise, in doing linguistic research, even some universal principles. These universal principles might be due to our general cognitive abilities, based on innate brain architecture, stabilisation processes, and even biases in neural nets due to phylogenetic adaptation to orderings in the world. Additionally, universality is induced in ontogenetic development by our common semantics, which is formed in learning to adjust to our surroundings and in acquiring knowledge about the world and the public language. These learning processes work according to principles of similarity and contiguity in sets of data and lead to stabilised clusterings of data, which represent concepts. A concept is the equivalence class of such a data-cluster or data-set comprised by similarity or held together by contiguity relationships. Two data-clusters or sets are equivalent if the can be united without destabilisation, i.e. while keeping intact their internal similarity measure or their coherence by contiguity relationships. In this way there is a process of concept formation going on (cf. Bartsch 1998), whereby public language use guides the individual learners concept formation.

The biases and established states of neural network connections are not rules or principles themselves, nor can they be understood as algorithmic programs. The neural system, in the state it is in at some stage, even if it is stabilised, is much more flexible and dynamic than any algorithm of symbol manipulating rules can be. It will always be able to produce utterances which are not quite according to the rules formulated by linguistic science. These deviations, when publicly accepted, are one source of language change, whereby the systematic orderings are changed, i.e. the rules we can devise for a language at a certain stage. Still, we can say that we show linguistic behaviour which more or less is conform to the rules that describe it under idealisation, but we cannot uphold that we follow the rules in producing our behaviour. In the light of these considerations and the new developments in connectionist network modelling, the traditional cognitivist positions of Chomsky and Fodor are no more convincing. In order to understand our language ability, we need not postulate an internal language as a set of rules and representations, formulated in some hidden mind language.

When speakers are generating data especially for linguistic analysis, they are normatively aware of their linguistic products in performing correctness judgements. They imagine situations of use of linguistic expressions and linguistic products and judge their correctness by simply testing whether they feel compelled to agree to them or to produce them in imagined linguistic situations. Though this method of testing against linguistic intuition cannot serve to test the correctness of one’s own linguistic ability, because for this it would be totally circular, it is a way of carefully controlling and monitoring data with respect to internal consistency and coherence, and of getting material for linguistic analysis in order to describe linguistic structures and devise linguistic rules, which are then not more than an idealised expression of ones own linguistic ability, as it is at a certain moment, monitored by a normative attitude. One’s own linguistic ability can be tested as correct by comparing its products to those of speakers who are accepted as providing models in the speech community, and if one belongs to these speakers one can only test for conformity with the others.

The rules, which are the result of linguistic analysis, provide a systematisation of a standard, exemplified by the linguistic products of speakers who have become adjusted to the norms exemplified by the models provided by norm providing speakers. The rules have to be correct with respect to these speakers' judgements of correctness about linguistic products. These model-speakers themselves need no rules or explicit application conditions for producing and using their linguistic expressions. However, when speaking in a formal style they can do this in a normative attitude, monitoring their speech and hereby adjusting it consciously to other products of this style. Also then no one needs the rules formulated by linguists; all we need is to focus on previous examples under certain points of view, like pronunciation, choice of register, etc. Rules merely belong to a meta-level, on which linguistic products and their situations of use are the objects of conscious systematising, as it is done by linguists.

Speakers produce their expressions from out their background ability, and they automatically compare their products with those of others for conformity, whereby some provide the models, and others follow. And speakers also naturally make their products such that they conform to their own previous products, if these have not been corrected. This all is not done by conscious comparison of new products to old ones, rather this comparison or adaptation happens automatically by making products such that they fit best with the previously established activation patterns in the neural nets. We can formulate it such: The new product is supported by the activation patterns of the previous examples to which it is similar. And when the new product is perceived, as it is in control, it recurrently triggers these activation patterns and at the same time strengthens the connection strengths between the neurones involved in bringing about the activation patterns. This leads to more or less adaptation to models of correct speech, because the input from these sources gets perceived by means of the neural nets of the learner, where it activates those patterns that have been caused by examples which, according to his measurements, are most similar to the input. These can also be examples he had previously produced himself. They will be adapted somewhat more to the model input, but on the other hand the perception of the model input will already be influenced by the state the neural network is in when receiving the input. We all have come across funny examples in which a language learner repeats his own wrong pronunciations even after having been corrected several times. He does not seem to hear the difference between his own linguistic products and the pronunciation of the model. His perception, i.e. what he hears is the product of his auditive input with the state of the auditive network in which the input effects activation.

The problem of completeness and of correctness of a system of rules for the representation of a language exists then in the same way as it would exist for representing how to bicycle by means of an algorithm of rules, or for representing our co-ordination of vision and motoric activity. Only a trained linguist, language teacher, or student will be able to compare his and others’ linguistic products to rules and application conditions. He has to be trained in linguistic analysis and its terminology. And also he, who understands the rules, will not use them to construct linguistic products in real conversation. However, the rules can help for consciously construing some examples or for highlighting the structure of a series of examples. This is equivalent to putting a set of examples in line and continuing the sequence in an exactly specified fashion. What is then learned is not the rule itself, rather the examples and constraints of modifications on these are learned. Rules can help to present examples in an efficient way in a learning process by alignment and opposition. We can learn from the examples which are provided by means of rules, but not from the rules directly. We can learn rules as *know that*, namely that expressions of certain forms are correct and are used in situations of a certain kind, and that certain structures are found in language, but we do not learn rules as a *know how*, as a competence of how to speak. And we never learn directly from rules. If rules are given a role in learning it can only be an indirect one, by providing examples in an ordered manner and emphasising certain aspects of these. And these examples then are used as the data for learning.

A system of rules is correct if it produces the right output, especially the right behaviour, and behaviour might furtheron be judged as correct by stating that it conforms to the rules. But there might never be a complete system of linguistic rules. To view language, along with the tradition of the Chomsky school, as a device for generating a potentially infinite set of sentences from a (small) finite set of syntactic-morphological rules and a closed lexicon suggests a wrong way of looking at linguistic ability and activity, namely seeing it as a program or algorithm for symbol manipulation on a computer. New insights provided by modelling sub-symbolic processing of input and of learning makes it reasonable to discard the old picture of language competence as a system of rules. However, sub-symbolic linguistic capacity is not something that stands by itself. It is only possible, and can arise merely in learning processes in which a public language is learned, and this language consists of public products, utterances in their semantic and pragmatic context, which are, for a good deal, systematically related to each other, due to requirements of efficiency and orderliness, which make it possible to capture these products, to some extent, by linguistic rule systems. The creation of norms is not a random business, rather norms are worse or better, depending on how far they fit with what is already there. New forms will preferably be modifications of old ones, formed by analogy and continuing what is already created, in a regular way. Analogies are created by substitution operations. These preferences concerning the systematicity of norms gives rise to also deeper linguistic rules that themselves can be understood as systematisations over the set of norms (Bartsch 1987). The preferences, of course, are due to our innate general cognitive ability of adaptation of our brain and our behaviour to our surroundings and tasks and to the regularities in our surroundings which are relevant for us and which also must have supported the formation of certain cognitive abilities during evolution.

When we construe a rule system that should reproduce, or even 'explain' a set of utterances, and permit us to make new ones of the same kind, we are performing on the level of consciousness, and in the degree of discreteness achieved there, an analysis of existing or especially elicited linguistic products. Using these rules, we would be able to form consciously new or old utterances by performing a synthesis based on the rules which were devised in this research activity. This new production of utterances on a meta-linguistic level is quite different from how the utterances were produced in the first place, on the subconscious level. It is like replacing a natural process of growing flowers by cutting the results, the flowers, into well discernible parts and putting them together again in order to reconstruct the old flowers, or getting new flowers by new combinations of the parts. Analysing the flowers into parts, classifying the parts and putting them together according to rules declared over the classes achieved by analysis, gives new results, new flowers. But this is not what growing flowers means. It is a method of producing artificial flowers. Nevertheless, we might learn a lot about flowers under the taxonomic point of view, like we learn a lot about sentences in linguistic analysis. But we do not learn everything about how flowers grow, and certainly not how to grow flowers. For that, we have to do genetic studies, and studies about the influence of the surroundings, soil, other plants, animals, and the weather.

Likewise, our philosophical and linguistic reconstructions of linguistic usage and interpretation are based on systematically observing our linguistic behaviour, including our judgements of correctness of language use in inferences. We make assumptions about kinds of orderings on the data and the guiding principles of language use. Additionally, mathematical and logical tools can be used, as they are available at the time of research, or as they can be developed for these purposes. Language use is embedded in action and interaction, and in how we operate on our surroundings and understand and describe our world. A theory of language use therefore takes into account the built-up of actions and situations and how we represent parts of world. Use also shapes the linguistic products which are used, and thus can explain some of the linguistic forms and their semantics, especially a basic logical compositionality in syntax and semantics. The main profit of this kind of linguistic analysis focused on language use is to gain insight in the relationship between language and the matters in which it serves. On the basis of this relationship, we are able to construct formal languages for certain purposes. It is here where linguistic analysis generally is useful.

Theoretical Linguistics provides a rational reconstruction of linguistic behaviour, including its products, the utterances, and their production. Linguistic analysis uses certain kinds of operations and relationships, such as forming substitution classes in context-types, testing for dependencies, and paraphrasing. These first two methods are also applied in the analysis of the situations, actions, and objects that satisfy expressions. There is no way to observe in an interesting way the neural workings steering linguistic behaviour. Observing which areas of the brain are activated in which tasks and making contrast analyses by contrasting tasks or input data and contrasting activated neuronal patterns in a parallel fashion, or finding out in which way connections are strengthened by forming stronger synapses between the axon and dendrite fibres involved, in a biochemical way for short term memory, and in an anatomical way for long term memory, all these recent research results are of greatest interest to neurologists or neurophysiologists. For the philosopher, the linguist, and the cognitive psychologist, they show that the connectionist cognitive model is in principle the right one. But all these findings do nothing for linguistics itself, except that these general cognitive conditions suggests to take into account flexibility in language use for the construction of grammars, and to take the dynamics of concept formation and grammar formation into account for the explanation of language change. There must be structural parallels on a certain global level of description.

Introspection into our cognitive functioning is impossible. What is often called introspection is intuition about ones dispositions towards linguistic behaviour and correctness judgements. *Intuition* is awareness of certain dispositions, under conditions one is not fully aware of, but which are nevertheless operative in experienced or imagined situations of use one is aware of. These dispositions answer to conditions of use of expressions, which are partly found out in semantic and pragmatic analysis and can be seen as theoretical entities in semantic theory. Application conditions for expressions are such theoretical entities. Intuition about regularities in application of expressions and reactions towards them requires that one becomes aware of both, and this requires some training as a language researcher, and of course, consciousness about using language. Consciousness about language use is required for deliberately creating language, for designing a language, for scientific language research, and for making normative statements about language use, especially judging its correctness and devising rules, but it is not required for using language itself. Suitably programmed computers can use and analyse language. Their use of language is based on algorithms programmed in, or also on a learning process, where not rules are learned, but behaviour is induced that conforms to the rules, without being produced by the rules. Also detecting syntactic structures in sets of data and parsing new data on the basis of the structures detected so far can be done by computing probability based procedures. Such automatic analysis, for example in Bod 1998, does not require consciousness. However, the design of such programs does. Without consciousness, one would not even be aware of the fact that one uses or analyses a language.

Linguistic intuitions are in the first place replays of previous language use, and in this replay one becomes aware of the use and its conditions, and on the basis of this, one imagines new situations of language use, especially what one would say appropriately in this or that situation. These are our linguistic intuitions. Knowledge about language is only possible when its symbolic function has been understood and then some language analysis has been performed. Before, there is merely knowing how to use language. We do not consciously identify the conditions of application of an expression. Many possible conditions might be part of the previous situations of linguistic use, some of which we are consciously aware of and others not. In the analysis of our linguistic intuitions we rely on those linguistic products and situations of use of which we have been aware, or we make new combinations of parts of these in our imagination. Mistakes of rationalisation afterwards, i.e. after language use, especially in linguistic theory formation can bedevil our intuitions and give rise to wrong linguistic or logical rule reconstructions, which work in many instances by providing appropriate output, but at other occasions they might fail. Then we are inclined to speak of exceptions to the rule. The alternative would be to do away with the rule altogether and try to find another one. Norm-formulations are based on such intuitions and reconstructions by inventing rules. These norms then may get into conflict with language use on some occasions, which are thereby judged to be incorrect linguistic usage. For a descriptive attitude, it is rather wise to take into account the role of hidden factors and the fact that we never consciously realise all factors involved. Therefore, probability statements about cases of application, which are continuously revisable in the light of new data, are a good second choice solution instead of formulating fixed applicability conditions and rules.

We have explicated the difference between the two ways of looking at language, namely as a capacity or language faculty that steers our linguistic behaviour from out the sub-cognitive background, and as sets of rules and representations, which we come up with in linguistic analysis and which we also can, if we want to, deliberately follow by consciously following rules presented before us, and thus constructing new examples. Likewise we can construe programs that follow the rules. These are ways to implement linguistic norms.

The opposition between causal and intentional-denotational semantics treated in the first section, is a special aspect of the opposition between linguistic ability and linguistic normativity. Intentional-denotational semantics is a product of linguistic and situational analysis and of the normative attitude with respect to its results, typical for human language users. We have described linguistic ability as an ability to language use and, on the other hand, we have described the ability to linguistic analysis, to constrain deliberate linguistic use, to formulate and devise rules, and to consciously follow rules. The results of the scientific activity are implemented in the automatic execution of rules in traditional cognitive systems. We have pointed out that causal semantics, as it is established in our neural nets, belongs to our linguistic faculty and is also the essential background for our conscious denotational intentional semantics which explicates the relationship between language and world. The experience that linguistic expressions have a denotational semantics is the conscious result of our causal semantic processes. Our experience of what expressions denote is partially analysed and formulated in sets of semantic rules which we find in denotational and structural semantics. These rules we could then follow consciously, if we would like to. But, as Wittgenstein has pointed out, all such rule following presupposes the basic language ability to which we are trained. By this training process our causal semantics is established, and only if we are geared to the world in this way, can we interpret any representations or any rules. Such interpretation would be a prerequisite for consciously following of a rule. By training we do not achieve the ability to follow a rule unconsciously, rather we are trained to regular behaviour and to performing tasks which are, to a high degree, conform to the rule which describes the regularity in the behaviour or tasks in rational reconstruction.

7. Linguistic structuring is flexible and is not fixed by a single set of rules

Since language use of individuals is steered from out their neural basis and external factors, we can always expect that language users produce new kinds of data which do not quite fit with what linguists have figured out about previous language use, when ordering masses of data under methodological principles, and by tools we are aware of, such as substitution, movement, paraphrase and elimination tests. Even more, language use can be expected to be deviant in comparison with theoretical grammars which linguists have constructed on the basis of their own linguistic intuitions about correctness of linguistic forms. Deviant data are not just accidents of corrupted speech, but instances that in principle are in agreement with the physiological neural network organisation of linguistic ability at a certain stage of development. What looks unsystematic under one point of view can be systematic under another. Clashes between assumed norms and some particular language use will be resolved one way or the other, largely depending on social factors. Furthermore, what we have experienced in previous language use does not necessarily lead to a certain unique ordering and no other. There can be alternatives, i.e. alternative structurings on the sets of data. In principle all the alternative structurings may be taken into account, because they all can give rise to subsequent linguistic behaviour as continuation of one of these structures.

There are different ways of continuing similarity sets of old expressions by new ones, parallel with our semantic structuring of the growing set of experienced situations into similarity sets. This flexibility corresponds to different possibilities of activation patterns coming up under slightly different contextual conditions. They are delineated by contrasts on our neuronal networks, which continuously get adapted to accepted new input. There may be a flexible choice between several potential activation patterns built up in the course of learning in slightly different contexts, into which the new data fit more or less. The choice may depend on contextual factors and perspectives, we are even not all aware of. Examples of such alternative structurings, which do not exclude each other, are the different possibilities of syntactic coherence of prepositions with their neighbourhoods: the phrase *sit on the bench* can be structured by the constituents [[*sit on*][*the bench*]] or [[*sit*][*on the bench*]]. They come up by different ways of substituting and contrasting expressions. The first gives rise to new constructions like *the bench I sit on* parallel with *the apple I eat*. The second gives rise to constructions like *On the bench, I sit*, parallel with *This apple, I eat*. We have, in principle, two options in linguistic analysis: either admitting flexibility in rule formulation and admitting several rules that partly capture the same input, or rigidification, which allows only one rule and hereby one structure. The choice for algorithms means rigidification, which creates a wide gulf between natural and formal languages.

In the tradition of Generative Grammar always one of the possible structurings was chosen as the right one, while the others were discarded as being wrong. A lot of spurious argumentation was devoted on showing that a certain structure or rule was the right one within the system of rules that was assumed to be the right one, too. Likewise, the NP-VP, or Subject-Predicate dichotomy was declared to be the right universal structure of sentences, although a structure with the verb as head and flexibly arranged NPs are much more adequate for languages with a rich case system or adpositions (i.e. prepositions or postpositions) for expressing grammatical roles. The flexible positions of sentence constituent NP's were treated as results of additional movement transformations applied to the basic clause structure. In alternative theories of grammar such as Dependency Grammars and Flexible Categorial Grammars (Bartsch 1995) the flexibility of the NP-positions has been taken into account right away in formulating one rule-schema which generates all positions for an NP such that there was no need of transformations. The free syntactic order is much more in agreement with the free semantic order of the participants with respect to the action they take part in. Secondarily there are then some soft constraints for word order due to the preference for prototypical sentence forms and due to pragmatic and contextual factors.

It is advisable for theoretical linguistics to admit organising linguistic data in a flexible way, i.e. admit growing sets of data and alternative structures erected on them. Such flexibility follows the flexibility that can be expected from the context dependency of sentence forms and from the global and distributed organisation of activation patterns on the neural-physiological basis. The basic principles of substitution and analogy, which also must be encoded in the architecture of neural networks (cf. chap. 3), admit different structurings on the sets of data. Linguistics has to take seriously the linguistic data available in language use and develop theories that explicitly take care of these. It can profit from insights in neurological models by being open to kinds of flexibility, which have to be integrated into cognitive theory and linguistic theory. Linguistics should not push aside flexibility as a matter of performance, or even as performance errors. Especially systematic errors or deviancies have been part of the empirical basis of linguistic theories, since they indicate properties of the internal organisation of our language faculty. Flexibility of description has been inherent in data-based structuralism. Traditional descriptive linguistics, in the European and American versions of structuralism, has always been an endeavour of structuring sets of data. And there are now theories and programs of data-oriented parsing developed by Van der Berg, Bod, and Scha (1994) and Bod (1995, 1998), which combine structuralist methods with probability measures on sets of data.

As has been pointed out, traditional theoretical linguistics, assuming that there is a single system of representations and rules in the human mind as his internal language, has striven for uniqueness of a given structural analysis, and much of argumentation has been devoted to show that a certain structural description must be correct, while another is not. The data used as basis of the analyses were founded on the normative intuitions of competent speakers of a language. The intuitions had to be intuitions of natives speakers, especially of academic speakers, in formal situations of testing while the language use outside these test situations and language use of non-native speakers and speakers less exposed to the norm of the official standardised language shows more flexibility and leaves more room for the possibilities which are inherent in our natural linguistic abilities. The goal was to systematise the data by a unique set of most general and deep rules, of which it has been claimed by the Chomskyan school that they describe the language capacity which is part of the mind of speakers. I have argued, in agreement with Searle (1992), that the latter is not correct. There is no hidden mind as the place of deep rules. The mind rather is all that is in principle conscious, at one time or other. We may adhere consciously to certain norms when speaking, but not to deep systematisations of sets of norms. Systematising is an on-going process which is the result of unconscious brain activity of the language user on the one side, and it is the result of conscious operations of the linguist, such as substitution, insertion, and generalising over detected syntactic and semantic structures, on the other. Flexibility is a fact of language and language ability which is not compatible with the view that a language should be a unique and fixed set of rules and representations.

The flexible systematisations in structuring growing sets of linguistic data are an intermediate level between the flexible neural competence basis and the idealised and abstract linguistic theory. This theory is not so much a theory of linguistic competence as it is a product of rationalisation of conscious linguistic behaviour produced in eliciting intuitions of linguistic correctness in speakers well-trained in the official standardised language. The results of theoretical linguistics are impressive products of our will to rationalise, i.e. to explain and understand, and even guide, our linguistic behaviour by rules. It is part of the century-old endeavour to understand man as the rational animal. This attitude is inherent in rationalising all our behaviour, especially our linguistic behaviour. It always comes down to a very selective argumentation and explanation afterwards. This rationalising strategy repeatedly breaks down when used predictively. Also grammars never will be predictively adequate of natural language use. Rather they at most will function in a normative way for language users in formal contexts, and in a productive way for technical purposes in systems of traditional artificial intelligence. In this connection Schnelle (1996: 114) argues, extending his Leibniz-interpretation, that the language faculty itself as a brain function, "narrowly related to the general theory of cognitive units (i.e. monadology)", is describable by "numerical methods (the infinitesimal calculus as a basis of a dynamic theory)", while "symbolic calculi (formal logic) are necessary for external control and distinct reasoning". The latter serve as "external tools of cognition" and "discursive control of argument and deliberation". This perspective on the role of grammars gives a good picture of the use they can have for us.

Linguistic theories, which propose grammars that are unique sets of rules, are never descriptively adequate with respect to language use, and even farer away from explanatory adequacy in the sense of explaining how language is acquired and produced. They always are partial and fragmentary. Those grammarians that took into account our intuitions of inferences in natural language produced fragments of natural language grammar, for example Montague Grammars, which are semantically interesting. These are logical grammars, which formalise the relationship between linguistic expressions and our models of worlds. Grammars of natural languages, proposed by theoretical linguists and semanticists, can be understood as normative grammars presented in a very general, or deep formalisation. In them, well-formedness is defined by rules designed according to what we, as educated speakers, accept as linguistic forms when we look at our language, and what we rationally accept as inferences, given the information we have about the world. But these ways of looking at language do not very well represent language in all its practical purposes. They might become attractive under the view of simplicity and elegance of theories which are the best systematisations of normative grammars of such idealised linguistic means that serve for straight forward information transfer. Because they cannot fully take into account the empirical basis, i.e. language use, they give a merely partial outlook on language, namely a view on a normalised grammatical core which has in fact a lot of linguistic activity around it which is not captured by the rules proposed. These grammars are mathematical constructs of idealised or formal languages of information transfer, rather than theories about language, as it is used by individual speakers in a speech community.

Language is situated between three constraining factors, which themselves are not quite stable. Firstly there are the biological foundations of behaviour, perception, and cognition, and the physiological-neural registration within the neuronal connectivity in learning processes. Secondly there is the world around us which has to be represented by language in its relevant aspects. And thirdly there is the normativity of the language community, which is motivated by the need to conform more or less to the same standardised linguistic behaviour over very large groups of people across space and time. Linguistic usage is influenced from these three sides. The first side, as far as it consists of the neural network effects of learning processes, permits much flexibility, and more or less stability.

The second side, consisting of relevant regularities in the world, guides the conceptual built-up and basic syntax, reflecting the functionality of language with respect to representing situations we are engaged in. The third, the normative side, places constraints on what is acceptable. The first permanently gets adapted to the second and the third. But also the world for us is not stable. It gets understood by means of a conceptual capacity which is always developing further. Also our norms change under influence of the language abilities of individuals, when some of them present new models of linguistic usage, which get adopted in subgroups and larger groups. Not everything is allowed that is possible according to the structurings we lay on the growing set of data. Much is not continued, or does not come under attention at a certain stage or never. Much linguistic output, created by individuals or by neural networks that after training behave quite agreeable, is not accepted. Normativity, in judging phonetic, syntactic, semantic, and pragmatic appropriateness, cuts down on the possibilities which at each point in general and individual language development are available to language users. On the other hand, the admittance of creative language use at certain occasions and in certain contexts, for example in poetry or in the speech of subgroup leaders that become models of behaviour, frees this sub-cognitively available potential from the normative pressure enacted in normal circumstances, and gives rise to language change.

8. Learning from rules is possible via the construction of examples

An interesting question is how we can learn from our own mind's designs, such as rule formulations and norm codifications, and how we can teach others by confronting them with linguistically formulated rules. The problem is how norm formulations, explicit rules, and designs can have an influence on neural coding, and how norms of language can cause change of linguistic behaviour. Norms can have such influence, because they are exemplified in normative linguistic behaviour, which enforces linguistic behaviour via a learning process. Normative force influences behaviour by correction, by giving models or examples of accepted speech, and by providing training. Persons who issue, promote and enforce norms, offer models for correct linguistic usage. They provide data, i.e. linguistic products and behaviour, according to their own neural background, and others adjust via registering these data and being trained to accommodate their linguistic usage to these models. Even without models provided by other people, a person can construct examples following a rule step by step, and then he can learn from the examples, and hereby gain the ability for coming up with further similar examples without using the explicit rule.

An individual who is compentent in the language in which the rule is formulated, understands the rule formulation. In the process of understanding the rule formulation, the rule content is construed as the rule's satisfaction conditions. This happens on the situational episodic maps via the conceptual indicators on the conceptual maps which determine activation on the sensorimotor maps. This results in imagining the possible satisfaction situations of the rule, and thus in selecting the relevant features or aspects in the perception of the data given as positive examples for the rule. Likewise it facilitates seeing why the data of negative contrast are not conform to the rule. This control of the rule against the data presented speeds up learning the praxis of acting conform to the rule.

There can even be a creation of norms from the scratch. Whole languages have been designed and then learned, for example Esperanto, or formal languages like logical ones and programming languages. Large or small sets of newly designed norms have been constructed theoretically, by conscious deliberation. For designing a language we need imagination, based on analytic awareness and theoretical understanding of the built up of one or more languages. We have to imagine new combinations and new forms in analogy with linguistic experiences we have had. A design is a product of imagination, guided by previous experiences and rationality with respect to certain preferences and goals. How can then, furtheron, behaviour be produced which follows a design? This seems a difficult question for everybody who thinks that all our behaviour is the result of purely bodily states and processes, without assuming guiding actions from out the conscious level itself. An answer to this question may be the following:

Like we can bake a cake following a recipe, we can make linguistic products also by consciously following linguistic rules step by step. These are external operations. We then have externally built these products and view them as examples of our rules, i.e. structured according to the rules. Making designs and following designs, especially rules, is partially a conscious process. Though we might not know how we come up with certain aspects of the design, these aspects have to be conscious because we have to control and evaluate them with respect to their agreement with the whole design and their functionality or aesthetic quality. Because of their very form and structural properties the structures that are experienced, permit to make overt projections, modifications, and continuations, which lead to new products that may be evaluated. We can construct products consciously following a design more or less freely. In this activity of making examples by construction and then perceiving them, the neuronal assemblies and processes that support these examples best, become activated and modified according to the input examples. The neuronal pattern which is strongest activated by the constructed examples as input, supports these and every further extension of the series of examples. It will be strengthened and modified in order to fit better to the incoming data that were constructed following the design or rule. Now examples can be produced that are no more constructed following the design or rule. Nevertheless, they conform to it to a higher or lower degree. They can be checked in control by generating descriptions of the examples and finding how and whether they fit with the general rule formulation by seeing whether the example descriptions can be derived from the rule formulations. The rigid process of construction is replaced by a soft, flexible process of performing in a reasonable agreement with the design or rule.

Also imagination mediates between rules and behaviour. In imagination, internal construction of examples according to designed rules can be performed, and we experience these internal products like external ones and can learn from them. Such imagination can be exemplified additionally in externally realised products which also can be perceived. The design is realised by constructing and imagining examples following the design step by step, and then we adjust our behaviour to these examples. Thus the consciously constructed examples serve as input into a neural base which hereby learns how to produce these and similar examples directly. Thus, the products that have been constructed consciously by applying rules step by step, can be learned by the same underlying neural system, which has made these constructions and additional modifications of these, from which we then consciously have made selections and new combinations in our design. And in learning the products which are made according to the design, the system acquires a new ability, or an extension of an old one. This is the only way in which a mind that is thought of as merely being emerging experience of certain large structures of neural processes and states, can on its part influence what happens on its neural basis.

The constructive ability of the mind, as far as it is an ability to design by way of selecting from modifications of products that associatively emerge from the neural basis, seems to be an ability of the neural basis to monitor and control itself and to improve itself. This control and improvement is achieved by operating on one's conscious, i.e. perceived or imagined, products, namely on expressions, symbols, pictures, perceptions, and imaginations, and by playing back the results in receiving them as new input data. Also associated products, which only become possible by certain additional contextual input, can be selected and in design strengthened in a conscious manner. This enables an individual to do more than merely react on what it has received as input from the outside. It can transform its own output products by making new analyses, selections and syntheses of these on the conscious level, and it can use these newly constructed products as new input. Learning this new input can further the individual's own abilities in strengthening weak and only contextually available associative activation patterns. New concepts can be formed in this manner and put to use in planning, designing, or analysing new situations. Consciousness is involved in being aware of the products and behaviour which become, in consequence, analysed in different settings, and consciousness is involved in constructing new products and new behaviour step by step, by selection of parts and new compositions according to rules formed or adopted consciously, or according to structural patterns recognised and laid out consciously.

When our brain expresses certain results on the level of consciousness, it can, so to speak, see its own products on a par with perceptions of things from the outside world, and it can deal with its products like it deals in acting on outside objects and situations it has perceived. It gets its products, especially behaviour, represented in the same manner and degree of coarseness as perceived objects and situations are represented, and thus it can deal with them on this level, the intentional level. On this level of representation and understanding we can do the relevant substitutions of patterns within patterns, smaller parts within larger parts, getting new objects, new situations, and new sentences by modifying old ones via substitution and analogy. The products that appear in consciousness can be analysed, rebuilt, and changed like we do with things and actions in the world. Of course, our brain does this work; in a strict sense our mind does nothing at all because it has no brain-independent existence. But the brain can perceive some of its products in the mind, on the level of conscious expression. This happens in the style in which it experiences outside objects and thus can compare them with these and evaluate them. The comparison must be a recurrent play-back, whereby examples in imagination, constructed ones and simply imagined ones, are treated like perceptions of situations and objects. Substitution performed in the mind must be based on associations between parts of products with other parts that fit into the same context types. For example, in the situation in which we perceive or imagine a cat on the mat, the cat can be replaced by a dog, or a parcel, or a toy car, whatever gets mostly activated. All these fit into the context 'on the mat' by having an appropriate size with respect to the mat. By such replacement we construct a new situation, and likewise a new sentence: From *the cat is on the mat* we get *the parcel is on the mat,* etc.

To begin with, only properties in the world which affect us causally such that we can perceive them, can be part of our knowledge and thus be captured conceptually. Therefore also all our linguistic theory building is about external objects, which we can perceive, i.e. which can come into our mind. These are real or imagined linguistic utterances, examples produced on the basis of our intuitions, and their cases of use with respect to satisfaction situations and speech situations. As objects grasped in our mind they can be observed and seen in relation to each other. Our theory of language is a theory of the linguistic products, not of the production. We are conscious only of the products, not of the production process; and our mind, when we try to analyse it, is found to contain products, perceptions of linguistic utterances and situations, or imaginations of these. A theory of mind has to be product oriented; the so-called mind-processes are only postulated but cannot be experienced in the mind. They therefore do not belong to our mind. They, at most, are theoretical constructs of global properties of subcognitive brain processes. These constructs serve as steps in a rational reconstruction of a process that leads to the overt mind products. However, the constructs are superfluous when we are able to directly describe the brain processes that lead to the products. They could only be part of the mind if they would be perceptions, memories of perceptions, or imaginations of behavioural data, or processes which are sequences of conscious steps in overt behaviour, like, for example following an algorithm step by step. These steps can be seen and controlled because they are noted down in a series of products in imagination, or on paper, for example as written expressions, which are visually perceived.

Although our mind is somehow emerging from the brain, mind and brain are ontologically totally different. We have to distinguish the reception of data resulting in activation patterns in neural networks, as subcognitive processes, from experientially and linguistically achieved information, the latter being conscious knowledge. The first do not belong to the mind, rather only to the brain phenomena. Some of them have an expression on the mind level. Situational or linguistic triggering from outside is conscious as far as it is perceived, and not merely received by the organism. In perception it is part of the mind. The basic and prominent mode of the mind is perception, and furtheron imagination. We can perform conscious operations on these mind phenomena. Internal activation and internal triggering caused by the external sources, or also generated by internal stimuli, need not be part of the mind. Its products, though, appear as perceived or imagined, as utterances or other behaviour, in our minds. They are expressed in consciousness and thereby they are part of the mind. Products that come to mind may also be generated by subcognitive activity for which we see no external causes, though they may have causes nevertheless, even if these do not come up in our mind. In those cases it seems accidental that such a product comes up in our mind, and it is taken as an expression of freedom of the mind to form such products or entertain such ideas. Always, if we think about some imagined or abstract topic, we deal with products, namely with unspoken linguistic expressions, with experiences or imaginations of situations, and with series of modifications of utterances and of situations in our imagination (cf. also chapter 4).

Language use can be learned just by training, and it does not require consciousness as long as the semantics of such language use is merely a causal one. But consciousness is required for intentional denotational semantics, i.e. for being aware of the relationships between language and world, as it is required for deliberate change of a language, for deliberate change of knowledge, and with this of theoretical concepts, for explication and setting of norms and codifying these, and for all linguistic analysis and rule formation. In all these cases we deal with the linguistic products or data available, and form new ones according to rules designed. An unconscious system of artificial intelligence never can know that it uses language and never can design a system of language and language use, or for that matter a system of artificial intelligence itself. Even if a computer is programmed to create a simple language for perceived patterns, he just relates patterns of one domain with those from another domain without knowing that on the one side he refers to things which he uses as signs for the objects from the other domain. Because he has no conscious experience of both domains he can not intend that one kind of thing must represent another kind of thing. He thus has no understanding of what he does. The intentional act of symbolising something by something else can only take place in a conscious mind.

The level of consciousness is the perceptive and cognitive level on which we also formulate rules. When the rules get described and procedurally implemented on a computer, this implementation amounts to a rule-driven extension or reduplication of operations designed and performed on the cognitive level by the computer scientist and programmer, which finally have been transformed in series of smallest steps. The construction of new products according to rules can be done by a machine and program designed up to smallest steps of constructive action, conform to the physical basis used in digital computing. We then deal with an artificial system as a product of the mind of the designer, whereby it does not matter who runs the execution of the design, the mind or the computer. The latter will do this in a more perfect way than the human mind can do. Traditional Artificial Intelligence is a truly artificial intelligence, which is a design product of conscious human design activity, an activity in accordance with principles of rationality related to certain goals which the programs have to achieve. This design product can be seen as an extension of the human mind, merely working on its superficial level according to the kind of operations that are the conscious operations performed on the mind phenomena, including operations on known objects, actions, and situations. But this design product is not at all a model of the workings of the basic human intelligence, which lies on a sub-conscious, sub-cognitive level of brain activity. The design product has an artificial intelligence which humans have developed by rationally reconstructing their own overt behaviour, task performance and products, by means of rule systems formulated on a cognitive level, i.e. on the mind level. It presupposes consciousness in design and setting it up, such that it conforms to all kinds of normative notions, semantic correctness, goal directedness and goal adequacy, true beliefs and preferences, which all require in their application a conscious awareness of the data to which they are applied. All this suggests that an artificial intelligence, as long as it is unconscious of everything, never could by itself create another artificial intelligence by observing and analysing its own products, simply because it is not in the position to observe these. Observation implies conscious understanding on the intentional level in its manner and coarseness provided by our language's conceptual apparatus. This is not only sensory reception transmitted onto the neural networks where it activates strongest those parts that previously have been built up through activation by similar observational data. Rather it is having an expression of such activation in form of a conscious experience as the result of the received input. Of course, an artificial system can also be designed to take up its own products as data and do certain things with them, for example according to certain conditions of fitness to certain goals. But this presupposes that a conscious being, a human mind, has been one step ahead such that he has designed the system to perform in this way.

9. Consciousness and intentionality are prerequisites for cognition and knowledge

One of the metaphors about the mind discussed above sees it as the monitor or screen of the brain, or as the conscious interface between the brain on the one side, and the natural and social world on the other, whereby the mind is structured in conformity to the middle-sized objects and events, to which our behaviour is adjusted and among which all our dealings and activities take place. Another metaphor is the one that says that the mind is an 'expression' of certain brain states. All these metaphors are insufficient for really understanding the mind-brain relationship. Everybody nowadays seems to agree about the fact that there is no mind without a functioning brain. But to formulate this relationship as such a supervenience relationship does not add anything to understanding it. Even to speak of a relationship is already problematic, because the mind does not exist independently from the brain, which rather suggests to speak of an internal relationship between mind and brain. It explains why some authors, for example Chomsky, writes about the mind/brain, which even is more obscure than that it clarifies anything.

In the philosophical discussion arguments have been given against the possibility of reducing consciousness to brain states. The arguments are based on the totally different ontology of both, or they are based on the self-reference, which occurs because all knowledge, also theoretical knowledge about brain states and phenomenological knowledge about our mind is developed in our mind, and is therefore itself a mind phenomenon. Brain states are scientific objects outside our mind, but as scientifically known objects they are at the same time in our mind, and are known in the way our mind can know something which is outside itself. On the other hand my mind knows by self-experience what is on my mind. Here is an analogy: We can only talk about objects in space and their spatial relationships to each other and cannot say what space itself is in terms of speaking about spatial objects. Likewise, we can only talk about mind phenomena, including objects of scientific knowledge, and we can talk about their relationships to each other, but we can not say, in terms of mind phenomena of outside things, such as brain states as known objects, what the mind is in which knowledge of these objects is situated. Without spatial objects space seems to be nothing. We cannot say more than that space is the condition of the possibility of spatial objects. Likewise mind is nothing without mind phenomena, i.e. consciousness about objects, feelings, etc. We cannot say what mind is except that mind is the condition of the possibility of mind phenomena and all our knowledge about them, including knowledge about the world. But that does not explain what consciousness or what the mind is. There is always an epistemological primacy of the mind as the condition of all knowledge.

There is no explanation possible about what the transformation from brain states to consciousness is like, in terms of the kind of scientific and philosophical explanation we are capable of (McGinn 1993). This is the case because in doing so we would need to stand outside our mind in order to see and to describe its relationship to the brain. McGinn's argument is that our mind is 'cognitively closed' to the relationship between mind states and brain states. This means that our concept-forming procedures cannot be extended to grasp this relationship, because introspection cannot look into the brain to see the brain properties of the relationship and neurophysiological research cannot see the mind properties of the relationship in the brain. Thus there can never be seen more than a correlation between both aspects of the relationship, without any understanding of the relationship itself. Flanagan (1991) does not accept that the cognitive closure is one of principle; he thinks it is possible that there might be developed a theory that explains the relationship between mind states and corresponding brain states, although it is not based on really seeing the pairs of states that form the relationship. According to him, such a theory must be based on abduction, i.e. it must be a theory that provides the best explanation. However, this does not seem possible as long as the notion of explanation used in this context is not understandable. It would only be understandable if the steps from the brain part of the correlation to the mind part could be formulated in detail. And that is not imaginable, because there will always be the point where we get stuck with just a correlation between physical and mental phenomena.

Mind, the whole of conscious phenomena, is epistemologically primary to any statement of the correlation between states of consciousness and brain states. The latter only can be identified as states evoking consciousness, by correlating them with conscious states. Since not all brain states have an expression in consciousness there must be some observable property in consciousness creating states to which we can attribute the generation of consciousness. Such a property could only be identified by correlations to a state of consciousness. But a correlation by itself is not a reduction or explanation by which one could understand the correlation. Furthermore, it cannot be excluded that even different types of brain states might be correlated to about the same type of state of consciousness in different stages of the history of a learning brain, or even within a short time span, because different contexts might be involved, which cause different brain states to occur, though we are not conscious of these different contexts.

Another way of trying to explain consciousness is pointing out what consciousness is good for and that it therefore must have been an advantage with survival value by which it got supported in emerging in the course of evolution. A metaphor for the mind, based on this view, is the one of the mind as a mirror of the world. This mirror reflects according to its own dimensions and perspectives only what is relevant for the individuals and the natural sort. Taking a Darwinian stance, as Dennett (1995) has done, we can understand that in our natural history we, and all higher animals, got adapted to the middle size objects and events of which we have to take notice in order to survive. Consciousness about these, extended to include self-consciousness, makes rational explanation and prediction of our and others behaviour possible, and thus it makes behaviour to be action, and makes it possible to plan our dealings with the objects around us. The method of rational explanation presupposes the development of a linguistic medium of public and private representation, in which we can clearly delineate situations and beliefs in order to deal with them in organised and flexible ways. Intentionality is primarily consciousness in taking account of the outside world in perception and in producing behaviour as rational action, i.e. in a goal directed and goal-adequate manner. Self-consciousness makes it possible to judge about our doings as correct or incorrect, which, when applied on a broad scale, presupposes sufficient representation of our actions in language, in imagination, or in pictures restricted in their interpretation by aspectualisation, especially by language. Judgement of this kind gives us a way of conscious improvement of our activities and beliefs such that we can achieve growing social and practical abilities and knowledge. These advanced kinds of improvement are outside the reach of other animals. All this is very well, and explaining the emergence of consciousness in this way is at most a historical explanation about why the emergence of consciousness was supported by evolution.

If consciousness has come about in natural history and includes self-consciousness because the above mentioned improvements by judgements about own and other conduct have a survival value, then we can argue, against Dennett (1989), that intentionality is *intrinsic* in the human mind, and to some extent in higher animals as a natural result of the adaptation of our brains to the middle-sized world around us and our place among them. That intentionality is intrinsic to the human mind, and probably to higher animals generally, is Searle's (1983) claim, and he seems to be right about this, since this is what we experience and what we believe about other animals, with whom we make contact and who are also able to perceive their world and seem to have some knowledge about their own doings. That we have intrinsic intentionality is the best explanation for our ability to take the intentional stance towards our own and others behaviour, and towards all kinds of so-called intelligent machinery. It seems that the human mind can take an intentional stance precisely because it is intrinsically endowed with it, and because it is conscious about this in reflecting about his own and others condition in the world.

We can think of stages in the genesis of consciousness and intentionality. Firstly, only that can be conscious what is externally manifested and in this way can be perceived, i.e. experienced. It has got pictorial, procedural, conceptual, and furtheron even linguistic representations by being placed into the orderings of our experiences that are typical for these modes of representation. Our perception, phylogenetically, is adapted to chunks of middle size, some of which we need to avoid, others we can deal with. Positive and negative bodily encounters, as joy and pain, desire and avoidance become consciously experienced, as emotions which are primarily directed towards objects and events. Ontogenetically, our sort is equipped to learn fastest what is most favoured by biases in our neural organisation that have been formed in evolution. We have become endowed with sensors measuring input in certain dimensions or similarity spaces, and neural nets which favour reaction towards the middle-sized objects and events of our natural ontology, and especially favour emotionally activating objects, which become the centre of attention. These objects are the goals of our intentions, and our intentionality is structured by the contents that in some way represent these objects.

Secondly, we can then make more fine grained or more global partitions in the world. On the one hand, we can perform an analysis on the middle-sized chunks, by which we order them into similarity sets under certain perspectives. On the other hand, we can look for general relationships and properties in theory building. Under a perspective we order the chunks according to similarity and opposition. Thus we get under a perspective several concepts which are in opposition to each other and serve for characterisation as well as discrimination of objects. Over these generalisations are obtained. Achieving general knowledge we build up a conceptual semantics according to the fine-grainedness or coarseness of analysis or conceptualisation of objects and actions. In a process of analysis and synthesis we can rebuild old objects, build new objects, and envisage new situations.

Our consciousness of an object goes as far as the analysis goes by embedding it into sets of similar situations and objects salva stability of their similarity measures and thus characterising the object. By embedding it into sets of situations related by contiguity relationships salva stability or coherence we identify it. Our imagination about new objects depends on modifications and new combinations in accordance with the fine-grainedness of the analysis or conceptualisation at a given point. In principle, all the steps taken in a synthesis of different aspects or parts into a whole we can endow with intentionality. This means that we view them as steps of construction taken intentionally, and in this way we can reach finer grained levels of intentionality in rationalising a new product as a construction which seems to have been set up deliberately, although it might be that the product in fact was not constructed intentionally. It rather may be a product of merely unconscious neural processes without a conscious expression in the mind.

The same happens with the syntactic built-up of utterances. We can view every syntactic part of the expression retrospectively as intentionally used in its context. However, taking an intentional stance is only a rationalisation afterwards. It is not the normal way of producing or interpreting utterances. Likewise we can even go down to algorithmic constructions of expressions or actions and view these as intentionally built up according to the algorithm. But this does not mean at all, that in the human mind these intentional steps have been taken in producing the utterance or the action. They are merely intentional actions by the designer and programmer of the algorithms, designed in order to transform our symbolically formulated knowledge into the likewise symbolic formulations of programming languages and finally machine language, physically implemented into the machine's circuits. In the human mind nothing like this normally takes place, though such algorithms can be devised and followed in special scientific and design activities. Rather quite different processes take place in the parallel working neural networks. What we see in a logically built algorithm for performing a certain task or action, or for generating a certain perception, is based on a rational analysis of intentional behaviour, whereby the intentionality which normally is restricted to the larger chunks we are consciously aware of, is now directed towards small parts which are normally not produced or perceived intentionally. For these smaller steps no claim should be made as to their functioning as part of the human mind, as also Searle (1992) has argued. They have no function there, though they have a function in artificial intelligence as algorithmic levels by which the programmer bridges the gulf between a symbolic description of our intentional level of perception and behaviour, and the level of states and processes in a digital computer.

Also on our normal level of perception and representation of objects and actions these are not all produced intentionally. They can be constructed intentionally, but most of the time they just pop up as the products of many parallel brain processes caused by certain inputs. They afterwards may get rationalised as being intentionally constructed in control, performed in a self-monitoring process. Being able to deal with these products in such an intentional manner makes correction and improvements possible in an effective way. When we judge about our behaviour and make sense of it generally, and predict others' behaviour, we consciously relate representations of such behaviour to the world, i.e. we make use of our ability to entertain an intentional referential semantics. Our designs of artificial minds furtheron extend this level of conscious representations by devising finer grained representations of the same symbolic kind, down to finest symbolic representations and processes on the machine language level. And only by being an extension of what our mind produced intentionally through refining the analysis of the products emerging in our mind into smaller parts, the artificial intelligence is cognition, namely representation of our knowledge and a performance of procedures and tasks in a meaningful way. Our knowledge on the intentional level makes that the deliberately fine-grained representations of it, which are designed according to the requirements of the digital computing machine, can be viewed in large chunks corresponding to the objects on the knowledge level. The fine-grained representations can be seen as knowledge representations merely because they are transformations or translations from the knowledge level into another code developed according to a certain technical goal. Only because there is this level of intentionality and consciousness, on which we perceive, represent and know the world, the masses of very fine-grained representations encoded in the machine can be understood by delineating the large chunks, predetermined on the knowledge level, as information about the world. This also holds with respect to neural activation patterns. They only can be understood as representing knowledge by getting delineated according to observed correlations with conscious judgements about what is perceived, what is believed, what is thought, and what is done.

10. There is conscious thought without language

Carruthers (1996) defends the view that all our conscious thoughts are in natural language, and with this our thoughts are dependent on background culture and theories. Thus our mind consists in all the thoughts we can think in our acquired language and is universal as far as all natural languages have common properties. It is culture specific as far as a language is developed within a certain culture and contains concepts specific for this culture. Conscious states are higher-order discriminations of perceptual states and as conscious experiences they are available to acts of thinking. Such acts of thinking presuppose reflexivity because in thinking one has to refer to ones perceptions and thoughts. Reflexive thinking is only possible by means of a language which can represent ones experiences.

In his review Allen (1998) objects that many people think in pictures, not just in language. The generality of thought is then achieved through sequences of pictures which correspond to one thought, for example to the belief that grass is green. Carruthers' thesis implies that young infants and animals have no conscious thoughts, something the author accepts but others might not agree with. Allen concludes that the centrality Carruthers gives to thought in linguistic form in consciousness is questionable.

Allen's suggestion about the generality of thought which is not expressed linguistically but by sequences of pictures agrees well with the view about propositional contents and concepts in my theory of concept formation and understanding, presented as Dynamic Conceptual Semantics (DCS) in Bartsch (1998). Here general experiential concepts need not be linguistically framed. They are represented by growing similarity sets of experienced and imagined situations. Such similarity sets form equivalence classes, i.e. (quasi-)concepts, according to measures of internal similarity. Pre-linguistic quasi-concepts get furtheron stabilised and adjusted to public language in linguistically guided concept formation and in this way become concepts. Historical concepts, i.e. individual concepts and concepts of specific events, are formed as equivalence classes of sets of situations with internal space, time, and causal contiguity and also partial similarity relationships. Each member set is a partial individual concept, whereby the equivalence consists in that all identify the same individual, i.e. they are part of the same complete individual concept.

Basic thoughts, or propositional concepts, are construed as smallest situations (i.e. **B**-intersections, cf. Ch. 2) that fit into general and historical concepts salva stability of these concepts. By fitting into these they are analysed or internally discriminated, and identified. They can be linguistically expressed but need not be. Thus, having a basic thought or imagining a basic situation is construing a smallest situation as being integrated salva stability into similarity sets and contiguity sets, formed on the growing set of previously experienced or imagined situations. This integration requirement for smallest situations constitutes the satisfaction conditions or truth conditions of the basic thought. In the same manner, perceiving a situation means understanding it by integrating it, or fitting it salva conceptual stability into an experientially formed conceptual system, i.e. a system of such similarity and contiguity sets.

Characterisation of a situation happens by embedding it salva stability into general concepts, i.e. into similarity sets. Identification happens by embedding salva stability into contiguity sets, especially into partial individual concepts. This integration process can be expressed in consciousness by imagining an example of a situation that fits into previously established similarity and contiguity sets, which represent general and historical concepts, respectively. Thus, a basic thought does not need to be couched in language. It simply is this process of synthesising into smallest situations the constraints of integratability into historical and general concepts. It can serve as the content of a propositional attitude, such as a belief, or a desire. The integration process, and each play-back of it, defines the type of satisfaction situation for the attitude. Such contents are non-linguistic thoughts. They are not expressed in language, but they have structures that are isomorphic to syntactic-semantic structures of basic sentences. These contents or thoughts are close to language in its basic structures, though they are not real language because they miss the representational medium which we call a language. When they get linguistically represented these non-linguistic thoughts are the satisfaction conditions of the expressing sentences.

In speaking of sets of situational experiences, ordered according to similarity and contiguity, we speak about phenomena in consciousness. They are backed up by brain processes in neural networks which have been adapted to the surroundings in learning. The brain processes automatically perform orderings equivalent to measurements on which our conscious judgements about similarity and contiguity are based. Though there are no unconscious thoughts, there are the brain processes which activate those processes and neural activation patterns that give rise to conscious thought on episodic maps, as I shall argue in chapter 3. If the activation of the episodic maps is missing we still might have the brain activation that otherwise runs with a specific conscious thought, but now happens without consciousness. This brain activation can trigger further activation such that a certain behaviour comes up which in retrospect rationalisation of behaviour will be explained by assuming that the individual had an intention with a certain propositional content. Could we then say that the individual has an unconscious intention or an unconscious thought which triggered his behaviour? Such terminology would mix up brain with mind. Rather we should say that the brain activity was there that in other cases accompanies the specific thought or intention.

A brain activity by itself is not a thought. A thought rather is an equivalence class of stabilising sets of situations that make the thought true. Different people at a certain time entertain different representatives of this equivalence set, and at different times a person can entertain different representatives of the same thought, which are all representatives of the same thought or situational concept. The neural activity pattern of a thought is the result of a disposition implicitly contained in the neuronal connection strength after a process of learning. It is merely a causal effect of learning, and in a specific situation it is activated by a linguistic expression or experience of some aspect of a kind of situation or thing which had been regularly connected to the situations of learning that have been involved in creating this neural activation pattern or parts of it. Parallel with the parts of a thought and their combinations there must be neuronal activation patterns and relationships between these, which physically support the thought. A thought consists of parts which belong to our conscious experiences of situations in learning. The brain activation pattern of a thought, on the other hand, must contain a run-through of brain patterns corresponding to the parts of the thought and something corresponding to their compositions. These situation-induced activation patterns by themselves are not parts of thoughts. Only by way of being causally connected to our sensors and the stimulations that were and are received by them, they can function as the material carriers of our basic thoughts about situations which we can be conscious of in our mind. Without these connections to our sensory-motoric apparatus and the stimulations received by situations and behaviour outside and public, the neuronal activation patterns would not be carriers of thoughts and would not have a conscious expression in terms of thoughts. A brain activation pattern of a thought by itself does not contain the parts of a thought. Rather it just contains physiological states and processes which support the constituents of the thought. The activation patterns are not the aspects of the situations we have experienced, have thought of, and think of. The experienced situations, more or less indirectly, have caused these stable, recurrent activation patterns.

Consciousness is necessary in thought, but part of the underlying brain processes go through without consciousness, when they don't hit episodic maps. Such unconscious processes can work in producing behaviour which can be interpreted by observers in a rationalising manner as actions, although they might lack expression in thought and intention. Underlying brain processes by themselves are never thoughts, and there is nothing we can call an unconscious thought. Each thought consists of basic thoughts and they are all in terms of situational aspects we are conscious of. They need not per se be couched in language. Of course, more complex thoughts that are not just about basic situations, but about constellations of situations and logically structured knowledge and theory formation need language as a means of representation.

We conclude that neither a Fodorian language of thought built up from innate atoms nor a natural language is necessary as a vehicle of a thought, though, of course, much of what we think is couched in our natural languages.

11. Free will has to be learned on the level of consciousness

Action routines, in which activation conditions and execution of an action are specified, are procedures consisting of complexes of concepts related to each other by contiguity relationships, and bound to (pre-)motoric maps. As such they are represented in architectures of connectionist nets with recurrent fibres, in order to capture temporal sequences of sub-actions. They are indicated in the brain by potential activation patterns which are results of implicit dispositions in the set of neuronal connections, due to a process of training in the course of learning. Several action routines may be innate, as they are part of the natural way of securing survival of the sort. They can be further specified in a culturally determined way in additional learning processes. An example in which natural and cultural features are combined would be the action of eating something. When certain input directly addresses neuronal patterns which are strongly connected to certain other neuronal patterns which cause and direct the appropriate action, then the action is executed if there is no additional inhibition. This then is very much a deterministic picture of action. Action seems to be determined by the causal chain of stimulus and response. Free will seems to be excluded by determinism of action. I now shall show why this deterministic picture is not correct and in what sense we can speak of free will.

Hebb (1980) in his *Essay on Mind* has pointed out that although assemblies of activated neural cells in the brain cause action, there is still free will in the sense that many activation patterns are involved, which results in a lot of indeterminism. There is not just one sequence of stimulus and response, such that a series of actions will be realised in a causal chain of events. The mechanical picture of a cause resulting in an effect, like one ball hitting another and the other starting to roll, is not adequate. Many factors can be taken into account by the individual, and higher order strategies of invoking additional circuit activation, promoting or inhibiting activation patterns can be acquired. I now want to elaborate this more involved picture and also show why the assumption of a free will is necessary in order to promote learning processes which result in more freedom of action for the individual. Free will is a matter of degree and it is located between two poles of a spectrum of possibilities. One is the freedom of choice, the other is the acceptance of the necessity of the moral and public law as guiding principle.

A simple example of free choice is that I can choose out of several numbers one or the other. I can say, 'well, I take this', and then, 'no, I rather take another' and 'well, maybe I just take that one' and then possibly just decide 'well, why all this fuss, I will stick to the first, after all'. Here I had no special reason for taking one or the other. In normal cases of action we have reasons, desires, short time and long time preferences. The reasons and long term preferences are given by public norms, moral norms and social norms, and accepted values. We take reasons into account in our decisions to act in certain ways. To decide to act according to the necessities of the moral law has been described as freedom of action by Kant, in opposition to a mechanical or natural drift of an action coming about as a natural response to a stimulus. Our consciousness of the law makes possible our freedom of action against the determinism of action according to animal instincts.

Emile Zola in his grand series of novels *Les Rougon-Macquart. Histoire naturelle e sociale d'une famille sous le Second Empire*, written between 1870 and 1893 illustrated his 'scientific' theory that all doing of people was determined by inherited strengths, weaknesses, and instincts, and by the learned patterns of behaviour, determined by their cultural and social surroundings. In any way, the individual is a product of nature and nurture which reacts to accidents and social conditions within its surroundings. Zola's work provides an illustration and elaboration of determinism, as it was shaped by Darwinism and by the great impression the workings of mechanical inventions had on psychological theorising.

In connectionist network models there is no relation of mechanical causation between stimuli and effects. Rather there is a probability based relationship in which many factors are involved that cannot all be spelled out. Very little effects on the micro level may have or not have, depending on further conditions, a large effect on the macro level of activation of neuronal assemblies which are relevant for causing certain actions to take place. Anyhow, the model of mechanical causality is not adequate for understanding the causal basis of action, and the indeterministic character of probability based models of activities of cell assemblies seems to give new openings to the possibility of free will. However, that kind of indeterminism by itself is not a guarantee for the possibility of free will. It just seems to be a property of our models of description, and the possible levels of description of brain activities that make for indeterminism or determinism. The argument would be that if we could describe the processes on the micro level of cell activity, and if we would be able to take into account all determining factors, the description would look as if our actions were causally determined. And the argument could proceed that on the even finer grained micro level of quantum mechanics, indeterminacy would again be part of the descriptions of the goings on in the brain which are the causes of our actions. We conclude that the property of determinism or the property of probability induced indeterminism of our alternative descriptive models generally has no bearing on the question of whether free will exists, because we can have the one or the other, determinism or indeterminism, depending on the relative macro- or micro-levels of description.

However, the argument above is not to the point, because only indeterminism in the prediction of actions on the level of our ontology of situations, objects, and actions, and its parallel macro-level of neuronal assemblies in the brain has some relevance to the question whether free will is possible. This we may conclude if we assume that conscious phenomena, such as actions and felt dispositions towards certain actions, are expressions of certain actualised neuronal activation patterns. I shall point out below in what sense indeterminism is possible on this level of macro phenomena which constitute the contents of intentionality and action. Especially I want to show that to have a free will is a question of development and learning. To put it crudely: an individual has to learn to have a free will.

Free will means that there is the possibility of choice between courses of action. It may very well be that the choice is not free in the sense that an individual is caused by brain processes to choose one option. This then does not seem to be a free choice but a forced one; the individual is forced by internal brain processes towards one of the options. This, in fact, is always so. The individual's mind and behaviour, all the individual is conscious of, and even the conscious individual itself is an expression of brain processes. I want to show that nevertheless this fact is compatible with having a free will. This is so, because on the other hand, the individual can gain insight in several options to choose from, and for each of these reasons can be gathered, desires and long term preferences can be considered, and in this way many brain processes take place that lead to different activation patterns causing the different possible courses of action. By being able to make explicit possibilities of choice, different activation patterns get actualised and the whole process becomes one of competition between activation circuits and patterns for strength. In the whole process of deliberating between options by consciously taking into account reasons and long term preferences, one activation pattern will come out strongest and thus one option will be realised. By consciously considering alternatives, choices are introduced and the chance is provided that not the activation pattern which is strongest by nature or by habit will win automatically, but one of the other potential activation patterns established in learning values and norms gets sufficiently enhanced by reasoning, consideration, and emotional activation of publicly accepted values. Then an alternative action pattern, or a positively valued natural drift or acquired habit can be the winner. In such a case we would say that free will has gained a victory over a natural drift or habit which is socially unwelcome.

I shall now elaborate on how such a situation can be achieved. For that let us imagine a young criminal who has the tendency to commit criminal act B in situations of type A. We say that he has a criminal habit, i.e. he has internalised an action routine A->B. What is necessary to restore in him a free will with respect to actions in situations of type A?

Punishment following realisation of B in situation A is one means. By creating an adverse emotion with respect to B, it might establish just enough inhibition of his tendency towards A->B in future cases to give room for the introduction of a situation of choice between possible alternative actions in situation A. In order to get a choice, the young criminal has to be trained to formulate alternative action concepts for situation type A. He furthermore needs to have a training to evaluate the alternative actions. Learning the ability of evaluation implies that he has to learn to formulate and take into account his short term preferences, his long term preferences, public norms and values. The publicly acceptable preferences, norms, and values have to be attached to favourable emotions in training. He also has to be trained to deliberate between the evaluations of the alternative actions and to attach a positive emotional value to outcomes which he has learned to be publicly approved of. The strength of enforcement of public norms and values, and the public emotional valuation of these, has to be exemplified in training processes and thus be learned from examples. This internalisation gives strength to certain potential activation patterns in the brain which favour a socially acceptable outcome of the choice between actions in situations of type A. This means that education has to take place in the way described, such that learning processes give rise to further potential activation patterns in the brain which promote choice towards a publicly acceptable action.

The considerations above imply that free will is taught and learned in planting inhibition by punishment, by teaching conscious reasoning and consideration, and by teaching emotional, publicly accepted, evaluation of actions by examples. What comes out in each situation of type A cannot be predicted for certain. Choice of action is not deterministic on the levels on which we describe situations and actions or, for that matter, global brain activation patterns established in learning about these situations and actions. There are too many factors involved concerning internal and external stimulation and strength of potential activation patterns and their internal connections. They make that all our predictions will be a matter of chance. But anyhow, there are at least chances for realising alternative actions C, D, E, or even new surprising ones, instead of just doing B according to the strong primary tendency A->B, we started out with. To have these chances means to have free will of action in situations of type A. This implies that the employment of free will can be established by guided processes of learning, and this is in principle part of the goals of imprisonment and social training of persons with criminal habits.

To those philosophers who would object that what is described here is not free will, we may answer in the line of a Wittgensteinian approach to meaning that this is what we call 'free will' in our social practices, and that there is no other meaning for the term than the one established in these practices of law enforcement, of learning socially acceptable behaviour, and of pointing out to individuals the possibilities of choice and the ways of reasoning about alternatives. Scepticism with respect to the existence of free will then can be seen as meaningless as is scepticism about the nature of knowledge, namely whether there can ever be truth of beliefs. Like we have a clear everyday notion of truth with respect to sentences and texts, we have a clear everyday notion of free will with respect to actions. And like telling the truth can be learned, also acting according to free will can be learned. We have to be content with these everyday notions of free will and truth without searching for absolute notions, which have been assumed to be stable, ideal, and context free.

The discussion about consciousness up to this point justifies the assumption that there is, like in connectionist models, no central control unit for cognitive and emotive processes. However, there is control on the results of brain processes, and this control happens in consciousness and takes into account comparison of the resulting imaginations of situations and actions, represented on our episodic and pre-motoric maps, with external and especially intersubjective data, i.e. with other situations and actions, and with correctness judgements, in theoretical as well as in practical affairs.

Chapter 2

Formal Constraints on the Relationship between Understanding and Connectionist Neural Network Activity

In the previous chapter the relationship between denotational semantics and causal semantics was at issue in the discussion about the roles of consciousness. Denotational semantics, a semantics from a realistic point of view and with normative constraints is the hart of interpretation of linguistic utterances. From a cognitive point of view not primarily interpretation, rather understanding is the topic. Our goal now is to make some formal assumptions about the mapping of the cognitive structures of understanding on processes and states of the brain, such that there we can speak of structures with a degree of coarseness that is appropriate for discussing the causal backbone of cognition. The connectionist neural networks discussed in this chapter are fairly abstract models of some structurally relevant aspects of brain processes. We formulate desiderata for these, which then can be transferred onto the brain, leading to hypotheses about global brain structures.

In this chapter a formal relationship shall be postulated and spelled out between, on the one hand side, understanding of situations and basic sentences, and on the other hand, global neural network activity that models the causal background of understanding in an appropriate coarseness in the brain. Understanding is couched in the model of Dynamic Conceptual Semantics (Bartsch 1998), which will be explained presently. Especially we will try to constrain the relationship between the level of experiential concepts in Dynamic Conceptual Semantics (DCS) and concept formation in distributively representing connectionist neural models (CNM). Both approaches model learning of concepts, especially linguistically expressed concepts, and at the same time they imply theories of understanding situations in perception, and of understanding simple sentences that describe situations. DCS gives a logico-philosophical reconstruction of concept formation and understanding on the cognitive level of conscious phenomena, namely on growing sets of data which are experienced situations. CNM tries to model on an abstract level of neuronal connections and activation patterns the learning of classifications and associative connections of contiguity in perception and behaviour, also linguistic behaviour, as generalisations and associations over sequences of examples. Special attention shall be given to the problem of syntactic systematicity, which is a notorious problem in CNM.

The formal relationship proposed between both models of concept formation and understanding will be a postulate motivated by semantic considerations. Both positions ground language learning in perceptual and behavioural experiences, on which the operation of selecting aspects of situational experiences under perspectives leads to generalisation, while the structuralist operation of forming substitution classes of items within identical contexts leads to paratactic classes of contrasting parts which take part in the same syntagmatic and contiguity relationships. These form syntagmatic and contiguity patterns.

The data of linguistically guided concept formation are pairs of situations and utterances, primarily simple sentences with their experienced satisfaction situations. Generalisation, on the one hand, yields situation concepts or types, especially action types, and on the other hand it yields utterance types, i.e. sentences. Substitution and contrasting within identical contexts leads to distinguishing parts of situations and parts of utterances, as well as parts of situation types and parts of basic sentences. Paradigmatic classes, short: paradigms, are formed as sets of situation parts which can be substituted for each other within situations in order to yield new situations. Conceptual paradigmatic classes are similarity sets, i.e. sets with an internal similarity between the members, due to an identical property or relationship. Likewise linguistic paradigms are formed as sets of expressions which can be substituted for each other within sentences in order to yield new sentences. Within situations, parts of situations stand to each other in contiguity relationships, and likewise sentence parts stand to each other in contiguity relationships, called syntagmatic relationships. These syntagmatic relationships are generalised over all members of two different paradigmatic classes which combined together constitute a new constituent of a situation or of a sentence, respectively. This is the common structuralist basis of our conceptual and linguistic structures, especially of basic syntactic structures. With these syntactic-morphological structures of basic sentences correspond the semantic roles of the sentence parts. They express a part of the possible satisfaction situation for the sentence and the role this part plays in building up this situation. The paradigmatic classes are the categories that are characterised by the syntactic role their members play in the construction of sentence constituents and sentences. In parallel with their syntactic role in sentences, they express types of actions and other relationships within situations, the role participants of these actions or relationships, and the circumstances of the actions. Basic sentences express a situational concept which characterises the possible satisfaction situations of the sentence. Relationships between basic situations and their circumstances are contiguity relationships of spatial, temporal, causal and other kinds. They give rise to sentence connectives of causal, adversative, concessive, temporal, and local kind, and in parallel to compound sentences with adverbials phrases, especially adverbial clauses describing the situational circumstances.

Generally, syntagmatic relationships express contiguity relationships within basic situations and between situations. Paradigmatic relationships express generalisations, i.e. similarity relationships based on identical properties or relationships. Contiguity relationships further play a typical role in forming historical concepts, such as concepts of individuals, actions and events. They can be expressed in texts, such as life histories and descriptions of actions, events, and their position and course in space and time. Over these, further generalisations can be formed leading to general concepts of kinds of individuals and other general historical concepts.

Concept formation starts out on the experiential level, which provides growing sets of data. The data, as far as we are conscious of them, change for us in the process of concept formation, because for us the data are always data which are understood in so far that they, in uptake, are embedded directly into sets of previous data by similarity and contiguity. On the experiential level, which consists of growing sets of situational experiences and their delineating linguistic expressions, we form two kinds of experiential concepts, general concepts and historical, especially individual concepts. Our general concepts are represented by series of examples consisting of linguistic experiences accompanied by the satisfying situational experiences, whereby each member of the series is seen in the light of the previous members as similar to these under certain selecting perspectives. In each stadium of the process of gathering more experiences, such a representation of a general concept as a sequence of examples is partial with respect to the full set of possible examples. Partiality works as a constraint: Each partial representation of a general concept has to be embeddible into the complete representation salva stability of the internal similarity of this complete set of situations, which is the extension of the realistic and publicly endorsed concept.

A *general concept*, strictly speaking, is the equivalence class of all partial representations that are equivalent by being embeddible into the complete representation salva stability of internal similarity. Each member of this equivalence class represents the concept. Since we are conscious only of these sequences of examples, these are the representations of a concept we, in fact, have at a given moment, and this concept as we have it is partial with respect to the complete concept. *Historical concepts*, especially individual concepts, are sequences of situations formed by contiguity associations in space and time connecting each member situation with previous ones. Such a sequence is a partial representation of the complete individual concept, which would be the full life history of the individual. Also here, partiality works as a constraint: The partial individual concept has to be embeddible into the complete one salva stability of the contiguity and coherence of the complete individual concept. The *individual concept* is the equivalence class of all the partial life histories which are part of the same individual life history. Also here all the partial representations are equivalent in that they are parts of a complete individual concept. Our individual concepts are always such partial representations, i.e. parts of life histories of individuals.

On top of this situationally based linguistic stratum, which is grounded in our perception of situations, performance of actions, and reporting about these, we get the logical stratum consisting of a vocabulary of logical connectives, negation, quantifiers, and modal operators of different kind. The logical vocabulary gets its meaning from the ways we organise our information in forming generalisation and existentialisation over particular information by means of general and existential quantifiers, by seeing information in conjunction or disjunction, in conditionalisation, and by seeing it as reality, as possibility, or as necessity. This stratum of the organisation of information, which makes it possible to keep information at a distance from sensory and motoric input and output, on a level of knowledge which is not directly manipulated by situational contingencies, is not captured in the discussion below, because it lies outside, or above, the treatment provided by typical connectionist modelling.

It has already been pointed out (Bartsch 1996, 1998) that connectionist models of concept formation share certain welcome properties with Dynamic Conceptual Semantics:

1. Perception is the core and basis of concept formation.

2. Sensory input is directly embedded into classificatory and contiguity content and thus understood as a perception of a situation. This amounts to direct content addressability by input. A situation is immediately perceived within the conceptual system available at a certain moment, which means that it is perceived as categorised and identified within the generalising classificatory, and the individualising contiguity structure established so far on the set of experienced, imagined, and reported situations. Likewise linguistic utterances are understood by directly addressing the linguistic type in contiguity relation with the type of the given context. Hereby also the semantic content is addressed, due to the association of situation parts with the parts of the expression in the respective contexts, and according to their syntactic contiguity connections, i.e. their syntagmatic relationships, within the expression. There are efforts to work out the corresponding operations of syntactic systematicity in connectionist frameworks by architectures with recurrent fibres using vector concatenation, by vector addition, and by tensor products (Horgan and Tienson 1992), some of which will be discussed below.

3. The structures stabilise in the course of learning from experience (conceptual stabilisation).

4. The structures are flexible in accommodating new and even deviant input (conceptual flexibility).

5. Conceptual growth and change are part of the system (conceptual dynamics).

6. Traditional cognitive representations of mental contents by means of an innate language of thought are avoided. In **DCS**, representations are always in the form of expressions of some natural or formal public language, or in the form of series of pictures and examples. In **CNM**, representations are causally effected neuronal activation patterns, which I rather shall call "indicators of concepts or categories", in order to reserve the term “representation” for representations in the strict sense.

As Wittgenstein has pointed out, representations always presuppose understanding, because they have to be related to something they represent, and thus they function as representations. Therefore they cannot function in establishing understanding and interpretation. They are what has to be interpreted. Interpretation is the relationship between expressions or pictures and their extensions in real and possible worlds. Understanding is the cognitive operation which the interpreter has to perform in order to be able to establish interpretation. In **DCS**, understanding basically is embedding a situation or utterance into sets or sequences of previously experienced situations or utterances salva stability of their internal similarity and contiguity relationships. In this way the conceptual semantics of representations, like sentences or pictures, is part of the cognitive level of perceptual and behavioural conscious phenomena. These consciously experienced phenomena and embedding constraints are publicly co-ordinated in the intersubjective acts of acceptance of satisfaction situations for utterances.

On the other hand, states and processes of neuronal connectivity and activation are the causal background for the conscious results of understanding. These results are representations. Our ability to handle a publicly co-ordinated conceptual semantics is implicitly encoded by the causally effected results in learning. They are encoded within the connection strengths between units in neuronal networks. These connection strengths embody a set of dispositions towards producing sets of neuronal activation patterns within the network. Their realisation in the network depends on appropriate input stimulation. The neuronal system also includes the neuronal units in the sensory and motoric organs. The activation patterns arising from dispositions of a neuronal net, can be assigned to types of input which triggers these patterns, given a certain state of the network. On the output side they can be seen as dispositions towards reactions to perceived situations and to situations which satisfy or dissatisfy utterances, and furtheron they can be seen as dispositions towards actions and utterances connected to these situations in contiguity relationships. With every input a vast cascade of reactive neuronal patterns is generated and it is difficult, though not impossible, to isolate by contrast analysis the activation pattern that primarily corresponds to an expression. Anyhow, all the typical contextual effects of the expression are also activated. They thus belong to the content of the expression represented on a neuronal net, and they can possibly be isolated in contrast analysis by being strengthened or inhibited.

7. In **DCS** semantic taxonomic systematicity is taken care of by inclusion and overlap relationships between sets of data. Based on this, we would like to require in **CNM** a reversed structure of inclusion of activation patterns in neuronal nets. This is motivated by the fact that we can see an activation pattern as corresponding to a set of examples which trigger this pattern when they are input, and some of these examples have been in the learning set that has resulted in convergence towards this pattern for this kind of input. Some sub-assemblies of activated neurones should correspond to semantic features of the examples. These current internal feature indicators have been created in the stabilising process of learning from examples. Such sub-patterns of activated neurones are

not real features, though they must correspond to features of the input examples; and such assemblies or sub-patterns must play a role in bringing about our conscious awareness of similarity between the input examples. They must be the causal basis for the internal similarity we find in the course of growing experiences in the sets of examples which correspond to this pattern. We recognise certain input data as examples of a concept due to that activation pattern. This neural pattern then works as an internal cause for classifying something as being an instance of a concept, and thus can be called a current network "representation", or better "indicator" of the concept. Thus the internal similarity in a set of examples is not just brought about by the sensory measurement of the input examples by the sensory units, which serve as micro-features or dimensions of the input vector. Rather the similarity is established by an reaction of the trained internal neuronal nets, whereby sub-assemblies of internal units identify features in the perception of the data. This also implies that in the course of learning our perception of the data changes due to perceiving other and new similarities between data. This model also leaves open the possibility that we measure similarity by internally indicated features which have no expression in consciousness, i.e. we have no phenomenal awareness of the identities that determine measurements. We might judge two data as similar without knowing on the basis of which identical features we do this.

In standard connectionist networks, the internal units have connections to all input units of the sensor. This means that an internal unit can be activated by quite different combinations of activity from out the input units. In the learning process some connections get strengthened, others inhibited. But still an internal unit is multifunctional, because its activation in combination with other units' activation can indicate quite different input; it can be activated from out different input units. This makes that it cannot be seen as a constant internal feature indicator. It at most can be seen as a secondary context-dependent micro-feature of an input example: only in the context of the activation of certain input units, which are strongly connected to the internal unit such that enough activation is transmitted, it can indicate an internal feature, namely the combination of the activation of these specific input units. The multi-functionality has as a consequence that a neurone in different contexts of other activated neurones can indicate different things. This hinders a neat correspondence between semantical structure and pattern structure in terms of sets of activated neurones and set theoretic operations on these sets. Intersection between two patterns does not always mean similarity of the input. Single neurones in the intersection may indicate nothing in terms of semantic similarity. But the larger the amount of neurones in the intersection the higher the chance that this indicates similarity. In order to make the set of activation patterns of a network indicative for semantic similarity and inclusion, we have at least to require that certain large enough assemblies of neurones have structural properties of inclusion and intersection. These assemblies must be the ones corresponding to features.

It would not be a good idea to take as a concept representation such sets of sub-assemblies of a distributed pattern which correspond to the features included in the concept. On such sets we could apply the set theoretic operations, like we can on sets of features. However, the disadvantage of representing a concept by a set of features, like it is done in linguistics, would be the loss of contiguity relationships between the features, and the loss of those features we are not aware of and have not expressed in our language, but which still might play a role in measuring similarity. All that which makes a neural network's concept indicator correspond to the richness of a gestalt-schema of a concept would get lost in such a set up of representing a concept by a set of features.

Because parts of internal neuronal activation patterns function in our perception of data we get the following correspondence: With forming hypernyms by forming unions of sets of examples for concepts, i.e. unions of concept representations in **DCS**, there should correspond intersections of neuronal activation patterns in **CNM**, and with inclusions of concepts in **DCS** there should correspond reversed inclusion. The formal relationship between both models should then be an injective homomorphism which maps a set of examples of a concept, i.e. a concept representation in **DCS**, onto a corresponding neuronal activation pattern in **CNM**, maps unions of concept representations on intersections, and reverses inclusions. We shall see presently that this requirement is not generally fulfilled.

8. Contiguity relationships in **DCS** can be represented in **CNM** by closeness and association of activation patterns. But this certainly is not specific enough for the different kinds of contiguity relationships, for which special conceptual maps for recognising local, temporal, causal, means-ends, action-result relationships would have to be trained additionally.

9. Syntactic structuring of the process of understanding amounts to adding up specifications of the set of possible satisfaction situations, whereby the binding of the specifications into a basic situational concept is is essential. This binding is crucial for semantic-syntactic systematicity. The final result is a representation which functions as such merely by the way in which it is built up.

Syntactic systematicity on the level of simple sentences is displayed in substitution based on contrast analysis between sets of data, stating identity and difference between simple sentences on the one hand and between situations on the other. Structural hierarchies in conceptual semantics, according to the semantically relevant syntactic hierarchies in a sentence, are realised in the constructive steps in the process of construing the set of situations into which the possible satisfaction situation of a sentence must be integratable salva stability. These constructive steps of a progressing specification of the possible satisfaction situations can be represented in what is called **B**-intersection ("binding intersection") of two concepts (cf. Bartsch 1998). These are intersections of two sets of situations, i.e. intersections of the two concept representations, such that the intersection contains the basic situation into which the two concepts are synthesised. This means that an intersection of two concepts is a **B**-intersection if and only if it contains the basic, i.e. the smallest situation which fits into both concepts salva stability. This means that in concept composition the basic situation fits into both concept representations salva stability. By performing **B**-intersections between the concepts in their syntactic-semantic role, a stepwise specification of constraints on the set of possible satisfaction situations for the sentence is achieved in the process of understanding (for details see 2.2.1 of this chapter). In order to find a formal correspondence between dynamic conceptual semantics and neural network activities, it has to be decided what operation on neuronal nets might correspond to **B**-intersections. It cannot be a simple union of neuronal patterns, or vector addition if the net is represented in the form of a vector. There must be an additional synthetic operation in which forming a union plays a role, like intersection plays a role in the synthetic operation of **B**-intersection. This point will be discussed in section 2.2.

This requirement of parallel and reversed properties between **DCS** and **CNM** is based on a postulated strong structural similarity between the ways in which concepts are represented in both models. Furthermore both models share a common background assumption, namely a non-classical position in cognitive science. From the semantic properties of understanding, explicated in **DCS**, which are based on principles of interpretation and truth evaluation in semantic models, semantically motivated conditions can be derived for constraining the structures which stabilised neuronal networks must exhibit, will they be semantically successful. I finally shall show that understanding a sentence in terms of a resulting conceptual semantic representation can be paralleled by circuits of activation in an appropriate connectionist architecture of neural maps. The process of activating certain neuronal patterns, which are causally related to input from previous experiences in learning processes, in partially fixed syntactic orders is significant as the neuronal basis of sentence understanding. It guides the interpretation of expressions in the world by providing the referential relationships to parts of the world in action and perception. Architectures in neuronal network models which let the process run in some syntactic order, have to be postulated as a neural basis for the process of understanding situations in perception and action, and for the parallel process of understanding sentence utterances. It is questionable how far the postulated formal relationship and requirements on neuronal nets can be realised successfully in current connectionist models which still are restricted in their capacity by their implementation on serial computers. However, at least some beginnings have been made.

2.1 The non-classical cognitive model

The classical model of cognition is designed following the standard computer set up, consisting of a central executive module and peripheral modules for output and input. It has as the centre a unit of rational reasoning, including the execution of rule application on representations in some formal language, corresponding to what has been referred to as the language of thought. Around the centre, peripheral modules are located, namely the perceptual and the motoric module. Information comes in via the perceptual module and output goes out via the motoric module, having trespassed the central unit of rational processing, which connects the peripheral units. This model strokes well with the ideas of rationalist philosophy, which saw the human mind as the rational core of man, getting input and securing output via peripheral sensoric and motoric modules, under influence of an emotional module, mostly competing with, and sometimes supporting the rational processing.

Both models of concept formation treated in this chapter, however, assume a non-classical cognitive model. The non-classical model is rather based on empiricist philosophy, taking perception and motoric behaviour as central for cognitive ability and deriving ideas, i.e. concepts and their combinations in thinking, from perception and memory by operations of similarity and analogy in processes of perspective dependent selection and substitution, and by operations of contiguity in registering factual relationships. Speaking in terms of centre and periphery, the sensoric and motoric faculties and the conceptual orderings established on their processing results on conceptual maps form the centre of primary cognitive activity. The rational faculties form the periphery, which consists in controlling the conscious results delivered from the central basis by applying to them rational and public constraints and making use of the general goal directed and logical procedures in analysis and in composition according to rules.

It has been shown in neuro-biological brain research that the activated neuronal assemblies for performing certain actions, for example for grasping certain kinds of things, and the activated assemblies for perceiving these actions, also when performed by others, largely overlap. Furthermore, perceiving the kind of things that regularly are grasped in this way additionally activates the neuronal assemblies for the respective actions of grasping (Fadiga and Gallese 1997). Thus not only the classification of the actions is represented neurally, associating first and third person perspectives on the actions and representing their perceptual and motoric aspects, but also the contiguity relationships of the actions with the type of things used in these actions are associated in the neural activation patterns. In fact, the whole frame or schema of an action is represented including its types of role participants and its typical circumstances. For example, starting out from seeing a certain kind of things, which we are used to grasp, grasping them is an action possibility which is offered straight away by the neuronal assemblies. Of course, such association comprises much of what is rationally reconstructed in practical reasoning and it follows the ways of performing practices while using certain things. This reminds us on Wittgenstein's example about the use of language by masons in handling different kinds of stones. The stones are what they are by the way they are used, and likewise, we can add, our neural indicators include in the activation pattern for a certain kind of stone the activation pattern of the way in which we normally use it. What appears as practical reasoning in rational reconstruction is not just some rational faculty innate in individuals, but it is guided in a context dependent way and organised by great amounts of practical experience, and by being trained according to public norms and conventions. Therefore practical reasoning is for a large part learned in education within certain types of situations.

In the way sketched above, perception, motoric, and emotional ability form the core of the non-classical model, while rational reasoning rather is the controlling periphery, possibly partly an externalised and explicated continuation of inborn preferences for stability, conservativity, monotonicity, which are publicly endorsed in education. Reasoning is induced mainly through public normative and controlling behaviour, directed towards the goal of getting true, reliable information. Reasoning works on representations, especially in public language, and trained in schooling. This implies that rational reasoning is to some part natural to human minds, but also very much a culture bound and enforced set of constraints for control concerning reliability on what otherwise is produced directly as output from the centrally encoded relationships between sensory input, emotional disposition, and motoric activity. The possibility of control is provided by our faculty of consciousness, i.e. by our awareness of others' and own behaviour, of emotional evaluation, and by our awareness of previous and current perception in memory.

A non-classical model is proposed, for example, by Smolensky (1989), who takes connectionist neural networks as models of the sub-symbolic centre of cognition, while rules and representations are the publicly induced and endorsed ideal rational constraints, imposed from out the periphery. The latter require symbolic and especially linguistic processing on a level of logical vocabulary. This is compatible with the possibility that specific practical and logical constraints are not merely publicly induced in the symbolic mode, but emerge also from sub-symbolic processes resulting into generalisations over neural indicators of concepts, and association (union), overlap, inclusion between such indicators. Parallel we have relationships in consciousness as operations on concept representations, which are here available as stabilising series of instances of a concept.

Concept formation is primarily based on the experiential level of perception and motoric activity, though it is furtheron developed on the level of linguistically explicated theoretical concepts, constrained by public language and theory-formation. These latter concepts are constraints on possible perceptions, actions, and expectations from out the periphery of cognition, which works via operations on symbolic representations of such perceptions, actions, and expectations. These linguistically explicated, i.e. analysed, concepts form the second level of concept formation. They are theoretical concepts in a broad sense which are explicated by coherent sets of general sentences held true. Linguistic labelling on the experiential first level of concept formation is a prerequisite for the linguistic explication of concepts in theories. Experientially introduced concepts can be elaborated, or even be redefined as theoretical concepts, i.e. as linguistically explicated concepts. And in different occupations, technologies, sciences and cultural achievements, there are different theoretical elaborations of experientially introduced concepts, according to their different perspectives, goals, and contexts in these different cultural fields. Hereby a single experiential concept might have several corresponding theoretical concepts as elaborations in different language registers and varieties.

In this chapter I shall concentrate on the experiential level of concept formation, because there we can hope to formulate a relationship between connectionist neural network models and the data-based dynamic model of concept formation. Connectionism, as it is developed today, cannot say much about theory formation, in as far as it is more than generalising over associative relationships. Therefore the genesis of theoretically and formally explicated concepts will not be treated here. If we presuppose that a theory has already been built up on the discursive level, the similarities between theoretical concepts, their opposition and their inclusion relationships can be represented on connectionist networks, because linguistic distributions, i.e. sets of contexts of an expression in normalised sentences can be represented in these networks. Therefore the semantically characteristic syntagmatic distribution of a term X, roughly the set of sentential contexts of the universally quantified term *all X, every X, X in general*, which form a true sentence together with the universally quantified term, can be ordered such that expressions are semantically more similar the more semantically characteristic sentential contexts they have in common. Comparing the semantically characteristic distribution of two terms results in a similarity degree according to how much their distributions overlap. A subset of such a semantically characteristic distribution is a feature, which can also be part of the semantically characteristic distribution of another expression. This way, terms have common semantic features and they can be classified by the features created as sets of sentential contexts on the theoretical level. Hereby similarity, opposition, and inclusion between linguistically explicated concepts is represented in conceptual analysis. By applying these linguistically explicated concepts to objects these will be classified according to the explicit features, and thus will be judged similar according to common features explicated in the general sentences held true.

2.2 Projecting Structures of Dynamic Conceptual Semantics onto Structures on Activation Patterns in Connectionism

In the DCS-model of concept formation and understanding, linguistically guided concept formation on the experiential level consists in creating structures on growing sets of data. Data here are pairs of utterances of simple sentences such as *The boy eats an apple, Peter drinks milk, Fred runs*, together with their satisfaction situations, which are situations that make the respective sentence true. Two types of structures are established. On the one hand there are similarity structures or classifications under perspectives, i.e. similarity sets of situations formed, to begin with under certain basic perspectives such as colour, form, taste, touch, sound, form, motoric behaviour. On the other hand there are contiguity structures. These are contiguity relationships within situations, such as the ones between actions and participants and circumstances. And there are sets of situations which are connected to each other by relevant contiguity relationships. These are factual relationships in space and time.

Similarity structures and *contiguity structures* on the growing sets of experienced situations are compatible in that relationships under one structure do not destroy relationships under the other structure. A situation is determined as that what it is under both structures together. A situation is always an experienced situation the borders of which are not fixed and vary with the focus of attention, which implies making a selection of aspects by the relevant perspective induced, for example, by a desire, an action context, or a question asked on the basis of previous knowledge. A situation is not determined once and for all. Rather the result of applying the methods of structuring under one or more perspectives gives an analysis of the situation. The situation in this form can be used further in applying more structuring, and the results of these further steps can be used as a new data in applying the methods of structuring again. This recurrent structuring is a kind of hermeneutic circle applied in perception. In this way, our set of data does not only grow, but the data also change in the manner they are understood, and in this way old data can become new data. In principle this spiral like process of understanding a situation can go on, and will practically terminate differently in different settings and different states of development, i.e. according to the different structurings available under the relevant perspectives, created by different practical and theoretical contexts, and available at a certain moment in the history of an understanding individual. Results of the process of understanding a situation, whether momentary or final, become conscious in perception and in pictorial or linguistic representation, imagined or uttered. After such a classified and identified item is represented in consciousness, it can come up again in current memory, and the recollected item, or an imagined item of the same kind, can become a new input in a new round of the process of understanding. In what follows I shall formally characterise the two kinds of structures we lay on our growing sets of data, and I shall formulate what their compatibility consists in.

Parallel to the conceptual structure we assume a formal structure on the set of partitions on a neural network which are formed by stabilising activation patterns, actualised with suitable input data. An activation pattern corresponding to a concept in a certain stage of learning and in a certain context is the partition of the neuronal net that is actual whenever the net gets input by examples of the concept or is activated via the use of its linguistic expression. The effects of certain contexts may be extrapolated by using contrasting contexts for the same object, and by keeping the same context but using contrasting objects as data. The activation patterns are results of dispositions of the network to react to input by actualising these activation patterns consisting of assemblies of activated units or neurones. The dispositions can be looked at as functions built up in learning, analogous to semantical intensions, which have internal extensions on the sets of neurones by causing their activation under suitable input into the net, and which have as external extensions the sets of examples which are suitable inputs for getting these activation patterns. This set of dispositions is encoded implicitly in the connection strengths between the units, which has been built up during the learning process. The basic idea about building up connection strengths or weights was Hebb's 1947 conjecture that units which are repeatedly activated together strengthen the connections between them (elaborated in Hebb 1980). The formal relationships between potential activation patterns or net partitions are union, overlap, and especially inclusion.

A problem is the delineation of a pattern because the pattern for a certain type of object, for example a chair, cannot easily be separated from all those activated parts of the net that are due to typical contiguity associations with chairs, such as sitting, table, sitting persons, properties of typical materials used, etc. However, some delineation can be achieved by contrast activation, for example by contrasting the patterns activated by perceiving standing persons, tables alone, chairs alone, whereby the parts activated by what is visually present must be activated stronger than the other parts of the patterns which correspond to what is not present in the input, but is merely associated. Also the associative correspondence with linguistic expressions uttered helps to single out a strongly activated part from other parts of an activation pattern in a contrast analysis of the patterns elicited by appropriate input. Hereby the important point is that the stabilisation of activation patterns and the relationships between them are induced in the course of learning from data, namely from the linguistic utterances together with their respective satisfaction situations. Each pattern can be represented by a vector with the neurones of the net as the vector's dimensions and the activation values as the values in the dimensions.

We need not assume that all members of a speech community show the same activation patterns for the same kinds of objects, nor that for a single person these are identical across time. The only requirement is that the formal relationships between these patterns remain stable in correspondence with the relationships between situations and sets of situations that are causally bound to the patterns by the individual's history of learning. For an individual using language also the relationships to the linguistic expressions of the concepts have to become stable during his history of the learning.

Of course, also pre-linguistically experienced situations are ordered into similarity and contiguity structures, but language learning fixes or modifies these in accordance with the public vocabulary learned. For example, a common name strengthens the similarity relationships between its cases of application in satisfaction situations; a proper name strengthens similarity and contiguity relationships between the different occurrences of an individual object: syntagmatic relationships in recurrent small sentences, the so-called lexical solidarities, strengthen contiguity relationships, for example the ones between chair and sitting, dog and barking, etc. All this strengthening of class-identity or contiguity identity, for example the identity of an object, happens because in principle the use of the same expression in several satisfaction situations is an indicator of a recurrent identity. The linguistic utterances themselves are likewise partitioned and classified paradigmatically under similarity and contrast, and thus become utterance types or expressions, and they get at the same time represented in their contiguity relationships, i.e. syntagmatically within larger expressions. The linguistic and the situational systematisations on the growing sets of data constrain each other by correspondence within the data-pairs, such that situational information helps linguistic analysis, and linguistic information helps delineating linguistically expressed situational concepts.

Between the formalised structures in both types of models, **DCS** and **CNM**, there must exist a mapping from one to the other, which must be structure preserving with respect to the linguistically explicated semantic structures. Because of this mapping we may speak of a "representation" of our conceptual structures in the set of potential activation patterns, or corresponding dispositions towards producing these patterns. These dispositions are implicit within the state of the connection strengths of a network. As far as these networks are designed to model cognitively relevant structural aspects of the brain and succeed in this task, we can speak of implicit representations of our conceptual structures in the brain, which become explicit as internal extensions in a set of activation patterns, or net partitions, which we here shall call indicators of concepts. The structural relationships between activation patterns must have a formal relationship to conceptual semantics. In requiring such a formal relationship we are able, in principle, to capture formally the bridging relationship between the linguistic data and the perceptional data, as it must exist in the causal semantics of the understanding individual. The neuronal network's activation patterns form the bridge per occasion between the linguistic and situational data that are publicly co-ordinated and agreed upon. In the learning process they impose constraints on a person's structural relationships between possible brain activation patterns. The latter internal structures get adjusted in learning to the external linguistic and semantic structures by adjusting the connection strengths between neurones in such a way that the network embodies dispositions towards those activation patterns in which the external semantic relationships are preserved in the mapping from the external onto the internal structures.

2.2.1 The structure of Dynamic Conceptual Semantics

A model of **DCS** is a growing set of data on which two kinds of structures, classificatory and contiguity structures, are established in the process of concept formation. The structures are compatible. Each new datum is understood by being placed into these two kinds of structures in a way that preserves the stability of the structures. This holds for the structures on the linguistic data, the utterances, as well as for the corresponding situational data, the perceived situations. What stability amounts to will be explained below. These two kinds of structuring our data by embedding can be employed recurrently on different ontological levels of categorisation and identification, in discriminating and identifying qualities within situational impressions, situational impressions within (the life histories of) individuals, individuals within situations and constellations of individuals and situations.

The data are pairs (u, s), whereby u is an utterance and s its satisfaction situation. On the set of utterances U types of utterances are formed, and types of their parts are formed by substitution of parts in contrast analysis between utterance types that are partly identical. These types are sets of utterances and sets of parts of utterances, namely sentences and parts of sentences, and sets of sentences and sets of parts of sentences, which are the syntactic categories. They are formed as substitution classes of utterances and of their parts that can be substituted for each other in certain sentential contexts and context types. Categories are thus similarity sets of utterances or parts of these, whereby similarity is relational, i.e. given by an identity of sentential context. Categories are linguistic concepts, i.e. concepts of linguistic and semantic forms. Likewise types or similarity sets, i.e. concepts, are formed on the set of situations. These are situational concepts, or situation types. And as far as linguistically guided situational concept formation goes, types or concepts corresponding to linguistic parts are formed by a parallel substitution in contrast analysis on sets of situations. This implies that for non-negated parts of simple sentences we have corresponding sets of satisfaction situations, namely the satisfaction situations of those sentences in which the expression part occurs in a non-negated fashion. Thus for the phrase *a/the dog* we have a set of dog situations, namely those situations that satisfy sentences in which this phrase occurs.

To keep a concept *stable* means that its internal similarity degree under the relevant perspective is not diminished by adding new examples. - The similarity degree is reconstructible in a quantitative way, or by stabilisation on neural networks.- An evolving concept, a quasi-concept, expressed by *e* stabilises by growing towards a *conceptually complete* set of satisfaction situations for *e*. Such a set is conceptually complete in the sense that its internal similarity degree does not diminish any more by adding new examples of satisfaction situations of *e* under the relevant perspective. A conceptually complete set of satisfaction situations is a representative of the concept expressed by *e*, and the concept itself is the equivalence class of such conceptually complete sets. In dealing with concepts in Dynamic Conceptual Semantics we just take such representative sets as concept representations. The members of a speech community entertain the same concept expressed by *e* if and only if their concepts expressed by *e* can be thrown together into one equivalence class, which is the public concept expressed by *e*. This means that they accept each others' examples for the concept.

On the set of experienced situations, and especially on the growing set of satisfaction situations for sentences, and likewise on the set of the corresponding linguistic utterances, similarity based and contiguity based structures are established, which are compatible. Situations are categorised and identified by their place in both structures, i.e. on the one hand these are categorised by similarity to other situations, especially by identical parts in their internal contiguity, whereby parts get distinguished under identity and contrast with other situations, and on the other hand they are identified by contiguity relationships with other situations in space and time.

On a set S of situations a set structure is induced by forming the set of subsets of S on which the normal operations such as union, intersection, inclusion, complement $\{\cup, \cap, \subset, \backslash\}$ are declared, with the normal set theoretic axioms:

$A \cap B \subset A$, $A \cap B \subset B$,

$A \subset A \cup B$, $B \subset A \cup B$,

$A \cap B = B \cap A$, $A \cup B = B \cup A$,

$A \subset B \ \& \ B \subset C \rightarrow A \subset C$,

$A \cup (S \backslash A) = S$.

From all subsets those are selected that are similarity sets, i.e. have an internal similarity relationship under some perspective among their elements, and also those are selected that have contiguity relationships among their members. Hereby we get a similarity based structure and a contiguity based structure on the set of situations and on the set of utterances.

Similarity based structure:

On the set of experienced situations we form similarity sets as subsets of S. Σ is the set of similarity sets on S. Similarity is a relationship with the following axioms:

x similar x;

x similar y $\rightarrow$ y similar x.

Similarity is graded: x can be more similar to y than to z.

A *similarity set* is a set S such that $\forall xy\ (x, y \in S \rightarrow x$ similar $y)$.

A similarity set has an internal similarity degree D which is the maximal similarity among all the members of the set. If $D > 0$, then there is a *perspective* **P** under which a transitive similarity relation holds between the members of the set by virtue of an identical feature or identical relationship that comes into view under the perspective under which identities and differences are discerned. *Stability* of a similarity set or class means that its internal similarity is kept constant in adding new members.

> A *class* is a similarity set with transitivity of the similarity relationship which defines the similarity set as a class: (1) $\forall x,y,z \in C$ (x similar y & y similar z $\rightarrow$ x similar z)

A *class* is the extension of a *classical concept.* Here all elements have a common property or feature which defines the similarity as the identity of the class in opposition to other classes under the same perspective, for example under the perspective Colour, or Shape, or Behaviour, or a composition of these. A perspective provides a selection of aspects of identity and difference which come into attention in determining similarity and contrast under the perspective. To view something under a certain perspective does not require that one is conscious of the perspective or has a concept of the perspective, and it does not presuppose that one already has concepts of the aspects that play a role in establishing categories under the perspective. In being directed towards situations under a perspective, we can become conscious of the aspects constituting identity and difference under the perspective. Hereby they can become concepts which function as features in a conceptual analysis. Although a perspective can become conscious and be conceptualised itself, it basically is a way of bodily orientation in activities and action by which certain channels of perception, understanding, or analysis are opened or selected. In advanced levels of gathering knowledge and in theory formation, perspectives can be constructed and chosen by making conscious decisions to attend to these or that questions and aspects. A perspective can be thought of in extensional terms as a set of classes of examples which stand in opposition to each other under this perspective.

> X is *in opposition* to Y: $X,Y \in \mathbf{P}$ and $D(X) > D(X \cup Y)$ and $D(Y) > D(X \cup Y)$.

This means that the classes which are in opposition under a perspective have each a stronger internal similarity than their members have with other members outside the class under the same perspective. We also say that X *contrasts* with Y under **P**. The concepts under a perspective are formed by identities and contrasts.

Complex concepts (a term borrowed from Vygotsky 1986) are sets of satisfaction situations for an expression which can be seen as unions of classical concepts that are related to each other by similarity or contiguity between their members under a common perspective. Complex concepts are similarity sets without a class-defining identity, i.e. they are not defined by an identity which must be present in all members and must be absent in all non-members. A special kind of complex concept is one that is based on similarity only. Such a complex concept with at least a family relationship among the members is Wittgenstein's example of German *Spiel* ('play', 'game'). Different plays/games still have a common property, for example they are all activities. However, this property is not sufficient to define the whole similarity set, because it does not discriminate this set from other contrasting ones, such as reading or working.

There are even weaker organised sets of satisfaction situations for an expression which are based on similarity or contiguity (or both), but where there is no general perspective under which similarity is seen as an overall common identity. These have a much looser internal organisation by similarity than complex concepts; they cannot be understood as unions of classical concepts under a single perspective. Rather they merely are complexes of concepts with metaphoric or metonymic relationships between these concepts. Metaphorical relationships between concepts expressed by a single expression typically require a change in perspective under which the similarity is seen. Such complexes of concepts are *polysemic complexes*. They are chain and heap complexes (Vygotsky 1986); the concepts in them are related to each other as members of a chain or a heap by common similarity aspects or by contiguity between the elements of two related concepts.

The more regular complexes of concepts are *centralised heap complexes*, whereby each concept in the complex has common similarities or contiguities with a central concept. If a centralised complex of concepts in union forms a similarity set, then it is a prototypically organised complex concept, whereby the central concept is the prototype to which the others are each related by some common identities. The complex concept contains sub-concepts which each overlap, under a certain perspective, by similarity with the central or prototypical sub-concept. Let now Σ be the set of concepts.

> A *complex of concepts* is a subset Σ' of Σ, such that there is an $A \in \Sigma'$ such that for all $B \in \Sigma'$ there is a chain of concepts $X \in \Sigma'$ from B to A whereby each two neighbouring concepts in the chain are connected by similarity or contiguity, i.e. if X and Y are neighbours in the chain then for $\forall x \in X$, $\forall y \in Y$: x similar y or x contiguous with y in a constant contiguity relationship.

> A *centralised complex of concepts* contains a central class C such that $\forall X \in \Sigma'$: ($\forall x \in X$, $\forall y \in C$: x similar y) or (there is a contiguity relationship R such that $\forall x \in X$, $\forall y \in C$: x R y).

A *prototypically organised complex* is a centralised complex of concepts which is based on similarity relationships under a governing perspective and a couple of more specific subperspectives. This centralised complex is not based on contiguity relationships between the members of its member concepts. Its central member is a class that furthermore has to each other member a stronger similarity than each other member has to all the other members together.

> A *prototypically organised complex* of concepts thus contains a central member which is a class C such that $\forall X, Y \in \Sigma'$: ($\forall x \in X$, $\forall y \in C$: x

> similar to y to degree D, and if x is similar to degree D' to all z, with $z \in Y$ and $z \notin C$, then $D' \leq D$).

The subperspectives under which the different member categories are formed also contain the prototypical member class such that their elements are comparable with the prototypical examples under the respective perspectives. This central member class is the prototypical concept. The union of the members of this kind of complex of concepts forms a *prototypically organised (complex) concept*, where the prototypical set is the central subset with which all the other subsets share a stronger similarity than a possible overall similarity across the whole complex.

> C is a *prototype* in Σ' if and only if $\forall X, Y \in \Sigma'$, $\forall x \in X$, $\forall z \in Y$, $\forall y \in C$: x similar to y to degree D, and if x is similar to degree D' to all z, $z \in Y$ and $z \notin C$, than $D' \leq D$.

Contiguity based structure:

There is a set **R** of contiguity relationships between situations, for example spatial and temporal overlap and adjacency, which define spatial and temporal connectedness, and also distance, and ordering in spatial and temporal space, cause-effect, action-result, means-goal, and other factual relationships. Our basic *ontology* of continuants is due to the most general spatial and temporal contiguity structure laid on sets of experienced situations.

> m is *contiguous with* n = there is a contiguity relationship $R \in \mathbf{R}$ such that mRn.

There are chains of situations formed by space-time and causal connectedness. The space time connectedness is provided by overlap and adjacency.

> C is a *chain* = $\forall s \in C$: $\exists s'$ such that s connected with s', and $\forall s,t$: $s \leq t$ or $s \geq t$ or s overlaps with t.

A *continuant* can be a patch of colour lasting for some time, or a portion of stuff, or an individual, or an event which is a process or non-permanent state or an action. It has internal and relational properties which define its identity through time. An *individual* is a special kind of continuant such that there is some continuous property during its existence and other properties which are changing slowly. Often the property which defines the identity is not known, it is rather merely assumed that there is such a property. However, the individual is identified by its slowly changing properties by means of which it can be followed in space and time as a continuous bundle of states changing only slowly, and by some constant relationships to other individuals. An individual in this model is its life history, and as a complete individual concept it is the complete knowledge of this life history, while the experienced individual is merely a partial individual concept, namely a part of this life history. We can have an individual concept of someone which is partially incorrect. If such incorrectness becomes evident it has to be corrected in a way that the partial individual concept can, in a consistent and coherent way, be part of the complete individual concept, the actual individual. Reference to an individual via a partial individual concept means that the partial concept is assumed to fit coherently into series of steps towards completion into the complete individual concept. *Stability* of a partial individual concept amounts to its growth into further partial individual concepts which are consistent and coherent within one possible completion, the complete individual concept.

An *event* can be point-like or can be a continuant. As a continuant it is a closed process with a temporal structure consisting of a beginning state, a couple of intermediary states, and an end state. An *action* is an event brought about by an agent due to motives and goals. For a thorough analysis of these ontological notions a large philosophical literature is available. To take it into account here would take us too far off. What is important rather is that our ontology arises by establishing a combination of similarity and contiguity based structures on sets of experienced situations. The *compatibility conditions* for these two structures imply that in historical concepts, i.e. in individual concepts and concepts of other spatio-temporal entities, both structurings co-incide: Each situation is *categorised* by embedding the situation into similarity sets, i.e. general concepts, and it is *identified* as a specific situation by embedding it into one or more individual concepts, partial life histories, or into relational contiguity networks of other spatial-temporal entities. Both structures intersect in determining, i.e. classifying and identifying, situations and more complex contiguity based entities in space and time, such as individual concepts and other historical concepts. Both kinds of embedding a situation have to be done into a stabilising system, or salva stabilitate within an already stabilised system of general and individual concepts.

In order to understand or analyse a situation, also as far as this takes place in experiencing a situation, we place it into these two kinds of structure, the similarity based and the contiguity based structure, by integrating it into similarity sets and contiguity sets salva stability. That both structures are *compatible* means that situations, as that what they are for us, i.e. as experienced situations, are determined by both structures at the same time. Both structures get developed on growing sets of data, i.e. develop with growing sets of experiences of situations, by ordering each new situation by similarity and contiguity into those subsets of situations which are already built up in the previous states of experience. In the course of development, the understanding of situations changes. Old experiences can be re-analysed in the light of new experiences, which give rise to new categorisations or new continuants into which they can be included.

Our sets of data grow. New data are added to old similarity sets and to old contiguity sets in a way that keeps these stable or coherent. To keep old structures stable implies to extend them by adding new similarity and contiguity sets for those data that do not fit into old sets salva stabilitate of their internal structures. This also holds for data that are nevertheless referred to by the same expressions as the old data with which they do not fit in a stability preserving manner. Such new language use is metaphoric or metonymic use of linguistic expressions. These newly created similarity sets partly overlap with old sets, on the account of similarity or contiguity. In correction of conceptual systems, on the other hand, old structures are partly cancelled in favour of new ones.

Conceptual growth comes about with conceptual change by extending and elaborating a conceptual system. Here complexes of concepts get extended by forming new similarity and contiguity sets from old and new data that do not fit together into one of the already existing sets salva stability. A preference for keeping old sets stable in their internal similarity and contiguity structure leads to forming new sets, based on some members of the old sets to which the new data are added because of their similarity or contiguity to these members, often under new perspectives. Thus similarity or contiguity under a new perspective leads to metaphoric or metonymic transfer of an old expression onto new data of a different kind. If the perspective remains constant in the transfer of a term to new examples and the internal similarity degree decreases, we deal with a case of concept broadening. If the labels of the old sets are used also for these new sets, either a polysemic complex of concepts is metaphorically or metonymically extended under a new perspective, or a concept is broadened under the old perspective. If new names are devised for these new sets of examples, the vocabulary is extended. - This all is extensively treated in Bartsch 1998. -

Basic syntactic structure: *The combination of similarity and contiguity structures in perceiving and understanding expressions and situations*

Our basic ontology, to start with, consists of experienced situations which are structured according to the impressions we have of the situations, our situational experiences. With growing numbers of such experiences they become analysed more and more by establishing more and larger similarity sets on this growing set of data. The ontology gets revised by embedding situations into both structures, similarity and contiguity sets. Contiguities give rise to further similarities, especially relationally based ones. Thus a banana is classified together with apples and pears as fruit because of its similarity to these due to qualitative identities of their contiguity relationships to eating, tasting, touching these things, and, with further experience about these things, to growing on trees and on other plants. Rice in a bowl, for example, is identified and classified together with rice as a plant by its contiguity relations with processing of rice seeds and growing rice plants. In this way we come to see situations and objects as belonging to certain classes and to the life histories of individuals or other historical entities.

Individuals, in this situation-based ontology, are sets of situations structured by several similarity and contiguity relationships. An individual is a completion of partial individual concepts, which are sets of situations embedded into a contiguity order of situations that typically holds within life histories of the special kind of continuant which this kind of individual is. Furtheron, similarity sets are formed over individual concepts, which have been established in this way. These are classes of individuals as there are the natural kinds and other general concepts over individuals. After having formed individual concepts, a situation can be seen as a part of an individual concept and thus can be perceived as a state, an action, or an activity in which the individual is engaged as an agent or in an other participant role. On this higher level of analysis and understanding, reversing the order of antecedency in concept formation, situations are now perceived not as basic, rather as being analysed according to our ontology of individuals, events, actions, states, and processes, which we have built up by combining similarity and contiguity structures in our uptake of situations. Since perception consists in embedding sensor-motoric data into these structures, the data as perceived data are already understood as such and such situations in the act of perception. With the dynamics of our conceptual systems also our perception of situations has changed over time, and thus our world as a perceived and understood world has changed.

The two types of structures, similarity and contiguity structures, in combination with certain restrictions, naturally or conventionally imposed for different kinds of entities, set up our ontology. The two types of structures are not only compatible but they together establish the coherent network of our historical and general concepts, as far as it is based directly on experience. Therefore a perceived situation, an individual, event, action, state, or process does not just consist of parts perceived due to sensoric stimulation within certain motoric contexts. Rather these stimuli are automatically embedded into similarity and contiguity sets and thus are analysed or understood in the process of perception. The perceived object contains additional aspects which are induced by the embedding of the incoming sensor-motoric data into general concepts in categorisation, and into historical, especially individual, concepts and in identification of the object. We perceive a car, although we do not really see the backside of the car. And we perceive the whole front of a car as a gestalt, although in rapid eye-movement we merely repeatedly scan over the front side, connecting stimuli of its parts in recurrent processes. By embedding the perceived data into contiguity and similarity relationships, the whole object is the perceived object, whereby what is perceived is hereby ontologically categorised and identified. Perception of something, i.e. what we perceive consciously, does already contain cognition of the object. This is a Kantian and Husserlian notion of perception, which also is implied in **DCS**, and also in **CNM**. What is perceived is not just a bunch of received sense data without categorisation and identification. In perception, data are always already embedded into previous organisations of data. Though at first we might merely have qualitative and quantitative categorisation and some space-time contiguity of situational impressions received via our senses, as soon as an ontological order is established on our growing sets of data, perception does not anymore consist simply of situational impressions within some qualitative and quantitative categorisations by similarity and contrast, and within some space-time contiguity relations. Rather by combining similarity and contiguity structuring in an ontological categorisation and identification, situations are then understood as built up of individuals and actions, events, states, and processes, in which the individuals are involved as participants.

The primary data are basic sentence utterances and situations. From a set of basic sentences with, in parallel, their satisfaction situations, whereby sentences and likewise situations are related to each other by identities and contrasts, concepts of the parts of sentences and concepts of the parts of situations are constructed in their relationships to the other parts of the sentences, and to the other parts of the situations respectively. Hereby also the relationship between the morpho-syntactically expressed (relational) linguistic case concepts with the semantic role-concepts are constructed in a context-dependent way; i.e. the cases can express different participant roles in the context of different verbs. An individual in a situation has a certain participant role, which is morpho-syntactically expressed by case marking, by pre- or postpositions, or by the syntactic position of the respective nominal term within the sentence.

Construing and understanding new basic sentences and new basic situations goes hand in hand: the structure of this process is recoverable in the structure of the result, the sentence and its satisfaction situation. In understanding a situation in perception, we can construe in parallel a sentence to describe the situation, and in parallel with understanding a sentence as a linguistic entity we understand its content by construing its possible satisfaction situation in a hierarchically organised process. The steps of the construction can be represented by **B**-intersections (Bartsch 1998), i.e. binding intersections, namely intersections between concept representations such that the intersection contains the smallest situation. This is the situational concept in which the combining concepts are bound together into a basic situation.

> A ***B****-intersection* of two concepts A and B is the intersection of a representative of A, i.e. a partial extension of A, with a representative of B, i.e. a partial extension of B, whereby the intersection is formed by adding the smallest situation s to both concept representatives salva stabilitate.

This means that we construe a smallest, i.e. basic situation s which must fit into A and B without impairing the stability of A and B. For general concepts, this implies that the addition of s does not diminish the internal similarity within the maximal, i.e. conceptually complete or stabilised, similarity sets under the relevant perspectives. For individual concepts, this implies that there is a part of it to which we can add s such that the coherence of the partial individual concept is not impaired by adding s. This is enough for understanding s as being part of a certain individual. For taking s as true more would be required, namely that situation s is embeddible into the individual, i.e. into its whole life history, preserving stability, which means that it is embeddible into the complete individual concept in a coherent way. Construing a smallest situation s in the intersection of two concepts is what in Kantian epistemology has been called the synthesis of concepts or categories by the power of imagination It is not simply the logical operation of conjunction or intersection. Rather both concepts are bound together into a smallest situation. This smallest situation is aspectualised to the composing concepts, i.e. its only aspects are the composing concepts. Note that the smallest situation defines the similarity measure of a concept, i.e. determines the similarity of a conceptually complete set of situations. It follows that any set of situations that contains a basic situation with this aspectualisation as a member is a conceptually complete set, i.e. a representation of the composed concept.

In imagination we can consciously represent the smallest situation, or the construed situational concept, by a series of imagined basic situations all containing the smallest situation as their aspectualisation. The concept construed by the **B**-intersection of two concepts can then be represented in consciousness by a set of situations each of which contains the smallest situation of type s as its aspectualisation. For connectionist networks, on the other hand, the difficult task is to devise an operation on the networks that will mirror this synthetic operation of binding and which might be considered as the kind of real neuronal network operation which has in our consciousness the result of the act of imagining basic situations which combine two, or more concepts. This problem will be tackled in section 2.2.2 and in chapter 3.

We shall presently see that the smallest satisfaction situation of a sentence is not construable without a minimal hierarchical structure of sentence construction. The result of understanding a sentence, and parallel of understanding a situation, is the construction of a smallest situation which is located as an element in a hierarchical architecture of sets of situations which are situational concept representations. Some sets are classes, others are selections from individual concepts. For example, the general concept Agent is represented by a conceptually complete class of situations in which someone is an agent, and the individual concept John is represented by a set of situations which are part of John's life history. Then a set of situations representing the combined concept John as Agent, a partial individual concept, is a subset of a partial individual concept of John, such that in these situations he is an agent. This is a subset of the set of situations that is John's life history. Within the selection expressed by *John as Agent* a step of a sentential hierarchy, for example of *John beats Paul*, is conceptually encoded. This requires that in understanding the sentence *John beats Paul*, the set of *John*-situations and the set of *Agent*-situations are **B**-intersected first before this whole is **B**-intersected with the set of *beat Paul*-situations. The last is construed by first **B**-intersecting the set of *Paul*-situations with the set of *Patient*-situations and then **B**-intersecting it with the set of *beat*-situations. It would be semantically equivalent to first construe *John beat*-situations and then form the **B**-intersection with *Paul as Patient*-situations. Likewise, but redundantly, we also could **B**-intersect the set of *John beat*-situations with the set of *beat Paul*-situations. Either way, the result is the same: The possible satisfaction situation for our sentence has to be a member of the set of situations that each contain the smallest situation construed by the above procedure, which is syntactically guided, though different syntactic routes lead to the same result. Of course, this procedure distinguishes *Paul beats John* from *John beats Paul.* The result of construing the smallest situation that has to be contained by all possible satisfaction situations of a sentence is dependent on syntactic structure.

In consciousness, relationships, especially actions, get classified by substitution operations employed on representations, whereby a variation of the participants of a relationship can be achieved on a set of representations of sentences and situations. We form constant one-place predicates by variation of the first place term, by subject variation, we get *x eat(s) bread* from sentences like *I eat bread*, *John eats bread*, *a girl eats bread*, etc. By variation of the second term, object variation, we get *I eat y* from sentences like *I eat bread*, *I eat pudding, I eat an apple*, etc. By variation of subject and object term we get the relationship *x eat y*. The relationship in fact is a growing set of sentences with their respective satisfaction situations, whereby the sentences and likewise the situations are related to each other by such substitutions of terms in subject or in object position, or by substitutions of the respective Agent-participants and Patient-participants in the respective situations. The internal similarity of the set is given by what remains constant under substitution.

B-intersections can be performed by substitution. In order to form the **B**-intersection between the partial individual concept John as Agent, i.e. a set of situations in which John is involved as agent, and the relational concept Beat, i.e. a set of beat-situations, we extend both with the smallest situation s that fits into both, salva stabilitate. This means we construe a smallest situation s that fits into the set John as Agent and into the set Beat, salva stabilitate. This now can be a Beat-situation construed from some other Beat-situation by substituting John for the Agent-participant, parallel with construing the sentence by substituting *John* for an other subject term in, for example, the sentence *Peter beats Paul.* An analogous step is taken to get the **B**-intersection between the set of John Beat-situations and the set of Paul as Patient-situations, and in this way the smallest situation in which John beats Paul is construed, parallel with substituting *Paul* for possibly some other Patient-term in, for example, *John beats Greta.* We can construe new complex expressions from old complex expressions by substitution, or variation, a method also called 'projection' from old to new expressions (cf. Ziff 1960). The operation of **B**-intersection is a parallel operation to substitution in sentences, now on the level of situational concepts, which are represented by stabilising or already stable, but still growing, sets of situations. Substituting Paul for Greta in a situation type or situational concept such that we get the situational concept John Beats Paul amounts to forming the **B**-intersection between the concept John Beats and the partial individual concept of Paul as Patient. Such substitution is the same as a unification of Paul as Patient with the Patient-aspect in the action concept Beat.

On the level of **DCS**, the sentence *John beats Paul* is understood by construing **B**-intersections between sets of situations representing the involved concepts:

A: John, i.e. a set of John-situations that form the partial individual concept of John;

A': John as Agent, i.e. a set that is the **B**-intersection between the concepts John and Agent, which is the set of all situations that contain the smallest situation in which John is Agent;

B: John Beats, i.e. a set of situations in which John beats somebody. It is the **B**-intersection between John as Agent and the relational concept Beat, which is a set of situations that contain the smallest situation in which John is Agent and the relationship Beat is realised. In construing the smallest situation the Agent-aspect of John as Agent is unified with the Agent aspect that is included in all Beat situations. That the agent aspect is included in all Beat situations means set-theoretically that the set of all Beat-situations is included in the set of all situations in which there is an agent, i.e. in all activity situations.

L: Beat, i.e. a set of Beat-situations

C: Paul, i.e. a set of Paul-situations that form the individual concept of Paul;

C': Paul as Patient, i.e. a set that is the **B**-intersection between the concepts Paul and Patient, which is a set of situations that contain the smallest situation in which Paul is Patient;

D: Beats Paul, i.e. a set of situations in which someone beats Paul. It is the **B**-intersection between the relational concept Beats and Paul as patient, which is a set of situations that contain the smallest situation in which the relationship Beat is realised and Paul is Patient.

E: John beats Paul, i.e. a set of situations that contain the smallest situation in which John beats Paul. It can be construed as either the **B**-intersection between B and C' or between A' and D.

E can also be construed as **B**-intersection between A' and C' followed by **B**-intersecting the result with L:

F: John as Agent and Paul as Patient, i.e. a set of situations each of which contains as a constituent the smallest situation in which John is Agent and Paul is Patient. This set then can be cut down further by **B**-intersection with the relation-concept Beat, which gives us a representative set of the situational concept of John beats Paul.

This account of alternative ways of understanding the simple sentence *John beats Paul*, by construing the situational concept it expresses in forming **B**-intersections, makes that the situational concept of John Beats Paul implies several set theoretic inclusions between the constituting concepts. The order relationships in the series of inclusions represent the possible alternative hierarchies of syntactic construction that all lead to the same result, a representative set of the situational concept of John beating Paul:

$A \supset A' \supset B \supset E$ i.e. E = **B**(**B**(**B**(John, Agent), Beat), **B**(Paul, Patient)))

$C \supset C' \supset D \supset E$ i.e. E = **B**(**B**(**B**(Paul, Patient), Beat), **B**(John, Agent)))

$L \supset B \supset E$ i.e. E = **B**(**B**(Beat, **B**(John, Agent)), **B**(Paul, Patient)))

$L \supset D \supset E$ i.e. E = **B**(**B**(Beat, **B**(Paul, Patient)), **B**(John, Agent)))

$A \supset A' \supset F \supset E$ i.e. E = **B**(**B**(**B**(John, Agent), **B**(Paul, Patient)), Beat)

$C \supset C' \supset F \supset E$ i.e. E = **B**(**B**(**B**(Paul, Patient), **B**(John, Agent)), Beat)

We see that the process of construing the complex situational concept is a gathering of constraints on the set of possible satisfaction situations for the sentence at issue; in understanding the sentence step by step, more and more constraints are placed on the set of possible satisfaction situations, by which the set becomes smaller. The process of understanding a sentence is thus a specification, which involves some fixed hierarchies, namely the combination of the role-concepts with the participant individual concepts. The concepts construed by **B**-intersection between two concepts are not empty, because either the smallest situation construed is already a member of both concept representing sets, or it is added to both sets, salva stabilitate.

We have seen that understanding a sentence follows hierarchies of constituents in performing the syntactic-semantic **B**-operations. For a simple sentence the combination between noun phrase concepts and participant role concepts has to be performed first. After that the formation of constituents can be done either way. This means that within the sentence the combinations of the sentence constituents are free, in other words, the main sentence constituent structure is flat. The traditional VP is not a necessary constituent in the built-up. The nomen-agent constituent can be connected first to the verb, and after that the nomen-patient constituent can be added. Even the nomen-agent and the nomen-patient constituent can be connected first in **B**-intersection, and then this new concept, John as Agent and Paul as Patient in an Action, can be specified further by **B**-intersecting with the Beat concept. Hereby the concept Beat is unified with the general action concept. Sentence constituent order is free and can be treated differently in different languages, without a difference in the conceptual semantics of the sentences. If we would also take into account extended sentences with adnominal phrases and adverbial phrases, we would have fixed hierarchies in the comprising nominal and verbal phrases, within which the adnominal has first to be connected to the noun or noun phrase, and the adverbial has first to be connected to the verb or extended verb phrase, according to the scope it has.

Since not all the participant roles in basic sentences are analysable by means of the general Agent and Patient roles, we can as well use the most specific roles for a given verb in the syntactic-semantic structure of basic sentences. In our example this would mean that we form the concepts of John as Beater and Paul as Beaten One by the binding operation and then bind these two concepts together into a smallest situation: **B**(**B**(John, Beater), **B**(Paul, Beaten One)).

In what follows we shall first formulate desiderata for any model of activation of neuronal connections and assemblies which also connectionist networks should have if they would be semantically adequate. This means that they should fit with the conceptual structures on the level of consciousness, as they are expressed here with respect to the dynamic conceptual semantics of basic sentences. Later on we shall evaluate some methods proposed in connectionist approaches for treating syntactic combinations in simple sentences with respect to their semantic adequacy. After showing them to be unsatisfactory, we shall make some suggestions about architectures of conceptual and episodic maps which provide richer models for how real neuronal networks in the brain could solve the problem of binding and of syntactic systematicity, and what role consciousness plays for understanding.

2.2.2 The desired structural properties of neuronal network processing.

Here I first formulate three desiderata for semantically adequate neuronal networks. Although existing connectionist modelling cannot live up to the second and especially the third of the desiderata, I shall suggest that human neuronal networks with some expression in consciousness can and must fulfil the desiderata.

First desideratum:

All the inclusions and non-empty intersections on the growing sets of experiences and construed situations in **DCS** must be reflected in **CNM** in the relationships between activation patterns that actually come up on conceptual maps in processing a sentence and in perception of suitable situations. Also there we have relationships of inclusion, union, and overlap.

A conceptually indicative activation pattern can be widely distributed over a compound network consisting of several parallel conceptual maps. Such maps can be maps devoted each to a different aspect of visual processing and maps representing categorisation of input to other sensory modalities. For example, all green things activate patterns that have at least in common the activation of the pattern that is indicative for green. Activation patterns due to reception of input data of green apples have also an intersection in common with activation patterns for input data of other apples, and fruit in general, and an intersection with activation patterns for other input of round things, and of green things.

With perception we find widely distributed activation patterns. The activation for seeing a car, for example, is spread over maps for form, movement, colour, and pre-motor maps for the action of driving a car. The activation of such parallel lower level maps being areas of perceptive modalities, of motoric activities, and other proprioceptic feelings can be composed together in a compound map. A state of activation of the compound map can be represented by a vector consisting of a concatenation of the vectors representing the activation states of the submaps. This can be a new input vector for higher level maps, where then a more compact activation pattern corresponding to cars, or to green apples can be found, which is merely locally distributed. Each submap, for example, can be understood according to the model of Kohonen maps, the compound map as a compound of Kohonen maps, and the higher maps can be again organised according to the Kohonen-model (Bartsch 1996, 1998): On a non-compound higher map we don't have intersections or inclusions between activation patterns corresponding to concept union and hyponomy. The concept indicating patterns there are not widely distributive, rather locally distributive over several neurones. Each concept indicating assembly of neurones is centred around a neurone which matches strongest with the kind of input, the best matching unit (BMU). Thus for roses and for tulips different assemblies would be indicative, though with the strong activation of the rose-BMU and its surrounding also the tulip-BMU and its surrounding might be weakly activated due to the general flower properties which roses and tulips have in common. The general concept of a flower would be 'represented' on a higher order map in which the overall flower input is registered in contrast with tree-input, bush-input and other classes that are given as input in contrast with flowers. All this input can be contrasted with animal-input and thus on a higher map we get contrasting BMU-surroundings indicative for the category of plants and for the category of animals. Thus more or less deep taxonomies might get formed by the availability of conceptual maps, whereby the compound input from the lower maps leads to classifications of the next higher map. The lower and more compound the maps are, i.e. the closer the maps are to the perceptive and proprioceptive maps, including motoric maps, the wider is the distribution of the activation patterns that are indicative for certain inputs. The wider the distribution, the better it is possible that similarity between concepts can be indicated by overlap of patters and that implicational relationships, hypo-relationships in taxonomies, can be indicated by pattern inclusion.

We require of conceptual maps that with the activation by data input from apples also data input from fruit generally is activated, and that on the lower compound maps the activation patterns for apples and pears overlap more than those for apples and bananas. On the other hand the higher level concepts have a function in that they create a stronger coherence between certain features by tying these together into a more general concept. Without this, activation would spread in almost equal strength to all sides; this would result in masses of unwelcome associations of lower concepts or features, because every two feature indicators that are activated together repeatedly get connected and the activation spreads from one to the other. A higher level concept A can bind features {a,b,c} on a lower level together by a circuit of activation between the A-indicator and the a,b, and c indicators (cf. chapter 3). Hereby the indicators for a,b, and c get stronger activated, highlighted or concentrated, among the other possible associative activations going from a to d, d to e, b to f, f to h and c to k and k to l and l to m, for example. This concentration also makes it possible to associate a linguistic item to {a,b,c} via concept indicator A, because hereby this set of features is delineated and highlighted among other feature indicating assemblies which are activated weaker than {a,b,c}. In chapter 3, I shall show how this principle of circuit forming with indicators on higher level plays a role in assigning lexical items to concept indicators and to processing constituents of basic sentences. A layered architecture of conceptual maps with circuit activation will enable us to combine both requirements, the one of having available logical relationships between concept indicators, and the one of singling out aspects of conceptual representation from a mass of possible associations and relating this selection to associated lexical items in forming a describing sentence.

Generally, intersections in **DCS** have to be mapped into unions in **CNM**, unions have to be mapped onto intersections and inclusions have to be inverted. In this way we should find potential activation patterns in **CNM** corresponding with the similarity based structures in **DCS**.

The mapping from **DCS** into **CNM** should be:

(1) $f(A \cap B) = f(A) \cup f(B) = A' \cup B'$, for example: for green round things we should find the activation pattern for green and the one for round activated, i.e. the patterns for green round things all should include the union of these two patterns.

(2*) $f(A \cup B) = f(A) \cap f(B) = A' \cap B'$, for example, for something that might be a donkey or a horse, is very similar to both, we should find the intersection of the horse and the donkey pattern activated.

(3*) $f(A \supset B) = f(A) \subset f(B) = A' \subset B'$, for example, if roses are flowers, then we should with the rose pattern also find the flower pattern activated.

If we represent the activation pattern on the assembly of units in a net by a vector, union there is vector addition and intersection is simple vector multiplication. The workings of the network should be such that the projection f from semantic conceptual structures to activation patterns on connectionist network structures is satisfied. But this requirement is too strong in this form: It holds for intersections of sets of satisfaction situations (see 1) because with growing specialisation the set of features grows that characterise the situation type, and these features somehow must be represented in the assemblies of activated neurones the net comes up with on the basis of the previous learning process. But it does not generally hold for unions of sets of satisfaction situations (see 2*): If we here intersect the features, too little remains to restrict the set of alternative satisfaction situations. Take, for example, 'apple or banana'. The intersection of features is not more than the features for 'fruit'. These would mean that at best the activated assembly in the intersection would accept any satisfaction situation in which a fruit is prominent. Therefore we need in the indicator of the alternative 'apple or banana' both indicator assemblies to be preserved, but in a way that this assembly does not imply one of these, except if the other one has already been excluded. If the net somehow should represent the satisfaction conditions for the expression, the net would have to represent a test procedure of the form: Test whether the apple-activation pattern is activated. Test whether the banana activation pattern is activated. If at least one is yes, then signal true, i.e. approve of the expression 'apple or banana'. The 'or' is not represented on the net, rather it means that a test operation has to be executed over the net. We thus have to exclude (2*), except for disjunctions in recognising things which are very similar, i.e. more similar to each other than to other kinds of things, such as apples and some fairly roundish kind of pears.

Also (3*) is problematic: A widely distributed pattern for rose, should comprise the one for flower. In this, (3*) his is right. But if the pattern for John as Agent should comprise the one for Agent, we get into problems. Our semantically motivated desideratum would have the consequence that with the activation of the Agent-pattern also the Agent-part in the indicator of Paul, who has the patient role in the sentence *John beats Paul*, would be activated. This would make it impossible to solve the problem of syntactic systematicity, namely that the network distinguished "John beats Paul" from "Paul beats John". The reversal of the inclusion relationship becomes problematic as soon as individual concepts are involved, which at different times in their life history have different properties and different roles which are incompatible at one point in time. In their indicator the different roles they have taken are represented as possibilities. This could be overcome if we would use an architecture of several maps in which the Agent-concept would also be represented on a higher-level map instead of merely being distributed over parts of all individual concept indicators that contain an Agent aspect and all indicators of verbs that have a frame with an Agent-participant. The higher level Agent-indicator then has to activate the agent-part in the John-indicator additionally in a recurrent circuit such that we get here a stronger activation than for the agent-part in the Paul-indicator, and the patient part in the John-indicator. The architecture can be such that it separates stages in the process of understanding according to major constituent structures, which are finally bound together in an episodic map, which has an expression on the level of consciousness. In chapter 3 such an architecture will be discussed and proposed as a solution to the problems of binding and systematicity.

Second desideratum :

Contiguity relationships generally might best be represented by sequential associations of activation patterns involving recurrent stimulation in temporal order. A special kind of contiguity is contiguity within a sentence and within a situation. Contiguity within sentences are the syntagmatic relationships which involve syntactic systematicity. This problem for **CNM** has been discussed by a.o. Fodor 1997. In this chapter I shall concentrate on the problem of syntactic systematicity in **CNM**, because it is insufficiently dealt with in these models.

Third desideratum:

Syntactic systematicity should be reflected in **CNM** such that the semantic properties, especially implications, of sentences are preserved. Before reporting about how that problem has been tackled, I shall first attend to the following questions:

1. What should the relationships between activation patterns look like in the process of syntactic-semantic understanding?

2. Should there be an activation pattern which is the result of the process of understanding and thus would that be something like a network internal semantic representation of a sentence?

3. What should the process of understanding a sentence be in terms of network architecture and mathematical operations?

I shall deal with the questions in that order. Conceptual Semantics, as it has been sketched in the previous section, places restrictions on what should happen in neural networks, will they have any chance of performing their tasks in the process of relating linguistic expressions to external semantics, and generally of being attuned to external situations and tasks. The process of understanding in **DCS** was one of gathering constraints on possible satisfaction situations, represented by **B**-intersection on concept representing sets of situations, which amounts to cutting down on the sets in a process of specification by intersecting expansions of these sets which are stability preserving, and by binding concepts together into basic situations by unification in an act of synthesis. At least some hierarchies in constituent formation were fixed. The fixed, endocentric, constituents have to be processed first, before they as wholes can be connected to other constituents in exocentric constructions. Except for these fixed hierarchies, the processing just can proceed in any order of major exocentric sentence constituents, whereby the concepts corresponding to these major constituents being processed first, are taken as units which get bound together in a concept of a situation, or a situation-type. How could such process of sentence understanding be mirrored in connectionist networks? We shall focus on the problem of binding concepts together and on the problem of syntactic systematicity.

Ad question 1: The formation of **B**-intersections should have the formation of unions of activation patterns as correspondent, because the activation patterns can indicate collections of features, established in the process of learning on the internal net, and mostly unknown to the language user. Concept indicating assemblies get activated in causal dependence of certain input features we may or may not be conscious of. They are internal indicators of outside features which we experience according to the way in which they have been processed or 'understood' by the network on the basis of its previous learning history, and according to the processing of the context available in the situation of experience. An internal activation pattern, due to an input and its effect on the established connection weights, can be represented by a vector which has as dimension the units of the internal net. In a real neuronal network we can think of such potential activation patterns as restricting and enriching our perception of possible external semantic correspondents, which are the input of the activation pattern. These internal feature indicators are learned and unconscious. They are activated by an input example according to the connection weights built up in the learning process. Assemblies as feature indicators, i.e. groups of activated neurones, must play a role in our conscious perceptions of input examples. They must have some indirect expression in consciousness by determining how we perceive sensory input in terms of consciously perceived features and relationships in and between the input examples; they are causally involved in recognising semantic properties and relationships. This causal connection induces the formulation of semantical adequacy for processes and activated patterns on a neural network.

Under a semantic point of view we desire the following structural properties realised by fully distributive 'representations': The greater the activation assembly or area the more restrictions or features it may represent in principle. The pattern for Flower should be the intersection of the patterns for the different kinds of flowers, and every activation pattern for a specific individual flower should comprise the pattern indicating its kind. Thus the pattern for roses comprises that for flowers generally. The Rose pattern should contain the flower pattern and also encode additionally more specific Rose features. Likewise John as Agent should be indicated by a potential activation pattern that contains the one for John and the one for Agent. The processing thus has to be such that potential activation patterns are added to each other. For example, first the pattern for John should be activated and then, according to the case marking or syntactic position of the phrase and the kind of verb, the Agent pattern should be activated. Likewise the pattern for Paul should be activated, and to this the pattern for Patient should be added. Furthermore, the pattern for Beat should be activated additionally. Internally, it always has some parts that indicate a generalisation of the Agent-participant, and some other parts that indicate a generalisation of the Patient-participant. The activation of the Beat pattern will again provide some activation for the Agent pattern that had already been activated additionally to the John pattern. Likewise the activation of the Beat pattern will imply some additional activation of the Patient pattern that is also activated additionally to the Paul pattern. This overlapping of activation can be seen as a unification: the Agent part within the Beat pattern gets unified with the especially activated Agent part in the John pattern. But this simple pattern union on a (compound) neural map does not work for syntactic systematicity because with activating the Agent-pattern in connection with the John-pattern it automatically also will be activated as being a part of the Paul-pattern, and with the activation of the Patient-pattern in connection with the Paul-pattern, it automatically will be activated as part of the John -pattern. Without taking into account the temporal order of processing we are stuck with the problem of syntactic systematicity; 'John beats Paul' could not be semantically distinguished from 'Paul beats John'. We shall see in chapter 3 that it does not just require temporal order; rather an architecture of neural maps and circuits of activation between such maps should do the job for binding and syntactic systematicity.

Ad question 2: Patterns of distributed 'representation' or indication of concepts are distributed over several parallel submaps, which together form a compound map. There should, though, be a unity that connects the John pattern and the Agent pattern by binding them to each other. Such binding operation, in principle, should hold for the whole sentence: the concepts that constitute a basic sentence should be bound together in the representation of a basic situation. But note that a simple union of all the involved concept indicating patterns would not do, because it neither can bind the concepts in the way required nor can it solve the systematicity problem.

A 'representation' of the satisfaction situations for a sentence which is sensible to structure must be such that the semantically relevant processing steps are somhow still recognisable in the overall pattern. Thus, we must

(1) either claim that a resulting final 'representation' of the conceptual semantics of a sentence by means of a potential activation pattern does not exist in a meaningful way, and that the syntactically structured procedure of connecting activation patterns is all there is to sentence understanding: Then the path to the goal would be the goal itself, and the resulting activation pattern on a net would be of no interest. Instead of a final 'representation' of a proposition there then is merely a series of activation patterns, as in type (4) connectionist networks according to Ramsey (1992). But note that in this series the binding requirement must be satisfied: the concepts must be bound together in the right way such that the perception of a situation in which John beats Paul is distinguished from a situation in which John does something else and Peter is likewise present and he beats Paul. In the perception of both situations the patterns for John and Beat get activated, but only in the first the Beat-concept is bound to the John-concept. Only the first situation may be recognised by the network as a satisfaction situation for the sentence *John beats Paul.*

(2) Or, if we want to cling to representations of propositions on a connectionist network, we must see to it that there are connections between successive patterns which are not just the union of two patterns, but imply a modification of one by the other, as has been proposed by Elman (1990, 1993, Garson 1994). Methods for trying to do this, as there are Elman's vector concatenation in recurrent processes, and Smolensky's forming of tensor products of vectors, will be discussed in the following sections 2.3.1 and 2.3.3. The problem with these proposals will shown to be that the relevant semantic properties are lost in these

'representations' and therefore they, by themselves, cannot be indicative for the possible satisfaction situation of a sentence. I shall show that such 'representations' are insufficient.

Our question right now is whether a finally resulting activation state that is typical, i.e. unique, for a certain proposition, is indeed necessary for understanding a sentence and relating it in interpretation to a model or piece of world. This requirement might be too strong. If we want to have some fixed representation for a proposition, we could best take an inscription of the sentence itself, or some semantically equivalent sentence, or some imagined or produced picture of the satisfaction situation. These would be representations in the full sense of the word, and they are part of our consciousness. Such a picture can be seen as an expression or articulation of the steps in the process of understanding the respective sentence, whereby a binding operation on the neuronal network expresses itself in seeing or imagining John as Agent and Paul as Patient in the Beat situation. There is no need for a final or single activation state of a connectionist network that should be the representation of a proposition, and which would be stored as such in some memory space. To require such a representation is a relict from the traditional cognitive model, where fixed mental representations were assumed to be construed and stored in memory as something that can be used and re-used in various contexts of processing. Instead of this, our memory either may be able to link up to a describing sentence presented, or to an approximate picture of a certain situation, or it may initiate forming a new representation of this kind. Thus, if we do not want to assume that the network somehow makes copies of resultant activation states and stores them somewhere in a big mental library, we have to think of a complex neural network (architecture) as being itself modified in its connection weights in such a way that experienced episodes are kept in it implicitly such that the respective patterns can be activated again if some parts of them get activated by some appropriate stimulus. Within the state of connection weight distribution in a network architecture, including the connections to the sensors, there is no representation to be found of a particular episode or of a particular sentence, although it must have had some effect in strengthening some connections and not others. The distribution of connection weights or strengths has been changed somewhat, such that with a current appropriate input the system is stimulated to bring about the representation.

In modelling understanding we shall restrict ourselves to those parts of the process of understanding in which brain activation patterns for the syntactic-morphological concepts and lexical concepts are run through constituentwise in connecting the sentence to the brain external public situations or to imagined situations. The sentence is in this way seen as a program or series of instructions for checking whether a situation is a possible satisfaction situation for the sentence. What is then kept as a representation in our mind is the image or percept of the sentence itself in its, in principle, public connection with situations. That connection must somehow be represented implicitly in the brain in a generalised manner as routines or procedures implicit as dispositions in a connectionist superstructure that works over the neuronal maps that indicate historical and general concepts. In connectionist processing the implicit procedures merely appear as an actualised sequence of steps in the concrete realised procedures, which serves in connecting the sentence with the external semantic conditions by measuring the sequences or rather circuits of activations, which it causes during the process of understanding, against the perceptual or imagined input which is hereby either accepted or rejected as a satisfaction situation. The syntactically organised processes select the relevant aspects in the perception of satisfaction situations in two counter-running processes of activation, namely by understanding the sentence and by perceiving the situation. these two processes are measured against each other. In case of fit the sentence is satisfied by the perceived situation. The procedure need not be frozen up in any resulting state of the neuronal network. Momentary results of the procedural steps may have as a conscious expression the sudden experience of having understood parts of the sentence or having imagined or perceived a respective satisfaction situation. The circuit architecture of neural maps in chapter 3 is designed to have this property.

Ad question 3: Connectionist networks typically work with operations on vectors, which are used to represent input and output, and also can be used to represent the activation of the internal net, such that all the hidden units are represented as dimensions of a large vector, which can have the values 1 or 0, depending on the state of being activated or not. An activation pattern then can be represented by such a vector which, of course, will be very long. In this purely theoretical discussion let us assume that lexical expressions, but also compound expressions, trigger activation patterns on the neuronal networks. The patterns have been formed as results of dispositions acquired in the course of becoming acquainted with the surroundings and in learning language associated with perceiving situations and performing activities. These dispositions are implicitly encoded in the connection strengths of the network. An activation pattern comes about according to the acquired dispositions by taking in and processing appropriate input, namely input of the kind that has served previously in building up and stabilising the occurence of this pattern in reaction to the appropriate input. What was discussed above already implies that for syntactically compound expressions certain operations should give a structure to the brain process of building up the resulting conceptual semantic representation in consciousness, which a resulting activation pattern by itself cannot express. The structure cannot be found back in such a resulting pattern; rather it has to be found in the process of producing and reproducing it following the syntactic form of the representing linguistic expression. In accordance with this view we can also assume that syntactically compound expressions which have been learned as wholes in situations, may correspond to a situationally caused activation pattern which can only then be an activation pattern indicative of an understood situation, if in its production a contiguity based, i.e. a syntagmatically organised situational built-up is followed. This also would mean that a situation can only be perceived in combining situational constituents according to contiguity relationships in space and time. - Of course, such a parallel between a syntagmatically organised understanding of sentences and of situations is a speciality of human perception and understanding, not found in speechless animals. But it leaves open the possibility that animals understand situations as organised by basic contiguity relationships and a by a restricted repertoire of general concepts and individual concepts. -

2.3 The problem of systematicity and binding in simple connectionist networks

The problem of syntactic systematicity arises because simple commutative and associative addition of patterns represented by vectors consisting of ones and zeros lead to a mix-up, such that sentences like *Paul beats John* and *John beats Paul* cannot be distinguished. We shall now see what operations on vectors can do, which have been proposed to deal with syntax in connectionist models. What network architecture can do will be illustrated in the next chapter.

1. Vector concatenation with recurrent fibres

Vector concatenation means that the vectors of two sentence parts in a sequence are fed into a net in such a way that first the vector of the first part is fed in with a certain activation result on the net. Then in the second circle, by recurrent fibres, a new vector consisting of this result concatenated with the vector of the second part is fed into the net. Like with all input vectors, the concatenated vector is per dimension multiplied with each of the connection weights. The sum of all these products per connection is measured against the threshold of the receiving unit of the connection in order to cause an activation of this unit or not. - See Garson 1994 and Elman 1990 for the details of representing information built up in time by recurrent networks. - If we have the vector for John on the net and then concatenate it with the one for Agent and feed this concatenation into the net we should get a modification of the John pattern whereby in contiguity with the John pattern some assemblies of units are activated due to the Agent-part of the vector concatenation that has been fed in. This seems to solve the problem of syntactic systematicity and provide for binding. In the course of time, i.e. if John has often been described or perceived as an agent, then with just activating the John-pattern also its associated Agent-assembly will be somewhat activated. This means that we expect him to be an agent, even if that is not explicitly stated. This method, if it can be refined enough to work in general, will result in networks in which contiguity in situations and between situations is represented by contiguity or nearness on network maps, or at least by association between activation patterns. These maps could be maps representing episodes and individual concepts. The representation of general concepts then would be distributed over individual concepts. On a second map the results of the first map could be generalised such that we there would receive activation patterns of locally concentrated assemblies of units as 'representations' or indicators of general concepts. We shall look at such a model later, in a more advanced architecture. The method of using recurrent fibres is essential for treating phrases in contexts and may be useful for treating recursion in linguistic structures as substitution of phrases into the context of other phrases, a problem not dealt with in this chapter, which concentrates merely on systematicity. How far this method in a more involved architecture of networks can also be semantically satisfactory, i.e. preserves implications and secures recognition of satisfaction situations will be discussed below, in chapter 3, where circuits of activation play a role in representing constituency.

2. Vector addition and vector multiplication

It is clear that simple pattern union, i.e. vector addition, gives rise to the systematicity problem. We shall now investigate whether we can use pattern intersection, i.e. vector multiplication, as an additional means to secure binding of concepts and by this also solve the systematicity problem. Let us assume that on the neural net we have individual concepts, i.e. concepts of individuals, 'represented' by locally contiguous activation patterns which indicate properties the individual has shown in the course of time. General concepts, such as Agent or Man are 'represented' globally distributed over all the activation patterns indicating individuals that have exemplified these general concepts.

Let us consider our sentence *John beats Paul*. The activation pattern caused by the utterance part *John* causes the corresponding activation pattern on the conceptual map. The activation pattern for the concept of John contains sub-patterns for Agent as well as Patient, because John has often been experienced as agent and often as patient. This, of course, gets generalised for all humans and also animals. By the subject marking of *John* in the context of the active voice of the verb, in English its pre-verbal position, the Agent pattern is additionally activated. In fact we here have an addition of the John activation pattern and the Agent activation pattern, which can be expressed by adding the two vectors that represent the activation by John and the activation by Agent. The vectors have as dimensions all the units on the conceptual map, and the values in these dimensions are the activation values. By performing vector addition, <JOHN> + <AGENT>, in each dimension the values are added to each other with as result a new vector which is the sum from the two. On the net, the addition results in additional activation of the Agent part in the John pattern, but unfortunately also in the activation of all other of the Agent-pattern's distributed parts, thus also the Agent-part in the Paul pattern. We rather want that only the Agent part of the John pattern gets activated. In order to achieve this we might want to form the intersection between the John-pattern and the Agent-pattern, which amounts to the simple vector product <JOHN><AGENT>. Hereby we have constructed the activation pattern for John as Agent which contains all the assemblies of units that are indicative for the features which John has as far as he is an agent. Since this would mean that we loose the features of his blue eyes and blond hair, because they have nothing to do with his being an agent, we have to add the vector <JOHN> to the product. This gives us <JOHN> + <JOHN><AGENT> as a representation for John as Agent. Proceeding like this we would get as a representation for the whole sentence: <JOHN> + <JOHN><AGENT> + <BEAT> + <JOHN><AGENT><BEAT> + <PAUL><PATIENT><BEAT> + <PAUL> + <PAUL><PATIENT>.

The resulting activation pattern is just the same as what we get by simple vector addition, i.e. pattern union, which is unsatisfactory. A difference can merely be detected in the process of construction. The overlapping parts expressed by vector multiplication get additionally activated. The Beat activation pattern generally has as sub-pattern a part that is due to the agents perceived in the set of Beat situations in the learning process, and it has a part due to the patients perceived in the same situations. By multiplying the vector for John as Agent with the vector for Beat, we get those assemblies which are indicative for the features John has as a beater. Likewise we get the part of the pattern which is indicative for the features of Paul as a beaten one. We in fact have a union of the pattern for John as Agent and the Beat pattern, whereby the overlapping parts are especially activated. There is a unification of the two patterns achieved by the overlapping part, namely the Agent pattern, which is identical in both and gets by repeated activation especially activated. Likewise the activation pattern for Paul as Patient is called up by multiplication of the vectors for Paul and the one for Patient, and it is united with the pattern for John Beats. Here the especially activated Patient subpattern in the pattern for Paul as Patient is additionally activated by vector multiplication with the Beat-pattern, where a Patient subpattern is included. The result should be a pattern for the situational concept of John Beats Paul. In fact, in the resulting pattern we cannot distinguish John Beats Paul from Paul Beats John, because it does not preserve the way in which it is constructed. But if we take into account the process, the systematic semantic distinction between John beating Paul and Paul beating John can be made. In vector addition and vector multiplication we have a semantically meaningful operation, whereby the process must account for syntactic systematicity. In the resulting pattern the semantic distinction made by syntactic systematicity is lost.

Note that the sentence *Paul, John beats* likewise is understood simply in the temporal order of its constituents. No transformation is required. But here first an overview of the sentence is necessary to see from the pre-verb position that *John* is the Subject. Since *John* is such marked as the Subject, *Paul* must be the Object. Then the pattern for Paul as Patient and for John as Agent are activated, and with the union of these, the pattern of Beat is united, whereby its Agent part additionally activates the already especially activated Agent subpattern of the John pattern in construing John as Agent and the Patient subpattern of the Paul pattern in construing Paul as Patient. Also here, the overall activation pattern finally resulting does not make the necessary conceptual distinctions, though the process of construction does.

If somebody would have mainly perceived beat-situations in which John does the beating, this would strengthen his concept of John as a notorious beater and his concept of Beat situations as situations in which typically John is actively involved. But if he never has seen John as a beater, there is a problem with the above representation in terms of a sequence of vector additions and vector multiplications: Some of the intersections can be empty, i.e. the vector product can be zero. If the network has never experienced John as an Agent the product <JOHN><AGENT> will be empty, if the network has never experienced John as a beater the product <JOHN><AGENT><BEAT> will be empty. Still, somebody who never has experienced John as a beater also must be able to understand our sentence. The above representation therefore cannot be achieved as an activation pattern in the process of understanding the sentence, rather it expresses conditions the network puts on possible satisfaction situations: a satisfaction situation for the sentences has to be such that its effect on the net, i.e. the pattern it causes there, contains at least what is expressed in the above representation. Not only must the products not be empty, but the situation-specific "John as Agent" features must include the Beat-agent features:<AGENT><BEAT> <AGENT><JOHN>. This requirement cannot be represented on the net as an operation on the activation patterns that are available in the process of understanding the sentence. The products could be empty. Rather the requirement has to be stated over the network: the network may accept an input situation as satisfaction situation only if the requirement is satisfied by the activation it gets from the input. Formulated in this manner, understanding a sentence means performing a process of generating requirements to be put on the net in case of perceiving a situation as satisfaction situation. The sentence then is the formulation of these requirements which can be activated in a sequence of instructions for checking on the activation of the net when it gets situational input which it has to accept or to reject as a satisfaction situation for the sentence. We shall see above how this can be achieved in a proper architecture.

The classical position in traditional cognitive science is that there must be a semantic representation, which must exist at some moment as a whole, and a copy of it can be kept and stored in an explicit memory. If we abandon this picture of a singular representation which can be stored in a memory like in a library, we still can be content with what is achieved, because we have pictures and sentence tokens as linguistic representations in external stores on paper, in notebooks, in computers, and in libraries, and we can call them up in reality and in imagination, consciously in our minds by stimulation through tokenings of some of their parts. This does not mean that there must be an unconscious part of our mind, as a kind of library, in which they would be stored in explicit forms. Rather they merely must be implicit in the connection strengths within our neuronal networks such that they can be called up by some appropriate linguistic or situational stimulus, resulting in a process of coming up with circuits of activation patterns with some articulation in consciousness, i.e. in a mental representation, which is a sentence token or a picture in our mind. - Note that our mind is constituted by what is conscious. Everything else is merely brain and the states the brain is in. - On our linguistic and pictorial representations we can do all kinds of explicit logical operations. We have an implicit memory in the connection strengths in our neural nets, which can come up with activation patterns in causally established connections with input from linguistic utterances or situations. Of course, the temporally ordered patterns occuring in procedures, like all other kinds of routines, must in principle be results of the capacities of our neural net architectures and must have been specified in learning processes. Then the presentation of an external linguistic or pictorial representation in perception, or the imagination of such a thing, can activate the relevant procedures of conceptual processing and episodic imagery.

3. Tensor product

Smolensky (1990), discussed in Horgan and Tienson (1992 and 1997), has made a proposal to treat sentences like the above ones by means of the tensor products between vectors, and not by simple addition. The reason for using this more involved operation is the systematicity problem, namely that the resulting conceptual activation patterns for a sentence like *John beats Paul* cannot be distinguished from those for *Paul beats John*, if we just use pattern union, i.e. vector addition.

In the tensor product, on the other hand, an n-place vector and an m-place vector are connected by multiplying each value of one vector with all the values of the other vector. The result is represented in a vector with n times m dimensions, or in form of a (n,m) matrix, having n lines and m columns. For vectors representing patterns on a net with n units we get as result vectors with n times n dimensions, or (n,n) matrices. For sentences like *John beats Paul*, Smolensky forms the tensor product of the vector for John and the vector for Subject, the tensor product of the vector for Paul and the vector for Object, and the vector additions of these with the vector of Beat as verb, which is also a tensor product. The result is something like

(*) $[\mathbf{v}(\text{John}) \times \mathbf{v}(\text{subj})] + [\mathbf{v}(\text{beat}) \times \mathbf{v}(\text{verb})] + [\mathbf{v}(\text{Paul}) \times \mathbf{v}(\text{obj})]$.

Generally, the semantic vectors of words are tensor multiplied with a vector for the grammatical category they are used in. We see immediately that this combination of tensor product and vector addition does satisfy the requirement of syntactic systematicity, but does not satisfy the binding requirement. The participants are specified as agent and as patient, here as subject and object, but they are not bound to the verb, because vector addition cannot achieve this. A net with (*) as an activation pattern, which must be the criterion or indicator for accepting a possible satisfaction situation, would also accept a situation in which John beats Peter and Mary kicks Paul. The vector addition in (*) cannot secure that only a situation containing the basic situation, i.e. the smallest situation, in which John is agent, Paul is patient and the action is beating is accepted as a satisfaction situation.

Moreover there are some further problems for this 'representation' in that the tensor product does not make sense as a semantically relevant representation for John as subject, Paul as object, Beat as verb, if subject, object, verb are just concepts of linguistic theory. The activation value in each dimension of the vector for the individual concept of John is multiplied with the indicating vector of the grammatical category 'subject'. However, if a semantic concept of subject is meant, something like Agent or Carrier of a property, then each positive value of the John pattern will be multiplied with that. An assembly of such activated units may stand for a known feature of John. Assume that an assembly of activated units is indicative for John's blue eyes, then also the value of these units will be multiplied with all the values in the dimensions of the Agent-vector, and where these vector dimensions have the value zero, the John-values will be multiplied with zero and disappear altogether. That does not make sense. The result of the tensor product has no semantic significance. The pattern does not in a semantically effective way contain the one of John, nor the one of Agent as sub-patterns, and thus the activation pattern of John as an Agent does not include the activation pattern of John, nor that of Agent. It represents nothing meaningful because it does not have causal connections to John situations, nor to Agent situations. That we can retrieve the values for the Agent vector from values of the tensor product vector and the values of the John vector does not give any semantic relevance to the tensor product. The point is that the tensor product does not have semantic relevance to the causal effects of the activation patterns of the John concept or the Agent-concept in the recognition of John or of Agent-instances, which have been stabilised in learning. The net has to recognise the causal effects of input situations as being satisfaction situations of the sentence or not. For this it has to recognise the implications which the truth of the sentence has. The tensor product seems to be merely a technical device without any meaning.

Criticism of this kind, besides other points which I do not follow, has also been put forward by Fodor and McGlaughlin (1990) and Fodor (1997) against Smolensky's solution to the problem of systematicity: Fodor (1997) points out that vector composition by tensor product misses the causal connection to the possible referents of the expressions in composition. In short, semantic composition is not respected by tensor products. Compositionality of representations is not kept up in constituent combination by tensor product: BROWN COW does not partly overlap with BROWN TREE, as it should. The mental representation of Brown Cow has to include the one of Brown, as has the one of Brown Tree. That is not so in Smolensky's tensor product. Fodor (1997: 114) argues that a mental representation must be causally explanatory in the sense that "a tokening of BROWN COW will include the effects of a tokening of BROWN".

Horgan and Tienson, and also Garfield (1997), do not accept that semantic compositionality should be respected in the results of the syntactic composition of vectors and they support Smolensky's syntactic structures by claiming that they need not work compositional under the point of view of semantics and causality. It is enough, they argue, that the tensor product makes for a systematic change in the resultant pattern when we replace John as Agent by Paul as Agent, or Peter as Agent, etc. We therefore can abstract the Agent pattern by systematic substitution of John in the tensor product by Paul, by Peter, by Fred, and so on. Against this I want to point out, that if basic syntactic structure would not be semantically relevant and motivated, there would be no sense in having syntactic structures at all. In order to uphold semantic relevancy of vector composition by tensor product formation, we would have to assume that perceiving John as an Agent is already a syntactic process in which first a John-pattern and an Agent-pattern get activated, such that both corresponding concepts can be recognised as instantiated and then a syntactic step is taken such that also the result of the tensor product of both conceptual representations gets activated. This then should result in recognising the binding of both concepts as instantiated in the perceived satisfaction situation. Generally, the perception of a situation would then have to involve a syntactic process such that additionally to the result of the composition also the composing parts are preserved. This is required because the result of the syntactic composition alone does not contain these parts anymore, but we still should be able to automatically get all the semantic implications from a situation. This view on the perception of a situation as containing all parts and the syntactic-semantic processes is acceptable under the requirement of systematicity, since the process of composition distinguishes a situation of Peter beating Paul from Paul beating Peter. However, the result still has to contain all the conceptual parts in a semantically active way. The result of forming the tensor product alone cannot 'represent' the proposition expressed by the sentence; rather it has to be aided by also keeping intact the activation patterns of the composing concepts separately. This can be done in an architecture of several conceptual and sensory maps, as I shall argue below.

A formal problem of the tensor product is that the vector length is n times n, if both contributing vectors represent an activation pattern in the same net consisting of n units. A vector of the length n times n does not represent a pattern in the same net but has to be represented by a larger net with n times n units, which can be represented in form of a n times n matrix. This requires a special architecture of the network. A way out for a one-map architecture would be to have an operation of compression performed on the n times n matrix to reduce it to an n-place vector. A compression that preserves semantic compositionality would be preferred. Just taking the diagonal of the matrix would not do, because the diagonal is not more than the intersection between the composing patterns, i.e. the simple product of the vectors, which would cancel out too much. If, for example, on the assembly of activated units on which the blue of John's eyes is indicated nothing of the Agent-concept is indicated, then it is cancelled in the product John as Agent that John has blue eyes. The pattern has the value zero at these units or vector dimensions. The blue eyes of John could not be recovered from the pattern of John as Agent, even though in our perception of John as Agent this feature would remain stable. In this way, John as Agent may not be distinguishable from, for example, Peter as Agent, if such cancelling happens for several specific features of John. What remains of John in the product with the Agent-vector are merely the activated units that also are members of assemblies indicating Agent features, and of the Agent-features only those remain that are also found in the representation of John. The vector product then would amount to an intersection of features. This would be fine as a step in the syntactic-semantic process, where it not that in case John never has been experienced as an agent the agent features would be cancelled out. Whatever is the solution to this problem of forming the intersection between these activation patterns, in terms of satisfaction situations, the parallel to this intersection would mean that there is a union of John situations and Agent-situations. But such a union is much to broad, and thus is inadequate for indicating John as Agent. John as Agent is rather found in the intersection of the John-situations with the Agent-situations, which is indicated by a pattern union on neural nets. As has been pointed out this union alone would neither satisfy the binding, nor the systematicity problem.

Another semantically relevant way of reduction of the matrix would be to compress the matrix by adding up the rows of the matrix and by adding up its columns. By adding up the columns we would just get the vector of the first constituent, in our example the vector of John, multiplied with the number of ones appearing in the Agent-vector. By adding up the rows we would get the vector of the second constituent, i.e. the Agent-vector, multiplied with the number of ones in the John-vector. This would amount to an additional activation of either the John-vector or the Agent-vector. We can add up the sum of the columns with the sum of the rows, and maybe additionally with the diagonal. This way of compression of the tensor matrix would give us the activation of the pattern union with higher activation values for the activated units. This would be a semantically relevant matrix compression, since it contains both vectors of the composition, though with extra activation of the activated units. If we would normalise the positive values to 1, we get by addition of the columns simply the first vector of the tensor composition, and by doing the same for the rows we get the second vector. Adding these normalised sums of the columns and of the rows would simply come down to vector addition of the two composing vectors. With or without normalising the positive activation values to 1, these semantically adequate compressions give us a good result for John as Agent and Paul as Patient separately, but when adding up the compressed tensor products according to (*) above we loose again the syntactic systematicity. The whole boils down to pattern union and intersection, and in the non-normalised case with some more differences in the strength of activation of the positively activated units in the pattern union and intersection.

3. Conclusions with respect to a one-map architecture

The result of considering these methods is that vector addition or union of activation patterns makes sense semantically as far as the implications due to the composing parts are concerned. This also holds for adnominal constructions. Take, for example, *red ball*. The Red activation pattern contains parts that generalise over surfaces, because we mainly experience coloured surfaces. This is parallel to a so-called Surface-slot in a frame representation. The Ball-activation pattern does not contain a specific colour but a generalisation over colours on the surface generally, an aspect that in a frame representation would represent a Colour-slot for specifying a ball's surface. It further contains parts for representing a round surface, including parts that represent surface with colour in general. In the union or vector addition of these patterns of Red and of Ball, these common parts get double activation, namely by the generalised Surface-parts of the Red-pattern, and by the generalised Surface parts of the Ball-pattern. In this way there is a union between the Red- and the Ball-pattern by unification of the common subpattern for surfaces. The intersections in the union take care of the unification, which amounts to binding the concept representations together. The important point is that we take into account that activation patterns contain parts that represent typical contexts or situations of occurrence of the phenomenon to be represented. These get especially activated in syntactic and situational composition by way of unification or overlap like in the activation pattern of (John[Agent)Beat{Patient]Paul} or (Red[Surface)Ball], whereby Agent-features have been especially enhanced within the John representation and Patient-features within the Paul-representation by morpho-syntactic information that the Agent-concept applies to John and the Patient concept to Paul, and these agent and patient features have again been enhanced by intersecting with the Beat-pattern which contains generalisations over agent and patient patterns.

However, all these combinations lead to a mix up in the resulting pattern which appears as a whole on the net at a certain moment. The process of construction contains syntactic-semantic information that is lost in the resulting pattern. The resulting pattern is only relevant for keeping up information about the involved concepts, the semantic binding and syntactically expressed semantic systematicity are represented in the process, not simply in the resulting activation pattern, which therefore never can be a 'representation' of the proposition. Furthermore, the unification or binding by non-empty pattern intersection and inclusion is something that cannot be realised on one single net. In order to represent these requirements on the neural activation due to a possible satisfaction situation, and to preserve the information represented in the syntactic-semantic process we need a more involved architecture of nets. This will be the topic of the following chapter.

We can conclude that a mapping between **DCS** and connectionist networks representations in **CNM** can analyse the potential activation patterns on a net by delineating them in their parts and by analysing the processes of understanding by a network in terms of semantically relevant steps. Understanding a sentence should not differ in structure from perceiving a situation which satisfies this sentence. Thus, as far as understanding a basic sentence is concerned, syntactic structure has to be taken into account and its parallel semantic structure plays a role in perceiving and thereby understanding a situation. Perceiving implies performing a conceptual synthesis in which **B**-intersections are formed, i.e. smallest situations are construed in a sequence of specifying the situation by classification and identification. This process runs parallel to the syntactic combination of the concepts in forming the constituents of the sentences that describe the situation. Note that in the representation by **B**-intersections the result contains the syntactic built up. A situation is perceived by embedding the sensory input into a conceptual structure: we attend to the action that takes place and to the individuals which are involved in the participant roles. The process of achieving the result of perceiving a situation, or of basic sentence interpretation, must consist of these processing routines for situations, which are likewise represented by the morpho-syntactic form of the describing sentence itself and get activated by understanding the sentence. In connectionist modelling we need an architecture of maps, rather than a one-map network, in order to capture the process of understanding and make it part of the result of understanding basic sentences and situations.

Chapter 3

An architecture of episodic and conceptual maps for understanding basic sentences

In the previous section it was shown that connectionist models using only a single map for the generalisation of input data are not able to provide for semantic transparency, systematicity, and binding. It was suggested that an architecture of maps with interaction circuits between them would be able to fulfil the semantic desiderata. Also results in neurological-cognitive research suggest such an architecture. Thus Saffran and Sholl (1999) write that semantic features of a concept of an object are widely distributed over the different modalities of perception and motion, for example as visual perceptional features, kinaesthetic features, action oriented features, tactile features and auditory features. They suggest that they are interconnected via reciprocal activation. The different features of an object will not all be activated with the same strength. "Properties will be differently accessed as a function of context. So, for example, if the scenario has to do with moving, the weight of the piano is likely to be the object's most salient feature, rather than its musical properties." (1999: 263). They find that the distributed model is supported by evidence; the loss of perceptional properties or of functional properties makes a difference for which kinds of lexical items can be accessed and how partial the access is. The loss of functional properties leads to a loss of artefact concepts, the loss of perceptional properties leads to a loss of animal concepts. Saffran and Sholl cite evidence from brain damage patients and PET and fMRI scans during semantic processing tasks. Research in Neuro-cognition is relevant to modelling cognitive processes because models, which are constructed such that certain functions can be fulfilled, should at least be compatible with brain research, although they abstract from the biological aspects of the functioning of the brain.

3.1 Episodic and conceptual maps

The basic distinction we shall make is the one between one or more episodic maps on the one hand side, and several layers of conceptual maps, ordered according to the generality of concepts these maps achieve, on the other. On the episodic maps happens all that is conscious, namely perceptions of situations, including perception of linguistic utterances, and imaginations of situations and utterances. On these maps, sensory and conceptual activity meet in consciousness, namely in perception and imagination. On the other hand, everything that happens merely on the conceptual maps is unconscious, i.e. below the level of consciousness. Even thinking, as far as it emerges in consciousness, is couched in episodes, namely in the form of imagined situations or inscriptions of sentences, pictures or graphic schemata. Schemata are compositions of linguistic and pictorial representations, which serve to represent, to hold and to control in consciousness large networks of logical and conceptual relationships. The episodic maps are the ones that provide for phenomenal content due to their intricate relationship with the perceptual system. Everything that is conscious must have phenomenal content, i.e. it must be couched in some perceptual or proprioceptual medium.

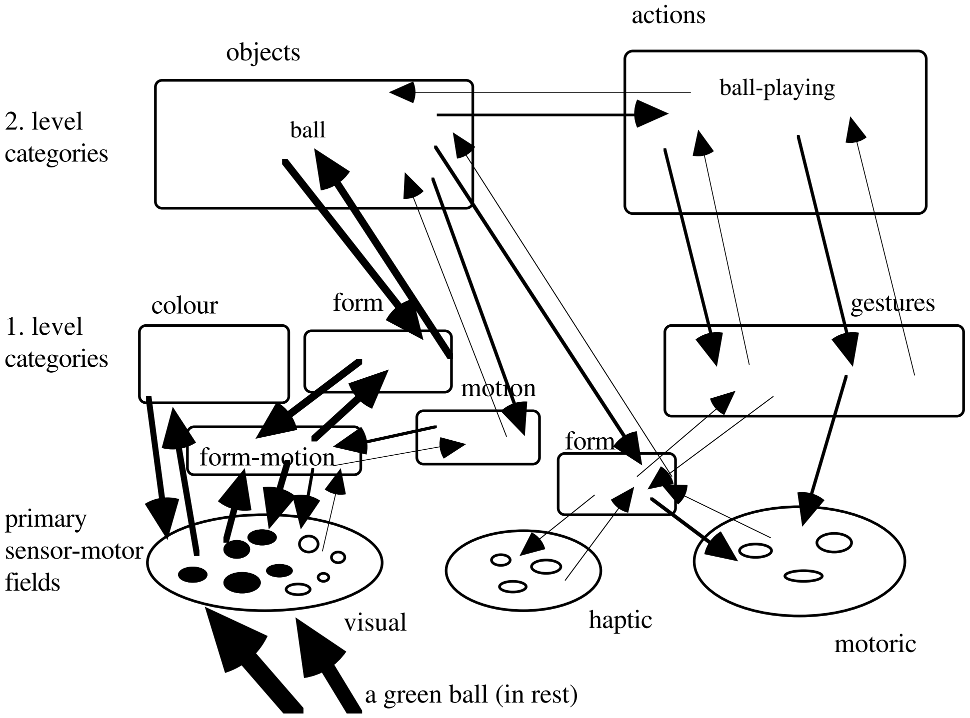
On the episodic maps phenomenal content is provided by including parts of the perceptual and proprioceptual system, especially feeling of our own motor activity. The phenomenal contents, being shaped by conceptual activity, make up feeling, perception, and imagination. The phenomenal contents are rooted in the sensory capacities of the episodic maps and include to a lower or higher degree categorisation and identification which consists in an interaction between activation on the episodic maps and on the conceptual maps. I take it that phenomenal content never is without at least some form or conceptualisation. It contains categorisation through the involvement of generalising maps, by which a perception or feeling is shaped via the neural network by the effects of previously experienced similar episodes: because of the trained, i.e. weakened or strengthened, connections, activation from an activation from a new instance of sensory reception travels most easily towards those neurones on the episodic and conceptual maps which had been previously activated by similar episodes. Hereby the sensory input of the new episodes get classified together with those previous examples, and hereby becomes a perception, a perceived episode.

As we know from brain research, there are different maps or regions for our different senses or modalities of perception: vision, hearing, touching, smelling, tasting, and also balance, motoric and other proprioceptual experience. And some of the maps are compounds consisting of several maps; for example, vision consists of a primary region through which all visual input goes, and more specialised regions such as a region for colours, a region for shapes, spatial relationships, and a region for motion and some shape and locality aspects. The specialised regions are inter-connected with the primary region. We can assume that on the primary visual region visual episodes are momentarily represented by the activation pattern caused by the visual input and the processing results achieved by the specialised visual areas. These results are categorisations as to colour, shape and movement, and they are received back on the primary visual region, which in this recurrent process becomes to be the map of visual episodes. A sensoric map, i.e. an area receiving sensoric stimuli, in interaction with categorising maps, which generalise over the different aspects of the stimuli, is called an episodic map. Thus a sensoric map is not by itself an episodic map; rather it is an episodic map only as long as it stands in interaction with categorising, i.e. conceptual maps. These are firstly modality specific conceptual maps, creating modality specific episodes, such as visual episodes, auditory episodes, etc. And these are secondly higher conceptual maps which categorise and identify individuals of our normal ontology of situations and entities. The episodic map of representing situations and objects according to our normal ontology consists thus of at least two spaces, whereby the space of modality specific episodic maps is embedded into the space of the normal ontology map. The smaller circuits, which run between the primary sensoric areas and the conceptual maps categorising aspects of the sensoric input in a modality specific way, are embedded into the larger circuits between the layer of conceptual maps of individual concepts and kind concepts and the sensoric input. There is not just a one-way path from input categorised in a sensory specific way to the higher categorisation according to the concepts of normal ontology. Rather reciprocal fibres also make the higher conceptualisation play back to the modality specific conceptualisation by injecting into the perception of the received sensoric data also such features which are imagined due to the general and specific knowledge couched in our concepts of individuals and kinds. In this way we do not just perceive certain clearly sensed parts of the front side of a house, but we see the whole front of the house, which even gets completed more or less accurately to the whole house by our power of imagination.

In the figure below the architecture activated in the visual perception of a green ball in rest is presented. Note that the circuits of activation are strongest for those aspects which are really sensed. The activation pattern for the concept "ball", which is called up by the perception of the form properties also activates in a weaker way the possible motion a ball can make and the possible actions which are the standard ways of playing with a ball, which have been experienced before. Also how it feels to touch a ball, the felt roundness, gets secondarily activated from out the ball indicating activation pattern, although the ball has not been touched. Via these secondarily activated aspects, the perception of the ball contains imagined features on the primary visual, haptic and motoric maps. Since the actually sensed visual features are stronger activated because of the (repeated) input stimuli, the imagined features are not dominant. The movement of the ball cannot not be imagined at exactly the same moments in which the resting ball is perceived. Rather the potential movement of the ball can only be imagined in alteration with the percept of the ball in rest. The imaginations on maps of different modalities can be entertained at the same time. Thus, while seeing the ball we can imagine what it is like to touch it, or to play with it in a certain manner. If the child who is observing the green ball, has the desire to move around and play, then the areas for indicating play-concepts will be activated and herewith the corresponding pre-motoric areas. Thus the concept-indicators for ball-play will be additionally activated and a ball-play action will be meditated and prepared on the pre-motoric areas, such that the child will move towards the ball and begin to play with it.

Circuit activation in perception: seeing a green ball

sensed: ●
imagined: ○



Similar to this architecture, models are developed in neuro-dynamical system research. Thus Grossberg, Mingolla and Ross (1994) have developed a neural theory of attentive visual search in which pre-attentive boundary segmentation, surface and figure ground perception, and spatial information activate certain object representations, which play back activation that helps to define (1994: 471) “the perceptual units that attentive spatial and object recognition mechanisms can further modulate and recognise during visual search.” Attentively organised search recursively applies to emergent groupings that may contain multiple items. Grossberg (1996) further explicates the object recognition mechanism by his Adaptive Resonance Theory (ART), where the bottom up activation from out stimulated sensorial maps activates a pre-learned concept representation, or rather concept indicator on a higher map, which then feeds back activation from the top down to the bottom. This top down activation, providing expectations about what is perceived, influences the perception of the received data by highlighting some features, suppressing others, and adding expected features. If the activated concept does not fit well enough with the data, the interaction circuit does not reach the synchronisation of the oscillatory activity of firing in which object recognition would result. Rather the mismatch causes an inhibition and hereby destabilisation of the circuit activity, while another competing concept takes over. It is also activated from out the data, and also possibly somewhat from out the first concept, which in some ways is similar to it. The second concept then has the chance to establish a new circuit of synchronised activity. Grossberg (1999) suggests that the circuits of resonant dynamics in between bottom up and top down activation processes plays a role in constituting conscious states in the brain; all conscious states are resonant states. He furthermore suggests that the ART-mechanisms, which have been shown in visual perception, are also operative in the other modalities. And, of course, the context-dependence of perception fits nicely into this model, because a context activates conceptual frames, which from top down influence the perception of the data presented in that context, i.e. interact with the bottom up activation from the reception of the data.

Another study in the same line of research has been done by Van Leeuwen, Steyvers, and Nooter (1997), who investigate perceptual segmentation by using a model of circuit activity organised by attractors, which are “stable limit sets of system states towards which nearby states converge”, while unstable limit states are called repellors (1997: 319). The model shows an interplay between stability and instability in chaotic behaviour. For calculating this interplay of non-linear change they use differential equations. Stability in these “discrete-time dynamic systems" means convergence towards attractors. An optimal segmentation of a perceived pattern of dots, for example, is one which is relatively stable. In one of their examples two sequences of dots form two lines which cross each other. Depending on from where one goes through the pattern, from the middle, from each one of the sides, or from the top or the bottom, the segmentation of the pattern may switch. The overall perception of the pattern as a crossing of two lines, or and X, is determined by the Gestalt-principles of proximity and good continuation. Hereby the starting points from which one takes up the pattern of dots serves in organising different attractors, between which one can switch such that “the two lines are desynchronised from each other “(1997: 337). Switching requires “an ability to disengage attention to a stable pattern by making it unstable and creating alternative stable ones” (1997: 320). Generally, switching between the interpretations of ambiguous figures is not organised by a control schedule, rather it occurs spontaneously. Perceptual destabilisation and switching are based on “deterministic chaos”, calculated by non-linear differential equations (1997: 322). They also remark, apparently following Crick and Koch (1990), that the synchronisation of neurone firing leading to oscillatory activity patterns, solves the binding problem in visual perception (1997: 322). Van Leeuwen et al. furtheron suggest that “two perceptual models could be assumed to operate in interaction: a perceptual segmentation system and a classification system. The classification system learns from the segmentation system and the segmentation system is corrected by the classification system. ...such a model would develop a bias for more stable perceptual organisation in the long run and combine this with flexibility in the short run" (1997: 342). This is a point also modelled in Grossberg et al. (1994) and in Grossberg (1996), and it also supports nicely the interaction between contiguity ordering and similarity ordering between data as the two ways of structuring growing sets of data in the process of concept formation and understanding (cf. chapter 2). Segmentation and classification in visual perception are ordering data by contiguity and similarity. In these dynamic models learning segmentation and object recognition takes place along synchronised circuit activation. Neuro-physiological and anatomical research by Kandel, Schwartz, and Jessell (1995) has supported this view, based on Hebb's conjecture, that units that are activated together strengthen their connections. Units that are connected in circuits of synchronised firing, strengthen their connections, which amounts to learning the combinations of input data, their forms and segmentations, and their classifications.

Also studies about deficiencies in perception caused by brain lesions show that binding of different aspects in perception is essential for object recognition. From experiments done with brain-damaged patients we know, for example, that there are patients who can see the colour of an object put before them, but not its form, and thus cannot identify the object by visual input. When they are allowed to touch it, though, they recognise the form by tactile input and are able to identify the object, for example, as a cup. This points to the fact that the specialised region for categorising visual form or shape is impaired.

It seems evident from case-studies by Sacks (1985) that also memory of scenes and faces, both requiring ordering of impressions by contiguity relations, has to result in activation on episodic maps, which have an expression in consciousness. An act of recollecting something from memory can be seen as similar to an act of perception or imagination. It is caused on generalising maps by some re-activation of parts of the neural net effects of previously experienced scenes. This includes maps that generalise, i.e. categorise, over spatial and temporal contiguity relationships between situations. From Sacks' case-descriptions it seems that episodic memory might function in parallel with the more analytic and abstract categorisations. It contains implicitly in connection strengths the partial orderings by contiguity relationships of the indicators of individual experienced situations. It can best be understood as the capability of ordering a new episode in space-, time-, and causal contiguity, while at the same time categorising it by comprising it together with previous episodes that are similar to it under certain perspectives. This two-fold ordering of the new with the old, keeping unimpaired coherence and stabilisation, is of course the basis for understanding in general. Parallel with the several levels of generalisation on the left side, we seem to have an implicit memory of schemata of episodes in their contiguity contexts and orderings between episodes. Parallel with the concepts on the several levels of generalisation, we have generalisations over sequences of episodes on maps of episodic memory, for example schema's of the kind "When it rains the ground becomes wet", "When a mushroom tastes bitter, I will not take it". Furthermore, parallel with the higher levels of generalisation, we have general concepts or schemata of spatial and temporal contiguity orderings, such as "in", "outside", "before", after", "at the same time", "next to", etc.

The episodic map for visual episodes is the primary visual region in interaction with the specialised regions, whereby the visual episodes consist in the primary sensoric activation of the primary region which is transmitted to the specialised regions categorising colour, shape, relative locality, and motion. From there the resulting activation patterns send back activity by reciprocal fibres to the primary map, where they cause a secondary activation, emphasising the features which have also been encountered in similar previous examples. The results of visual categorising, corresponding to features such as certain colours, forms, spatial relationships, movements must be bound together as belonging to one object or situation. I shall suggest that this happens by binding the repeated primary sensoric activation together with the categorising secondary activation in a small circuit that runs between the primary, sensoric map and the specialised, categorising maps of the respective modality. The sensoric map has to contain collections of specially dedicated neurones. They must be dedicated such that they contribute the phenomenal qualities as the medium in which all our experiences are expressed.

Crick and Koch (1990, 1992) have proposed that 40 hertz oscillations in the visual cortex solve the binding problem in the way that the common frequency of oscillation of neurone firing binds the visual aspects together, as they are distributively processed on the specialised visual maps. They even claim that this unifying frequency of oscillation must be the causal base of consciousness in perception. Of course, here as elders, the criticism by Chalmers (1996: 116) applies, namely that Crick and Koch cannot answer the key-question of why these oscillations, and hereby why binding, should be accompanied by conscious experience. However, I shall pick up the idea of binding by unifying frequencies, in proposing an architecture of constituent structures of sentences and situational perceptions by circuits of activation between several episodic and conceptual maps. The circuits might be individualised by certain rhythms, based on certain interferences of frequencies. Thus small circuits with certain frequencies in the firing of the neurones, which are involved within a larger circuit, determine the rhythm in the firing of the neurones in the larger circuit. This, of course, is speculation, but it would make sense.

For each sensory and proprioceptual modality we can assume a similar compound structure with a primary map and specialised categorising maps. Thus we have an episodic map for each of the modalities; we have visual, auditory, olfactory, tast+touch/texture episodes, arm and hand movement+touch, leg and foot movement+touch episodes. These specialised episodic maps are further bound together by activation interaction, via connections, with conceptual maps for identifying individuals and categorising objects and events into kinds of objects, events, activities, or actions. The interaction between the parallel episodic maps in a compound can happen via trained connections between these compound maps, or it can go via the conceptual maps in which groups of modality specific episodes are bound together as aspects of one object or one situation. The categorising activations on the higher conceptual maps are played back over the lower conceptual maps, to the primary sensoric maps. Such a large circuit of interaction headed by an activation on a conceptual map, defines an episode on the compound of sensoric maps. If the activation pattern on the higher level conceptual map indicates an individual concept, a kind concept, and a situational concept such as an action concept, we have an episode of an individual, being of a certain kind, involved in a certain action. The compound of sensoric maps in interaction with higher conceptual maps functions as an episodic map on the level of our normal ontology of individuals and situations. It is non-specific as to modality, but includes different modalities. In what follows I shall call this a normal episodic map.

The episodes on the episodic maps, caused there in recollection stimulated by activation from the generalising maps, are ordered in space and time via the connections to those generalising maps that show the activation effects of spatial and temporal contiguity relationships. They consist of compounds of parallel visual, auditory, touch and motoric maps and a higher level conceptual map categorising the compound input into spatial and temporal relationships between objects and events. Space and time are then for us instantiations of such spatial and temporal relationships on episodic maps. We experience space and time only in episodes involving spatial and temporal relationships. There are first and second order maps for storing the effects of contiguity relationships, which have been found in previously encountered constellations of situations.

Recognition of contiguity can be impaired by certain kinds of brain damage. Such cases have been reported by Sacks (1985). Somebody who is blind, wholly or for one half of his visual field, because of a certain kind of damage of his visual regions in the brain, cannot remember any scene which originally had been perceived in the part of the visual field now blind. Thus a patient who was asked to tell what buildings he passed on his way could only report about the buildings on his right side, and when returning on the same way he could report about the buildings on the other side, which were then on his right side. The left side scenes he could not remember, even if he had seen them before the damage of the visual regions in the right hemisphere. He was able to see and recognise only basic geometrical forms and patterns, and separate features. In so far he could report about the buildings, but he could not connect features in contiguity, and therefore he neither saw nor remembered faces or scenes, nor could he detect any emotional expression of faces. At most he could identify people by a single clear geometric feature, a very prominent form of the head, or the nose, or the chin, or a very perspicuous mark on the cheek. Because his right side visual region was impaired by a tumour, he could not perceive, imagine, or remember faces; and when he was asked to take his hat he grasped for his wife's head, as something of likewise round shape. He compensated his inability to perceive and imagine visual contiguity relationships in objects, scenes, and between scenes by getting things and actions into order through organising them in contiguity by means of singing tunes and songs, which he more or less did continually. His world was organised by musical contiguity relationships, replacing visual contiguity. Auditory contiguity got apparently connected with motoric contiguity of hand and foot movement, and thus he could organise his actions. On his auditory episodic maps he employed an especially rich gamut of contiguity relationships in which he embedded sequences of actions he had to do regularly, like washing and clothing himself. We can explain this behaviour by assuming that conscious recollection of sequences of actions and situations requires appropriate activation on episodic maps of perception and imagination in contiguity, which apparently are to be found mainly on the right hemisphere. Since the patient in question could find his way walking in his familiar surroundings, we can assume that he still had implicit memory of previously seen things and scenes, possibly on the right hemisphere, which does not come to consciousness on the visual episodic maps, which are now damaged. As far as his space-time orientation was conscious it was ordered sequentially in songs accompanying his activities. Also other cases reported by Sacks point to the conclusion that the right hemisphere is especially suited for perceiving and imagining contiguity within scenes, and in history, while the left side is stronger equipped for conceptual order, higher order concepts, and abstract thinking.

Visual episodes get connected by repeated co-stimulation with episodes processed at the same time in other sensory regions, such as those of sound, touch, smell, taste, motoric experience. Hereby these experiences are coherently bound into the unity of an object or situation. Thus, when seeing a plane in the sky we expect to hear typical aeroplane-noise, and the other way around, when we hear a typical noise in the sky from a certain direction we look up and expect to see there the aeroplane which causes the noise. A little boy I know, of one year and three month, who speaks no other word than 'ae-o' is able to connect the sight of an aeroplane and the noise of an aeroplane by using the word 'ae-o' when he sees or hears an aeroplane. Combining the several sensory perspectives of perceiving an object by establishing and strengthening connections between neurones of the different sensor-specialised episodic maps is part of the way to achieve the objectivity of an episode.

Objectivity is not achieved by just one sensory impression, but rather amounts to a connection of different kinds of sensory impressions bound together into a single object, linked by contiguity to other aspects of the situation and other episodes, and categorised together with other contiguous sensory impressions of different kinds on superimposed or 'higher' maps. Also these categorisations of different sensory kind must be bound together by neuronal connections established and strengthened in experience. This must be the case because we are able to draw implications from concepts like 'aeroplane' to the kind of noise, kind of form, kind of movement, kind of material typical for an aeroplane.

In what follows I shall simplify this very involved picture quite a lot in order to be able to design a map architecture of the relationship between situational episodes and the descriptions of these episodes in language. For the sake of representational simplicity, I throw together into one episodic map the various episodic maps of different modality, the activations of which have as expressions in consciousness the sensory impressions we are aware of in our perceptions and imaginations.

The notions "different layers or spaces of maps", and "higher" and "lower" maps, are used as notions of logical and cognitive organisation. It does not mean that these have to be different layers of the cortex, though that certainly would be a possibility of physical realisation of some regions of the different layers of maps. Future improvements on MRI- and PET-scan technics, showing small differences in depth of cortex activity, might give information about the position of layers of maps in and on the cortex.

Activation patterns on layers of conceptual activity, by themselves, cannot to be understood as representations of concepts. Such patterns have a concept indicating value only by contrast relationships to other such patterns, and by their relationships to perceptions and imaginations due to stimuli which originate externally from examples, experienced in a history of learning. Across different learning individuals and even across the life history of one individual the identity of such a pattern is not relevant, and does not even exist. Merely on the sensoric areas we must assume pre-established collections of dedicated neurones which are fixed in their connections to certain measuring devices of the input sensors. Different patterns on conceptual maps, due to the different initial situations of learning systems, the difference in their learning history, and the changes due to progressive learning experiences, will have the same discriminatory value, i.e. the same conceptual value, if the established contrasts and the relationships to the kinds of input are constant. This conceptual indicator value of a pattern is given by these two aspects and not by itself, i.e. not by the set of activated neurones alone. The identity of a neural pattern on a conceptual map is not necessary for indicating the same concept. However, on the episodic maps, in their basic areas which are directly sensitive to sensoric input, there should be a strong relationship between the phenomenal quality aspects of perceptions and the involved fibres leading to specialised sensoric cells and the spatial constellations of these, in order to have coherent phenomenal qualities within the experiences of things. Thus, specific frequencies of light should consistently correspond to certain qualities of colours, and for that one needs groups of neurones that are especially dedicated to certain kinds of sensoric input by being fixed in their connection to the respective measuring devices of the sensors.

Contrast between patterns has originated from contrast between examples and thus provides a structural representation of contrast relationships between examples. Concepts are merely implicit in the distribution of the weights of neuronal connections and their relation to sensors. A concept can be seen as a disposition to accept and create examples as belonging to a set of examples representing a concept. It is thus a capacity of the neuronal system and its relationship to the sensory and motoric organs. Conceptual ability is the ability to recognise, order, and create examples. The only representations of concepts are sets of examples and sets of general sentences we hold true, and which we are aware of in the conscious expressions of activations of the situational and linguistic episodic maps. Activation patterns on conceptual maps are merely concept indicators; they are what they are merely by their causal relationships to certain input examples, stabilised by adjusting connection strengths in the course of learning.

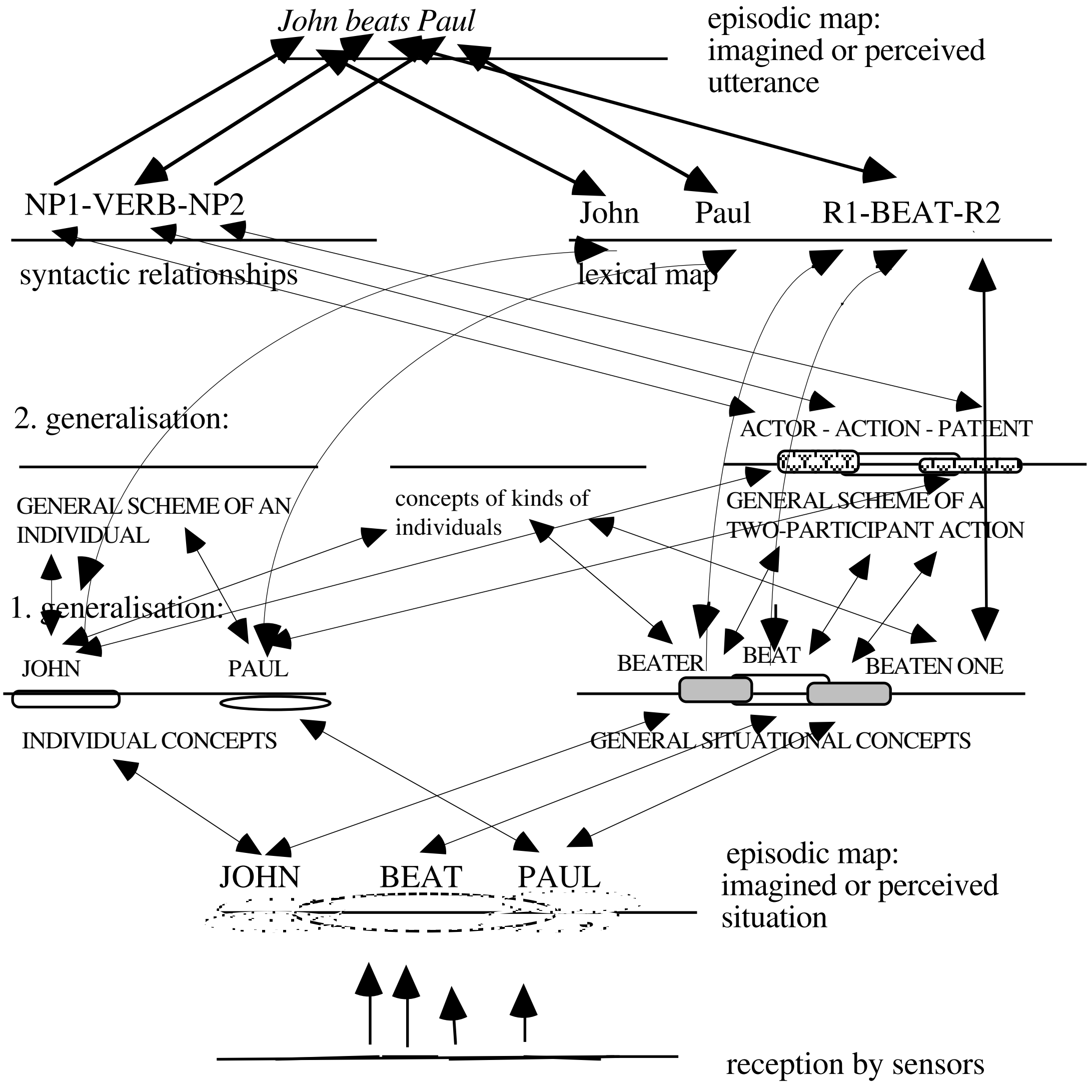
Only neuronal activity which happens on the episodic maps has an expression in terms of phenomenal content in consciousness. But that does not mean that phenomenal content consists just out of phenomenal qualities. Rather such qualities are never experienced as such, they always are imbedded in reactions from the conceptual maps, even if these are not yet stabilised. What happens on the episodic maps is created either as perception by sensory stimulation in space-time contiguity together with conceptual categorisation provided by the connections with the higher conceptual maps, or it is created as imagination stimulated by some other input, by processes based on inclusion and association, that takes place on maps of general and individual concepts and contiguity orderings with respect to other episodes. Imagination is especially triggered by linguistic input via the process of understanding, and perception can be modified by such imagination. Perception is caused by reception of data via the sensors, combined with registering of contexts consisting in bodily movements, and in interaction with the categorisations of the data: The data are sent from the episodic maps to the generalising maps; then the categorised data from the generalising maps are sent back via reciprocal fibres to the episodic maps. The activation from out the generalising maps calls up additional imagination on the episodic maps by adding further phenomenal content triggered by the generalised aspects that normally come into the picture under other perspectives than those under which the data are received by the external stimuli in this case of perception. Thus every perception of a categorised object or situation automatically is enriched by aspects which are imagined, and which are got from the effects of previous inputs which were contiguous with input similar to the current input. Which aspects are triggered depends on context, as has been pointed out in the example by Saffran and Sholl (1999) about moving a piano, where its heavy weight is to be considered, but not its musical aspects.

Imagination can best be understood as a derivative of previous perceptions: New situations or objects can be imagined in analogy to old perceptions, or there can be imagined really new ones of a kind that has never been perceived before. The process of imagination modifies previous or current perceptions by taking parts from other previous or current perceptions. The modifications are achieved by substituting parts of perceptions or previous imaginations by parts from other perceptions or imaginations, keeping intact contiguity relationships as they are classified in general situational concepts and individual concepts. Substitution is an operation with direct conscious results, which can be performed also consciously on representations due to activation on situational or linguistic episodic maps. On the neuronal level it is a change in association by blocking certain connections serving for associating contiguous content and by opening up other connections by which new content is associated to the contextual content which is kept constant in substitution. The results of substitution are conscious, and thus they can be controlled and evaluated as normal, as comical, as surprising or as revealing some deeper truth. General and individual concepts respectively characterise and individualise the imagination by being bound together in the examples generated on the episodic maps in analogy to previous experiences. The analogy is construed on sets of examples by the operation of substitution mentioned above. Substitution is constrained by coherence, i.e. by the contiguity relationships found in general and individual concepts. Note that imagination in the arts often reaches special effects of surprise and pleasure by breaking these coherence constraints. This means making a new connection against the normal expectations and current associations. Thus we can find on surrealistic pictures the strangest creatures, such as pigs with wings and a trunk; or we find an "artistic statement" presented as a toilet bowl used as a vase.

3.2 Understanding situations and linguistic expressions

That syntactic processing is not done by a separate linguistic organ or module but rather by an interaction of different brain areas involved in neuro-cognitive -perceptive and -motoric processes, some with linguistic task specifications, seems to be a widely accepted position. Hagoort, Brown, and Oosterhout (1999: 305) write: "it is most likely that both parsing and grammatical encoding are based on a concerted action of a number of different brain areas with their own relative specialisation. These relative specialisations can be memory requirements for establishing long-distance structural relations, the retrieval of a verb's argument structure and other lexical-syntactic information, the use of implicit knowledge about the structural constraints in a particular language for grouping words into phrases, etc." In what follows, the process of understanding a linguistic expression, and the reversed process of describing a perceived situation, shall be analysed in its functional components by an architecture of episodic and conceptual maps presented below. In the text we shall provide more detailed descriptions of the architecture than our figure of the architecture of understanding depicts. The figure is kept simple in order not to destroy the clarity of a global overview.

Circuit activation in understanding:



Explanation of the architecture

Typical of the architecture is a high degree of interconnectivity, which opposes the modularity view, as it has been proposed by Fodor and others, whereby processing takes place in a series of isolated modules; the processing result from one module is transferred to the next module for further processing until the final result of the serial order of processing steps is reached. Against this serial modularity, Mundale (1998) points out that, for example, the visual processing regions are very much interconnected. This also holds for other processing, such that "the degree of interconnectivity to the extend that it represents information exchange, also challenges claims of informational encapsulation (1998: 137)."

Mundale further writes: "But the discovery of the degree of connectivity in the brain, especially the prevalence of collateral and feedback connections between brain regions, cast some doubt on whether processing in individual brain regions is as encapsulated as Fodor suggests. Paul Churchland (1992) draws attention to the prevalence of recurrent connections, for example, in his negative assessment of the modularity hypothesis (1998: 138)."

On the episodic maps, by way of activation patterns and their interactions with conceptual maps, we generate representations of realisations of properties and realisations of contiguity relationships as they have appeared in the situational and/or linguistic input data, and as they have been modified by the circuit activation from higher levels of generalisation. On the first level of basic ontological generalisation, there occur activation patterns of properties, activation patterns of contiguity relationships in specific individual concepts, and activation patterns of the contiguity relationships in specific types of actions, events, and processes. Contiguity consists in space-time relationships, causal relationships, means-ends relationships, process-result relationships, action-participant-circumstance relationships, and others, which as general types of relationships have activation patterns on the second or on higher levels of generalisation.

We have seen in the previous chapter on mathematical modelling of neural networks that vector operations such as addition (union), multiplication (intersection), and tensor-product are insufficient as operations for combining activation patterns in a semantically adequate way, because these operations do not achieve binding and/or syntactic-semantic systematicity while keeping intact semantic implications between perceptual contents. The map architecture fares better since it has additional operations between maps (the episodic maps, 1. general maps, 2. general maps) due to the connections between these maps, which give rise to small and broad circuits of activation. The circles have slightly different starting points and are thus in different phases, and they possibly are in different frequencies of neuronal firing because of different connection lengths, different connection strength, and different activation states of neurones. In some of the circuits an episodic map with its primary sensory regions is partially included, and hereby the results of these circles of activation get a phenomenal expression in consciousness whenever they hit the episodic map. This means that the binding-operation that is performed as a unification, achieved by means of a circuit of activation, has a conscious expression in perception and/or imagination: we can imagine John as beater, Paul as beaten one, both being bound within a single act of beating. The constituent circuits keep running, repeatedly triggered from out the activation on the situational maps, and/or from the auditorily or visually based linguistic episodic maps, possibly recurrently reactivated by repeated reception of slightly different data, until the circles die out or get inhibited in order to make room for new acts of perception and imagination.

The above architecture has some similarity with Edelman's (1992: 132) scheme of higher-order consciousness, which not only contains a global component of perception and proprioception, but additionally contains "internal homeostatic systems" in the brain stem, hypothalamus, and autonomic centres which register internal states and values of the Self, and correlate them in the hippocampus, amygdala, and septum with activation received from the world signals through the primary and secondary cortex for sight, hearing, touch and other senses, where current perceptual categorisation takes place. These categorisations happen in interaction with the linguistic areas of Broca and Wernicke, whereby a semantic bootstrapping of the percepts is achieved. The combined input from the evaluative and the perceptive component is sent forward to a memory of value-category correlations established by previous experience in the frontal, temporal and parietal cortex. The activation there is played back to the perceptual part of the cortex and to the language parts, and in reciprocal fibres those parts again promote activation back towards the category-value memory. The circuits of interaction between linguistic-semantic areas and the value-category areas establish, according to Edelman, the higher order consciousness, in which understanding of the self, past, and future are achieved, while the circuits of interaction between the perceptual areas and the value-category memory establish primary consciousness. Chalmers (1996: 116-117) comments Edelman's model in the following way: "The central element of his theory involves re-entrant neural circuits by which perceptual signals can be conceptually categorised before they contribute to memory. Perceptual information and internal state interact in a subtle way ... to give rise to 'primary consciousness.' Though Edelman's model is very instructive for the relationship between perception, evaluation, memory, and language, Chalmers remarks that it does not make clear "how all this processing should give rise to conscious experience." Edelman's theory, though it does not explain consciousness after all, supports the architecture of understanding proposed presently, which can be embedded into his model. Edelman himself embeds the perceptive-conceptual understanding of a situation into a larger circuit with emotional, emotive and Self-expressing modules. I have omitted these broader aspects for the sake of simplicity of the model, although I think that they are very important.

The essential feature of our architecture, sketched in the figure above, is that there are circuits of activation which hit the primary sensorial and motoric maps: Firstly, there are small circuits in which activation goes via connections from one map to another, on which an activation results, depending on the strength of connections, on threshold values, and on the activation vector received from the first map. The vector of the resulting activation on the second map is fed into the connections leading back to the first map and results in an activation there, depending on the connection strength of the reciprocal fibres and the thresholds of the neurones on this map. Technically this can be modelled in connectionist models by using compounds of maps and a modification of Elman's (1995) SRN-technics for taking into account contextual effects: in a Simple Recurrent Network, a vector consisting of a concatenation of a vector with n context-units and m input units is fed into the net, whereby the pattern on the net of hidden units at time (t-1) is fed into the n context units in concatenation with the new input at time t in the m input units. Thus previous context can be used in processing new input. Such a network with recurrent fibres is a dynamic system, which means "that [its] state at any given point in time is some function which reflects their prior state." (Elman 1995: 203) Here we rather take instead of the prior state the activation states on the conceptual maps which can be fed into the lower maps and finally the sensorial fields via reciprocal fibres. The activation states on the conceptual maps are caused from the current input coming from the sensorial maps and the memory as it is based on previous experience, encoded in the connection strengths between the neurones involved actively in the previous experiences. The vectors representing activation states on the conceptual maps hit the lower maps and hereby emphasise feature indicators established by previous input of the same kind. Hereby the perception is built up from reception of input data together with the activation from activated concept indicators, which have been built up through previous experience of similar data.

Secondly, there are larger circuits in which in-between maps get activated and send along the resulting vectors on the connections available. The vectors are mathematical representations of the activation patterns on the respective maps. In figure 1 a small circuit is represented by an arrow pointing in both directions, in order not to overload the figure with lines. Such an arrow stands for two fibres running in opposite directions, i.e. represents two fibres, reciprocal to each other. The same holds for the curved lines between maps.

An example of a small circle on this rather coarse architecture is the following: by reception of visual data on the primary visual area an activation is caused there and categorised according to the specific visually categorising maps. Activation vectors of this sensorial map, which is by these (modality internal) categorising activation circuits a visual episodic map, are sent to the first map of normal ontological generalisation on which patterns results, depending on these vectors, the strengths of the connections involved, and the previous state of neuronal activation. If there has been achieved some stabilisation in the learning process, the arising activation pattern will be the momentarily activated effect of the conceptual schema, or frame, indicating a generalisation over beat-episodes: <The Beater - The Beating - The Beaten One>. A vector representing this concept indicating pattern is fed back into the modality specific episodic maps which together form the basic episodic map of perception and imagination. This feed-back serves to outline the pattern of the episode, the specific beat-situation, by highlighting the Beat-features above others which also might be present in the visual field. This feed back also serves to add to the perceived features those features which were regularly experienced in beat-situations up to now, even if they are not encoded in the present sensory input. By this circuit we have achieved the neuronal pattern of the perception of the specific beat-situation as a Beat-situation, which is a product of the sensory reception, its modality specific categorisation, its general ontological categorisation, and the enriching imagination due to the state of the conceptual system at a certain period of the learning history. In this way, perception in a fully developed conceptual system includes categorisation on the first level of normal ontological generalisation. This means that we perceive the world in its normal ontology of situations (events, states, processes, actions, activities) and individuals involved in these situations as participants (Cf. about the basic ontology a.o. Bartsch 1995).

3.2.1 Understanding a situation

The generalising maps of normal ontology together with their neuronal connections are the neuronal backbone of concepts of episodes, i.e. situational concepts, and concepts of entities, i.e. individual concepts. Both kinds of concepts involve some generalisation, since also an individual concept requires generalisation over several appearances of the individual, besides the fixation of contiguity relationships between these appearances. A general concept of a type of situation, for example of a beat-situation, requires generalisation over several specific realisations of the situation type, which includes that contiguity relationships within the situation are generalised, such as the ones between the typical participants of an action and the action, and the different stages of the action. Such generalisation, at a certain stage of stabilisation, results in the activation pattern of the action, which is a pattern indicating the action frame or action schema by virtue of its circuits of connections serving for inter-activation with the patterns of the types of participants. This means that not the activation pattern itself can be understood as a neural representation of a beat-schema; rather this activation pattern has to be taken together with the concept indicators of the kinds of role participants, each including the indicator for the participant role and the kind of individual that occurs in this role. The activation pattern of the most general <Agent-Action-Patient> schema has the value as an indicator of such a schema only by having connections with the general schema of individual concepts on the second map of generalisation. In order not to overload figure 1 with too many lines, I have omitted intermediate generalisations which determine the Agent-participant as an individual able to act, which is more specific than the general schema of an individual concept. I also have omitted the small circuits connecting the Agent- and the Patient-participant of the activation pattern of the general Action-schema with the general schema of an individual concept, by which these participants are determined as being individuals.

An example of a large circuit, running in two directions and involving internally several small circuits, is the one from the episodic map to the first map of generalisation for individual concepts, activating the pattern of the individual concept JOHN, whereby in the John-pattern the patterns of Agent-features are also activated, which become highlighted by the small circuit connecting with the second map of generalisation. By input from the first map of generalisation the Agent-part of the general <Agent-Action-Patient > pattern on the second map of generalisation is activated, which in turn in a small circuit highlights the agent pattern on the first level of generalisation. The <Agent-Action-Patient> schema has also been activated by being fed with a vector from the pattern of the <Beater-Beat-Beaten one> schema of the first map of generalisation. From both directions, clockwise and against the clock, the activation goes back to the episodic map, which still might be activated via repeated reception of the visual episodic data, which are slightly different in the course of the event. A large circuit and the small circuits within it might in this way get into a harmonised state of activation. The complex of activation circles amounts to a constituent analysis of a larger constituent into smaller ones.

We can imagine that this harmonisation between different circuits might happen by getting into a state of a certain rhythm of neuronal firing in the circuit, induced interference of the rhythms or frequencies of the constituting circuits. Combining interfering frequencies of neuronal firing results in an overall rhythmic pattern for the combination. This is an modification of the proposal by Crick and Koch (1990, 1992) for binding between different apexes in visual perception. This rhythmic state, which binds together different constituent circuits, then can be seen as the neuronal background of the perception of the episode, i.e. of the expression of the neuronal activation in consciousness. Of course, this all is highly speculative, but at this state of the art we can't do more than speculate by designing models which are compatible with scientific findings of neurological brain-research and which contain features that in an organised fashion correspond to semantically motivated desiderata discussed in the previous chapter about syntactic-semantic compositionality, and binding.

Circuits can be taken as that which corresponds to semantically motivated constituents of a situation or of a sentence, for example as the constituent 'John as agent in beating', the constituent 'Paul as patient in beating', and the maximal constituent, i.e. the whole situation of 'John beating Paul'. The two major constituents would have to be represented by circuits which are in slightly different phases. The maximal constituent then could have a rhythm calculated from the two constituent frequencies in interference. Since the circuits of the constituents keep operating at the same time, hopefully without getting mixed up in a muddle, the systematicity-problem would be solved by this model: We are able to distinguish John as the agent from Paul as the patient. The binding-problem, furthermore, requires that both are bound together with the action into a smallest situation. The action includes the frame of role-participants, namely Agent and Patient, or more specifically Beater and Beaten one, with whom John and Paul have to be identified respectively. This happens because on the episodic map the parts activated by the visual input of aspects pertaining to the beater and beaten one get highlighted and possibly enriched by activation of additional neurones by reciprocal input from the first map of generalisation. These parts, representing John as agent and Paul as patient respectively, coincide in an overlap with the patterns of the beater and beaten one, which are parts or constituents of the overall activation-pattern of the visual situational input. At the same time, this situational input is determined as a <beater-beat-beaten one> situation by the small circuit of the right-hand side. In this way, binding amounts to a unification by overlap of activation patterns on the episodic map, whereby the composing parts are kept separated by their respective circuits still in operation. Hereby we have a model in which systematicity, binding, and semantic composition are kept intact in understanding an episode, i.e. in the perception of a situation.

Note that this architecture supports the different alternative constituent compositions described in the dynamic conceptual semantics treatment of understanding a basic sentence (cf. Chap. 2). For example, we can form a situational constituent <John as agent, Paul as patient> unified in a circuit activation with the indicator of the general <Agent-Action-Patient> schema. This then can be unified by a circuit activation with the indicator of the Beat-schema, such that the action is specified as John beating Paul. The same mechanism of unification by binding together in activation circuits is made use of in understanding a basic sentence as describing a possible situation.

The relationship between concepts and lemmas, i.e. lexical items with a certain syntactic context specification, which represent a concept in language, is one of association learned in linguistically guided concept formation. With an activation pattern serving as a concept indicator, several other indicators for concepts are activated, which are related to the first one by similarity or contiguity relationships. Also their corresponding lemmas are activated by association. The question is how the right lemma is selected for the concept at issue. Generally we can assume that the concept which is most strongly activated selects its associated lemma most strongly. Levelt (1999: 96) refers for this selection to Luce's rule: "Lemma selection becomes a probabilistic affair. During any smallest interval in time the probability of selecting the target lemma is the degree of activation devided by the total activation of all active lemmas in the stratum. This is called 'Luce's rule'." In our model we can state how it comes that the chosen concept is stronger activated than the others which are associated with it. The reason is that the chosen concept is involved in recurrent circuits that repeatedly activate the concept indicator. This means that the concept in question is the chosen one. The other associated concepts merely receive stray activation from the chosen concept indicator. They are not involved in the constituent forming process of understanding the situation and in the constituent processing of the sentence in question.

3.2.2 Understanding a sentence

The linguistic part of figure 1 has been explicated very partially only. Utterances are also experienced on episodic maps. Parts of auditory and articulatory episodic maps become linguistic episodic maps by growing and strengthening connections to maps of linguistic concepts pertaining to categorisations of phonetic regularities, and to morphological/lexical, to morphological/syntactic regularities, and to regular syntactic contiguity relations. By virtue of circuits of recurrent activation, the linguistic data on the episodic map get categorised by these generalising maps of linguistic concepts. Therefore these linguistically specialised auditory and articulatory episodic maps can be called linguistic episodic maps. A linguistic utterance is a linguistic episode including the situation of speech, as well as the previous linguistic context which is already understood. We can be conscious of several aspects of this whole linguistic episode. Other aspects may play a role below consciousness and influence the emotional uptake of the linguistic episode. The perception of the utterance itself is formed by reception through the sensors which causes an activation on auditory episodic maps, which gets categorised via recurrent fibres by categorisation according to phonetic-intonational patterns and according to syntactic-morphological patterns represented on maps of linguistic generalisation. The map for phonetic categories and the map for their interpretation as phonemes as well as the maps of articulatory gestures have not been included in this architecture for the sake of simplification. This part is better explored by Levelt (1999).

In lexical maps we find for each lexical item a generalisation over its typical linguistic contexts, which means that proper names, for example, include a generalisation over their sentential contexts which we can express on the level of linguistic theory by saying that they are nominal terms (NP's). This means that in simple sentences they occur with verbs or other predicative expressions, including two- and three-place prepositions or postpositions, in the participant role-places indicated by case or syntactic position. Binding and linguistic systematicity is achieved by recurrent activation along small and larger circuits including the linguistic episodic maps, and it is achieved in a parallel fashion by circuits including the situational episodic maps. By smaller and lager circles of activation the linguistic phrase *John beats Paul* is analysed such that the individual concept of John is unified with the Beater-concept, and the individual concept of Paul with the <Beaten-one> concept, and the verb with the concept of the Beat-action.

Linguistic episodic maps and situational episodic maps are connected by association between linguistic concepts and semantic concepts. For basic sentences, syntactic and morphological marking of NP's are associated with participant roles, depending on the type of verb content involved. Lexical items in their syntactic category are associated with semantic, and also pragmatic, concepts.

In our example the individual concept of John is activated at the same time by circuits via the lexical item *John* and via the grammatical role-assignment by the NP1 marker (morphological case and/or syntactic position of the word *John*), connecting with the participant role AGENT in the general action scheme which is instantiated linguistically by the verb *beat*. This double activation results in a partial individual concept pattern of John as Agent, and by further activation of the Beater-part in the activation of the general concept BEAT, the <Beater-Beat-Beaten one> pattern, we have the activation of JOHN AS A BEATER spread over several maps in small and larger circuits, which even may cause a subpattern becoming part of the individual concept pattern of JOHN representing him as a beater: If such stories or experiences about John occur repeatedly, he will be characterised as a beater by a part of the JOHN-activation pattern, which is strongly connected to the general concept BEATER represented by the Beater-part of the general concept of a Beat-situation. Likewise we get an activation pattern for Paul as the beaten one as a partial individual concept.

The activation pattern of the verb *beat* on the lexical map activates by association via strengthened connections the <Beater-Beat-Beaten one> activation pattern that indicates the BEAT-concept by virtue of its connections to generalisations of individual concepts, to the <Agent-Action-Patient> generalisation, and to data of Beat-situations. At the same time the verb *beat* activates the general ACTION-pattern which is co-activated with the linguistic categorisation VERB. Though not all verbs are action-verbs, this is the prototypical category of a verb, and thus its activation pattern is strongly activated. The semantically based ACTION-pattern and the linguistically based VERB-pattern hereby get their associative connection strengthened. If this happens often, the expectation gets stronger that a verb represents an action and that an action is represented by a verb.

Parallel with the syntactic-semantic binding on the linguistic episodic maps, we get binding on the situational episodic maps. Finally the partial individual concepts and the general concept get bound together on an episodic map, in analogy with previous situations which effected activation patterns via their causally established connections to received situational input of satisfaction situations in association with their linguistic input, in the learning history of the language user. On the episodic maps happens the binding of the activations coming from the indicators of the general and individual concepts into the situations and objects, as we perceive them and imagine them. The activation, coming from the first level of generalisation via connections, which have been built up by previous experiences of the same kind, activates strongest those parts of the episodic maps which were previously activated by Beat-situations, John-situations, and Paul-situations. However, we do not simply get a union of these activation areas. Rather a circle of activation from the first map of generalisation, consisting of a <John as Agent>-vector and of a Beater-vector, fed into the connections and hereby multiplied with the strength values of the connections, results on the episodic map in a unified pattern, in which the parts corresponding to the Agent aspects are most strongly activated. The <John-as-Agent> activation on the episodic map partly overlaps with the pattern caused by the input coming from the activation indicating the Beat-concept on the fist map of generalisation. The same happens for the activation from the pattern of the concept of Paul-as-Patient and the pattern of the concept of a Beaten-one. On the episodic map we thus get an activation comparable to a perception of a situation in which John beats Paul. This can amount to a conscious act of imagination of such a situation. From such an imagination, or rather from its neuro-physiological base, originates a new circle of activation, similar to what happens in perception, but without the process of activation coming from sensory input. The imagination generally is weaker than a perception, but can be strengthened by repeated circles, possibly also fed by renewed processing of the sentence describing the situation. Imagination is supported by our implicit general memory of previous situations of this kind, and can also be exemplified by conscious recollection of such a situation, involving the episodic memory.

In principle, an architecture as the one described should exist in our neuronal networks, which makes use of small differences in the timing of circuits of activation, frequencies and rhythms of firings, inhibitions, strengths of neuronal connections, and the expressive power of episodic maps, causing conscious awareness in perception and imagination. On the linguistic episodic map we become conscious of understanding a sound-event as sentence-utterance with a certain meaning. On the situational episodic map we become conscious of the perception, imagination or recollection of a situation, often additionally partly represented by an associated linguistic description. This consciousness at the outer ends of the largely unconscious process of understanding, which are the cutting edge between world and brain, makes control possible.

The implicitness of our memory of concepts, words, situations and objects leaves an essential task to consciousness in imagination and perception. The processing itself cannot be observed, but its products or results can be observed in order to be controlled by checking against in a process that runs the opposite way: The processes marked by (1) and (1') below are reversed. The percept or imagination at the end of process (1) is conscious and thus can be perceived or imagined as new input in the reversed process (1'), which is the control process. If it does not lead back to the likewise conscious starting point of process (1), the word or phrase that had to be interpreted, the result of process (1) is found to be a mistake.

(1) imagining or perceiving linguistic utterance e ---> activation of indicator of the linguistic concept of the uttered expression ---> activation of semantic concept indicator ---> imagining or perceiving example E for the concept

(1') imagining or perceiving E ---> activation of semantic concept indicator ---> activation of indicator of linguistic concept of expression ---> imagining or perceiving utterance e.

The same is done in controlling whether we found the right linguistic utterance for expressing a certain image or percept. Here we start with (1') and control by running through (1). An example, which has really happened, is the following. In a discussion about different kinds of dogs I want to report some fact about a certain sub-species of dogs, which I know by being acquainted with three dogs of this kind in my neighbourhood:

(1') I imagine the three beagles ---> activation of the indicator of the concept BEAGLE (---> realising problem in finding the right word) ---> activation of the indicator of the word beetle ---> imagining or uttering *beetle*

(1) Imagining or perceiving the utterance *beetle* ---> activation of the indicator of the concept BEETLE ---> imagining a beetle.

I realise the conflict between imagining beagles and imagining beetles, or possibly a reaction of non-fit is already signalled from the conceptual level, even though the concepts themselves are not conscious. This conflict, resulting in control, shows me that the utterance *beetle* was a mistake. The similarity in sound between *beetle* and *beagle* helps me find the right word after a few seconds. Or in another scenario: my partners in conversation find the right word for me, being helped by the similarity in sound and the knowledge that I must be talking about a sub-species of dogs.

Connectionist modelling today still seems to be far away from getting working models of such architectures, due to problems of getting easily into chaotic and unstable behaviour because activations get mixed up in uncontrollable ways. Partly this is due to the small number of units representing neurones, the simplification of the connection strengths and inhibitions, and the choice of mathematical functions for representing activation processes and the updating of the connection strengths. Also because in today's modelling on serial computers, fine time-differences cannot be used effectively, it is necessary to build into the process retardations in-between the processing of constituents in order to keep the circuits apart, and to get a controllable interference between circuits, as it is needed in running large circuits. Such an architecture could model the main functional steps in 1. perceiving and understanding situations by characterising and identifying them, 2. describing linguistically how the situation or object is perceived and understood, and 3. understanding simple positive sentences by linking them to their conceptual semantic structure and finally to a perception or imagination of a situation that can satisfy them.

It is worthwhile to point out the main differences between the above map architecture and the system proposed by Levelt, Roelofs, and Meyer (1999) for modelling the lexical access in speech production. Theirs is essentially a feed forward activation-spreading model, consisting of two subsystems bound together, a conceptually driven system from concepts to lemmas (lexical items at syntactic property nodes), and an articulatory system from lemmas to articulated words, including context-dependence of articulation. They take concepts to be nodes in a conceptual network, whereby the relations between nodes are relations between concepts. They do not take lexical concepts as being represented by semantic features, since there is no evidence "that speakers tend to produce hypernyms of intended target words" (LRM 1999: 4). This is not a good argument, though, because if there is no reason to produce the hypernym, i.e. if it is not asked for, why should speakers do it. They tend to produce those words that provide the relevant contrasts in a set of objects at hand, and without a restricting context they will either produce the most specific lexical item, for example *poodle* if they are well informed about breeds of dogs, and think their listener is likewise, or otherwise they use the basic category word *dog*, as reported in studies by E. Rosch about what counts as basic categories. This simply is a question of what is relevant in a pragmatic setting, it does not say anything about whether hypernyms are included in an neural network indicator of a concept or not. Also in our present study no explicit semantic feature representation is assumed for concepts, rather occurrent widely distributed patterns on lower neural maps, and locally concentrated distributed concept indicators get established. From the activated units the further propagation of activation along established connections proceeds. These units are not semantic features. As was pointed out in previous sections, hypernym inclusion is a desideratum on a semantically interesting neural network, but it is only structurally explicated in the networks which have an architecture of several layers of lower and higher conceptual maps. Nodes for lexical concepts, as in LRM 1999, are much to local for a neural network representation, but on a more abstract level of cognitive representation they, of course, are acceptable for certain purposes of scientific modelling. But note that in using such representations one abstracts from such interesting phenomena as feature-probability, vagueness, and robustness of concepts. Furthermore, it is problematic that the LRM-model separates concepts from their syntactic properties. Linguistically expressed concepts are learned from their syntactic and situational contexts and thus in a sense carry their linguistic and situational context with them, a fact that is captured by representations in the form of frames and schemata. The model also omits the relationship between concepts and percepts, and it is strictly serial, something that cannot be upheld if one wants to understand the relationship between sensory input, percept and concept, and the broader relationship between situations and their descriptions in form of basic sentences. Their model thus abstracts from aspects that are important and explanatory for understanding production and interpretation under a semantic point of view.

3.3 The Binding-operation on different levels of description

We shall now compare what binding means on different levels of description, the logical level, the conceptual level, and the map-architecture level.

1. Logically, binding means unification: *x Beats y & x=John & y=Paul.* This form conventionally expresses by first and second argument who is the beater and who is the beaten one, respectively.

2. In Dynamic Conceptual Semantics, there is assumed a synthesis by the binding-operation (**B**-operation) within a smallest situation type, making use of the general role-concepts of Agent and Patient and a constituent order, out of several possible ones:

B(John, Agent) = John as Agent; **B**(Paul, Patient) = Paul as Patient;

B((**B**(John. Agent), Beat), **B**(Paul, Patient)).

Per constituent the **B**-operation forms the set of possible satisfaction situations containing the smallest situation received by synthesis of the components. In linguistic and semantic theory, it is of some generalising and explanatory value to use the general role-concepts. However, in the process of understanding situations and sentences, we can in principle do without the general role-concepts and just stick to the lexically coded specific role-concepts of a relationship or action, here the less abstract experiential concepts of the Beater and the Beaten One. The synthesis in forming a basic situation is possible by the power of imagination. Imagining can be described by the operation of substitution, or creating an analogy: In a previously experienced Beat-situation the specific agent there, the beater, is now replaced by John and the specific patient there is replaced by Paul. In order to get an imagined Beat-situation there has to be some remembered, previously perceived Beat-situation as a starter. Then other situations can be construed from perceived or imagined situations by substitution. In substitution, the Beater-aspects of the situation are preserved, while the specific individual features of Peter, the previous beater, are replaced by those of John. If there have been several previous Beat-situations with different beaters, this operation of substitution can also be understood as a sequence of first eliminating the individual concepts of Peter and the other beaters from the previous Beat-situation and then unifying the preserved general Beater-features with the individual concept of John. Unification thus makes use of generalisations over role-participants, by which we get the roles in their connections to the general characterisations of the types of role participant, for example, that the role is fulfilled by an animate individual, or a human individual, or even a fairly specific kind of individual, for example an elder woman is the typical filler for the role of grandmother in the grandmother-grandchild relationship. In the above example, the new role participant, the individual concept John, is unified with the agent-role concept, specified in the Beat-schema by the general concept of an individual. Binding is achieved by unification, which is based on substitution.

3. In the map-architecture, binding within a constituent is achieved through the embedding of an assembly of activated neurones on a sensory, and hereby episodic, map into a set of small and larger circuits of activation between several conceptual maps. By having on an episodic map activation patterns which are embedded into such a set of circuits containing activation of categorisation and contiguity relationships on generalising maps, we get as the conscious expression of these activation circuits the perceptions or imaginations of a scene, which hereby is understood in its basic categorisation and contiguity relationships. This holds for perceiving a sound-event as a linguistically understood meaningful utterance or speech act, as well as for all other acts of perceiving of situational data as a situation.

In figure 1 above, we see what unification could look like in circuits of activation. There we have assumed that a concept of a Beating-frame has already been formed, witnessed by the concept indicating activation pattern on a conceptual map. In the course of concept formation, on the other hand, the step of relative grading down the individual features of previous Beater-experiences is taken each time that a new Beat-situation with other individual features is perceived, whereby the effects of repeated features get strengthened and can be stabilised in activation patterns on a higher map of generalisation. - These activation patterns indicate role-concepts, by virtue of their activation relationships to individual concept indicating patterns, to patterns indicating generalisations over these, and to the pattern indicating the general concept of an individual. - The down grading is caused by the substitutions of the specific beater and the specific beaten one in experienced situations. It makes for the difference between the previous situations and the new situation. This down grading and final elimination of specific individual concept features on a higher level of generalisation is the process of generalising towards the general Beat-frame. This frame is implicit in the connection strengths of the neural network as its disposition to realise certain concept indicating activation patterns on conceptual maps and finally on episodic maps. If somebody is prevented by some neurological defect from generalising on higher levels, he still can perform the most basic step of substitution and thus understand a new Beat-situation on the basis of one or more previous ones by linking the new to the previous one as being similar by what is preserved, and as being different by that which is substituted. There are cases reported by Sacks (1985) which might be explained by assuming just this capability of substitution intact, while having a great deficit in generalising and forming categories, especially more abstract concepts. Sacks reports about people who are very good in narrative understanding and play acting, while they lack generalising intelligence (Sacks 1985: Part Four, case 21). These patients seem to operate well on a level of concrete scenes and especially actions, though they have no theoretic understanding.

3.4 Episodic memory in an architecture of neural maps

Episodic or historical memory has been called “declarative memory” in psychological and neurobiological research. I shall not follow this terminology because “declarative” suggests that recall from this kind of memory is necessarily expressed in language. This is not necessarily so. History and separate episodes can also be recalled in series of imaginations, and can be expressed by drawing, painting or by selecting images. It has been guessed that visual perception and imagery employ the same neuronal areas, and this is also assumed in this essay. Interestingly enough, Kreiman, Koch, and Fried (2000) have found rich evidence that this conjecture is right. They observed that visual recall of objects, animals, and faces in imagery involve for 88% the same neurones which also fire, at a slightly stronger rate, in the visual perception of the respective items. This involves areas in the sensory visual areas as well as areas in the amygdala, the hypocampus, the parahypocampal gyrus and the entorhinal cortex. By contrast analysis they showed that the selectivity of activated neurones was almost identical for the respective different input data in vision and in their later recall in imagery. The small differences may be due to the fact that on the one hand, more details are perceived in vision which are not all recalled, and on the other hand, that some extra neurones might be necessary for memory retrieval, and that some details might be added in imagery which where not part of the original perception.

Episodic memory has not been included in the architecture of figure 1, where individual concepts were represented in a simplifying manner on the generalising conceptual maps, although they have to be grounded in episodic memory. On the other hand, episodic memory as a function is supported by the architecture of episodic and conceptual maps. We never remember an episode in all its concrete details. This is so because not all details have been categorised in the original perception. The episode normally gets categorised under a few perspectives. When we remember the episode we can remember it from out the categorisations that had been achieved under these perspectives, and we might imagine other additional categorisations under perspectives which we apply later in the act of recollection. These other categorisations then do not originate from the primary categorisation of the specific previously perceived situation but are added, or imagined, by associating with categorisations that had been applied in other previous situations of this kind. This is what makes that recollections can be the subject of doubt, or even are false in some details. We also can remember a situation or individual from out its place within a surrounding of space-time contiguity.

In the proposed architecture of neural maps, semantic memory resides in the connectivity between conceptual maps on different levels of generalisation and the sensory areas of episodic maps. Semantic memory is the ability to remember something in general. Episodic memory rather it enables the recollection of a specific episode. The specificity, however, is not due to remembering the totality of features of detail. Rather it means that the episode is placed in its specific space-time surrounding within the context other episodes. Thus an episode is remembered as having occurred at a certain place and at a certain time. The recollected episode is experienced within a set of spatial and temporal relationships and thus within a history of contiguity relationships between episodes. Such contiguity relationships must reside in the connections between those generalising maps on which we find the indicators of relational concepts of space and time, with sensoric areas of episodic maps in which concrete space- and time relationships are represented, as there are in the parts of the primary and secondary visual areas in interaction with the visual areas specialised for the categorisation of form and movement.

The space-time contiguity of the parts and of the realised properties of a situation has to be the reference schema in which the categorisations and relationships within an episode are bound together. Current episodic memory involves activation of such contiguity schemas and their connections with conceptual maps. This episode-internal contiguity is part of the more or less specified situational concept characterising the remembered episode. The episode would still be a more or less general situational concept, were it not that it is integrated within the contiguity of the life-history of an individual. Its external contiguity with other episodes in individual histories makes it a individualised episode and not just a situational concept. Episodic memory must be organised by such space-time, and also other, contiguity patterns.

Like the <AGENT-ACTION-PATIENT> schema relates the individual John and Paul within one act or basic situation, the temporal schema's BEFORE, AFTER, AT THE SAME TIME, and likewise the spatial relational schema's, relate situations to each other within one situational constellation. By being part of a set of circuits of activation which include temporal and spatial relational concepts being bound with other situations, a situation becomes understood as a historical episode. Thus in order to implicitly encode situational concepts and action frames and to achieve a historic, episodic memory, our architecture of figure 1 needs to be extended by episodic memory maps connected to the conceptual maps that generalise temporal and spatial contiguity relationships between situations. Between these episodic memory maps and conceptual maps there run circuits of activation which, when hitting the episodic maps involving the sensory areas, constitute our perception or imaginative recollection of episodes, which are related to each other by temporal and spatial relationships. The episodic memory maps must include activation patterns which are indicators of space-time schemas of reference indices for historical entities, namely for concrete individuals, situations, and constellations of situations. They all must be connected to concept indicators on the generalising conceptual maps. In episodic memory, personally experienced situations are placed into the space-time contiguity that forms our life history. But we have also knowledge of the history of other individuals and historical entities. The history of others, individuals, groups, societies, and states, is remembered by placing situations and constellations of situations into temporal and spatial order, according to the narratives that make up the public history. This latter knowledge is totally mediated by language, but also partly carried into the areas of situational concepts and the episodic memory by understanding and imagining the narrated episodes.

Semantic memory of situational concepts is due to having experienced several situations of a certain type and keeping the effects of these accumulated in the growth and strength of the connections between the neural conceptual maps and of the connections with the sensoric areas. In the same way, we also can remember a single very peculiar and astonishing episode. Its neural effect on connections and its emotional marking can be comparably strong. This is so because the surprise and emotion increases the attention directed towards this single situation and might, by repeated rehearsal, promote the increase of connection strength. The repeated thinking about this situation, while imagining it by activation of the episodic maps in a process of wondering about it and trying to understand it, includes repeated activation of the connections involved and thus might have the same, or even a stronger cumulative effect than repeated experiences of a less surprising type of situation. This situation gets then implicitly encoded as a fairly specific situational concept, which is fixed as a historic situation within ones life history, by the circuit interaction with the space-time reference system of the episodic memory maps. The episodic map is not a ordered collection of episode representations, rather it merely shows activation patterns which function as systems of reference indicators, and which stand in interaction with conceptual maps, and especially with those that indicate generalisations of spatial and temporal relationships.

Also episodic memory must be implicit in the strengths of connections between maps of indicating situational and action concepts, which get their internal contiguity ordering by virtue of their connections to maps of temporal and spatial concepts, and get embedded into external contiguity ordering by additional connections to space-time reference schemas indicated on episodic memory maps, and get their perceptual and imaginary realisation via their connections to sensory and motoric areas of space and time perception and proprioception. As such this memory can have effects for actions and perceptional orientation via connections between conceptual and motoric maps. Implicit memory, general and episodic, without current involvement of activation on episodic maps is memory without conscious recall in recollection. It can be activated on conceptual and episodic memory maps and thus, even without conscious remembering, it can have effects on motoric maps. It then has effects on motor activity and attitudes, though it does not have an explicit representational expression on situational or linguistic episodic maps. Recollection, on the other hand, includes not merely implicit episodic memory but also recall, i.e. representation via episodic maps. This recall can be impaired in patients while they still have available implicit memory which automatically guides their movements in familiar surroundings. They find their way, but are not able to report about the surroundings or identify pictures about the surroundings. Such cases have been reported by Sacks (1985). Also the so-called unconscious, which gets revealed in Psycho-analysis, is implicit and in this way as a set of dispositions. It becomes explicit in recollections, which are conscious acts of remembering if triggered appropriately and if inhibitions are deblocked.

What is explicit for episodic memory are its results, i.e. the recollections, which are created by involving episodic maps. These provide not only for the phenomenal content of our perception of properties of basic situations and individuals, but they also provide for the phenomenal content of spatial and temporal relationships, which are constituted by our sensoric and motoric experiences, and generally of our proprioception in actions and all kind of activities and feelings. Memory, whether semantic or episodic, is unconscious and largely implicit. Its products, however, perceived and imagined examples for general concepts, and recollections in current episodic memory are conscious. Like all perceptions and imaginations they are formed as representations on episodic maps, which involve sensory, emotional, and pre-motoric areas in interaction with conceptual maps for the characterisation of objects and situations, and with episodic memory maps for their identification within the contiguity ordering.

Thinking in terms of the training and of the causality of neural map architectures gives us a way of understanding the cognitive structures of experience, as they might be rooted in neural effects of bodily activity. Of course, it does not help us to understand the basic gap between neuronal activation and phenomenal experience, the hardest problem in consciousness research. Our understanding of this relationship so far is only possible as a structural and functional understanding. We make use of assumptions about bridging between operations in consciousness and their publicly controlled adequacy and correctness with respect to the world on the one hand side, and brain structures and processes on the other. Some of the latter have been established by neurological research, others, in this state of the art, must be assumed under the point of view of semantic and practical adequacy. Future neurological research might give some independent confirmation to neural architectures, which have been postulated in order to serve cognitive abilities.

Neural net activations give rise to all kinds of associated activation patterns. This, by itself, would lead to very loose associative thinking, which is less valuable for acquiring orientation and knowledge about the world. Therefore the question arises as to what makes it possible to attend to what is relevant and is at issue in a certain situation. Here consciousness enables us to focus our attention to relevant selections. The attention to objects, situations and constellations between situations consists in running recurrently those larger activation circuits that embed other smaller circuits, including sensoric areas, such that we have episodic experiences. Our perception and imagination, including recollection from memory, is steered by default concentrators of attention, which are circuits of activation involving indicators of individual concepts, of general concepts of individuals (kind concepts), and of situational concepts. There are, as kinds of standards, larger circuits running as concentrators between these smaller circuits or concentrators, in order to enable us to really imagine and perceive situations, or constellations between situations, instead of floating along in loose associations. As has been posited, the circuits play a role as constituent conceptualisations in interaction with the sensory areas and thus constitute an episode on what has been called episodic maps. We are aware of constitually ordered episodes, situations, constellations of situations, and whole histories, and even theories comprising situations by laws in an explanatory manner. We are not caught in a stream of impressions, where we would wander aimlessly from one impression to the other, from one property to the other, from one individual to the other, or from one situation to the other. The larger activation circuits bind the lower ones together in a kind of constituent structure, and thus keep attention directed, thus preventing loose associations.

Concentrators, i.e. activation circuits, can be activated by perspectives, which are opened by questions such as: What kind of instrument is this? How can this event be explained? What situation is wanted for X in order to succeed in school? What properties are relevant for being able to achieve this goal? Questions express perspectives under which the answers can be found. The perspectives can be seen as higher order concepts, under which the concepts fall that are possible answers to the question. Thus, under the perspective of Colour, expressed by the question "Which colour is this?", fall all the colour concepts, such as red, blue, green, etc. We can think of a perspective, in terms of neural networks, as a certain activated area of concept indicators, by putting the neurones in these areas in a state of slight arousal and thus lowering their activation threshold, and by pre-activating certain connections such that they are more sensible to stimuli, or by temporarily grading down other connections to make these somewhat inhibited. Concepts of objects or situations have to be found and concentrated on, and the perspective determines a selection to attend to when looking for these. Perspectives constrain the area of attention, and then the activation circuits in these areas determine what can come to consciousness when considering this or that question. Higher order concepts play a role in forming concentrators over widely distributed activations on lower maps, which by themselves can flow out to all sides. From there the larger circuits organise the lower ones into one large constituent interaction which runs in itself without getting lost into all directions. This must be the basis of selection by attention and emergence of an episode in consciousness.

Chapter 4

Evaluation, Thought, Imagery, and Language

In this chapter it is argued that thinking presupposes consciousness, because it deals with representations and their evaluations. It is shown that this further leads to the assumption that thinking has to take place in the medium of imagery or of language, or in both. In whatever medium, thinking takes place on episodic maps. In consequence, the notions of a hidden mind and the unconscious are reduced to neurological structures and processes. A hidden mind as a collection of representations and rules for manipulating them just does not exist. Representations are not hidden, rather they are in consciousness emerging from episodic maps in their interaction with conceptual maps, taking place in activation circuits between these maps. Procedures for manipulating representations can be thought of as partially learned from publicly endorsed relationships between perceived representations, such as they occur in inferences, substitutions, and other kinds of transformations. As learned procedures on public representations, on objects and on actions, they are parts of the procedural memory implicit in neuronal connection strengths. This might be so for relationships between linguistic representations, such as the ones between declarative sentences and the corresponding questions. Such transformational relationships can be learned as patterns. Other relationships between representations, such as inference rules, for example modus ponens, can be exemplified and experienced in reflection on the conscious results of the inclusion, intersection, and associative relationships between activation patterns and circuits, and the respective implications caused in the perception of situations. Such a cognitive base for logical relationships between representations, and also for many substitution relationships, can be the reason for an a priori knowledge about their public endorsement. In such cases, a public rule is not a mere convention, rather it is cognitively based.

4.1 The relationship between understanding and interpretation

The purpose of this chapter is first to show that judgement must be conscious and even presupposes self-consciousness. Judgement is entertaining a representation together with a truth claim, substantiated by a truth evaluation of basic sentences with respect to the world, in coherence with narratives and theories describing the world. Perception is the basis for the evaluation of basic sentences, and I shall show that perception presupposes a notion of self as being in space and time. I shall reconstruct the experience of space and time from out proprioception in co-ordination with self-perception from the outside. By being in space and time a perceived situation, and a percept generally, is embedded in a coherent way into the contiguity of space and time, which can be partially described by narratives and theories. Secondly, by discussing interpretation and evaluation of sentences and representations in general, this chapter provides a background for examining the relationship between language and thought, and especially between language and the content of propositional attitudes.

Interpretation binds understanding of sentences to the world in acts of reference and evaluation. It is the public application of the result of understanding, and subject to public control. The public control on this application indirectly constrains our process of understanding such that it does conform to the public outcomes of reference and evaluation.

In the last chapter an architecture of maps was proposed in order to model the perception of situations and the understanding of positive basic sentences. This understanding of basic sentences enables us to interpret them with respect to real and possible worlds, and it enables us to understand and interpret complex sentences. Interpretation involves understanding and establishing referential relationships to real situations and, generally, to the world. Understanding is a matter of conceptual semantics, interpretation a matter of denotational, realistic semantics.

Interpretation is a conscious operation involving control, which presupposes linguistic and conceptual semantic representations. These are the conscious expressions of activation circuits that involve linguistically specified and situational episodic maps within the neuronal architecture. Hereby the situational episodes that serve to evaluate the sentence as true or false must be perceived episodes. A basic sentence is judged as true if the conceptual-situational part of the process of understanding gets embedded into the process of perception, including the interaction circuits between conceptual maps and sensorial areas, by which the latter come to function as episodic maps. In the case of such a completed embedding, the perceived situation is a satisfaction situation of the sentence.

The interpretation and evaluation of complex sentences, for example negative sentences and sentences with logical connectives (*and, or , if ...then)*, presupposes evaluation of basic sentences as true or false, and decides the evaluation of the complex sentences according to their recursive truth functional built up. The procedure of truth evaluation and denotational semantics altogether, i.e. interpretation as a whole, is a relationship between representations and world. A sentence is held true if and only if the perception of the world satisfies the sentence. Evaluation is mediated by the process of understanding and perception. Representations are the conscious expressions of the activation circuits that involve the primary sensorial areas, which hereby function as episodic maps. There, recollections, imaginations, and perceptions of spoken or written sentences, of situations, of pictures, and of schema's are created, and at the same time expressed in consciousness. In order to relate these to the world we need an experiential consciousness of reality, namely a consciousness of the reality of a perception in contrast with the pure imagination of a representation. This is not just a matter of more vividness and strength of a percept, compared to an imagination.

We are not conscious of a perceptional representation merely in the manner in which we are conscious of an imaginary, invented or recollected representation; rather we are conscious of the situation itself; in perceiving we see, hear or smell the situation or object itself as something outside us that makes itself known to us by enacting pressure on us, or by reflecting or emitting light waves, acoustic waves, or odorous gases which hit our surface, and especially our sensory organs. Perception is not just a situational representation. Rather it is an experience of the real situation itself, with which we get into some kind of interaction, registered in our perception and proprioceptic experience of sensory-motoric activity. Within the perception of a situation there is already included the 'projection' of the perceptional representation into the outside world. We do not perceive a representation of a situation in our head, rather the situation itself is perceived as being around us or at some distance from us. This is only possible because the visual, auditive, or other perceptive data are connected with actual and potential motoric and pre-motoric data, which we are aware of in proprioception.

The perception of a tree includes the direct path of moving towards it. It thus includes the direction and the distance between ourselves and the tree. The perception of a tree in this way includes self-perception. The self experiences itself as part of the world, i.e. in relationships of direction and distance to other things it perceives. The experience of direction and distance is based on a co-ordination of motoric activity with visual 2 1/2 D information, i.e. the slightly different input for the two eyes, and also with the different acoustic information received by the two ears. Such world-experience is absent in representations which are pure imagination, although it can be rehearsed in the recollection of personal experiences, and it is mistaken in illusions. Illusions become evident by incoherence of perceptions with new perceptions in other and better perceptional circumstances, for example in taking other perspectives for seeing a rod in the water that appears broken to different degrees, and in using other senses, such as touching the rod and finding that it is not broken. We also can imagine or remember an object in some distance and direction from us, and as such we imagine it as if it were real. The specific feeling of reality is partly due to the vividness of the perceptual data perceived, and although we do not always consciously feel the variable impetus of input data on our motorically active sense organs, this feeling of differences in the reception of data must be responsible for the vividness and the spatial and temporal quality of perception of a real object or situation. There is no need of an extra inference from a perceptional representation to the existence of the real situation in order to claim reality, via the causal relationship between the object or situation and our reception of the data. The feeling of reality of the self and its surrounding, its being there, is a basic experience accompanying all our perceptions.

Our experiential notion of reality, which can be analysed as evolving in three phases, is essentially connected with the notion of self:

(1.) The self has to experience itself (self 1) in an internal representation of the body in its motions and its keeping of balance. This experience is called proprioception. Sacks (1985) describes cases where patients have lost this capacity of self-perception.

(2.) The self has to identify this self-experienced self 1 with its body as it perceives it from the outside by looking at it, touching it, and hearing it (self 2 = <self 1, body 1>). In Sacks' cases of lost proprioception, the individual can steer itself only via the outside perception of its body and the knowledge that this is its body, though it does not feel it as its body according to presupposition (1).

(3.) The self, via self 2 (or body 1 in cases of lost proprioception) experiences itself as located in space, i.e. located with respect to other bodies. It does this by movement in space of body parts or of the whole body. The (2 and 1/2)-dimensional visual perception, (2 1/2D), due to the slightly different input via the two eyes, becomes a real 3-dimensional perception (3D) by being combined with the feeling of motion for the stretches of the time appropriate for the movement to cover a certain distance. This movement can be experienced as self motion or as a movement caused by being carried or pushed around. Of whatever kind the movement is, it creates a feeling of distance which is part of the perception of distance. Likewise, hearing is combined with movement for experiencing directions and distances.

The reality of a situation experienced within the perception of the situation consists in the embedding of the situational representation into spatial experience, according to step (3.). Hereby visual perception of an object A includes the perception of the direction and distance to A by means of a pre-motoric calculation of a virtual movement directed from the self 2 (or body 1) towards A, and of the time span appropriate for reaching A. Distance is automatically measured in time for a standard speed of movement. This distance measurement by movement is different for hand movement (in order to grasp or touch), for being carried, for walking, running, riding or for driving a certain vehicle. Arm- and hand movement gives rise to a feeling for short distances from ourselves, the speed of walking gives rise to a feeling for longer distances between self and an object.

Percepts normally are not accompanied by a truth claim or an evaluation as true. Being true or false is a property of control of the values of representations which are understood as separate from the piece of world they are supposed to represent; it is the relationship of satisfaction or dissatisfaction of a representation by the world. But perceived situations and percepts generally are not separate from the world; they are for the perceiver part of the world; they are in the world. Therefore for them the truth relationship collapses. The percept is an experience of something being so, accompanied by the self-experience of being there, whereby the percept always is in ones surroundings, standing in the coherence provided by its contiguity relationships of space and time with the world around.

Perception reports, on the other hand are representations separate from what they represent. They therefore can carry a truth value. They are automatically experienced as representing something real by way of the causality of activations in the neural network architecture, leading from perception to the linguistic expression of the perception. If nothing has gone wrong in that process, the reverse activation from the perception or imagination of the linguistic expression to the situational representation results in an identity of this control representation and the recollected or repeated percept. This then is the individuals control of his own perception report, which results in the feeling that the report is reliable. Report and perception stand in interaction within a circuit of activation. Other perceivers never can deny the perception report of an individual, though they can find that it does not agree with their own perception reports in the same situation. Such deviation then can only be due to a failure in the individuals perception apparatus, or to a failure in his language, in the sense that his language is not in intersubjective agreement with the public language. By claiming that a perception report of an individual with an apparently normal perceptive capacity is false, the public forces the individual to change his local position, and if that does not help, to re-learn the language, i.e. to change his terminology in order to agree with the public use of words.

Perceived episodes (situations, utterances, inscriptions), which are represented on sensory and motoric episodic maps, are experienced as being in the world. As recollections of perceptions they are remembered as having been real. When we evaluate basic sentences as true or false we compare them in the process of understanding with currently remembered, previously perceived episodes or with momentarily perceived episodes. An utterance of a basic sentence is found true if and only if its representation achieved in the process of understanding on the situational episodic map, its conceptual semantic representation, can be embedded into a perception or recollection of a perception achieved under appropriate perceptual circumstances. We can then say that the basic sentence expresses a fact, which means that there is a real situation which can be perceived under appropriate perceptual circumstances as satisfying the sentence. Satisfaction of a sentence means that there is a situation, a satisfaction situation, into the perception of which the conceptual semantic representation of the sentence can be embedded.

Our experiential notion of reality established by positioning ourselves within the context of our surroundings lies within the realm of experience of what we perceive as being outside, and feel as bodily activity from within, such as motion in space. We are in this way never outside our own consciousness. Reality outside consciousness, metaphysical reality, cannot be experienced in this way, though it has to be postulated as a "Vernunftidee", as an idea of reason, as Kant and also by Husserl have pointed out. Experienced reality and postulated reality get unified by placing experienced reality within the overall reality which must persist also independently of our experiences, as the ground of all experience to be made.

The world as we understand and know it, our epistemic world, can be thought of as a growing set of information, which we assume to be true about the world. This information is always partial and can be extended or corrected by information reached under better epistemic conditions. The world seen from the epistemic point of view can be understood as the limit of complete information, which we never reach, but which has to be assumed in order to strive for more, and for more reliable knowledge, and in order to accept correction of information. Our partial information is only good knowledge as long as it is embeddible, in a coherent way, within a growing set of information which can be relied upon. In this way, the "Vernunftidee" of a reality, governing our process of acquiring knowledge, is a necessary assumption for the possibility of knowledge.

The basis of evaluation of all sentences is the evaluation of situational representations in perceptually appropriate circumstances as real or not-real, true or false, by ourselves or by others such that the evaluation can be accepted by ourselves. On this basis the evaluation of complex sentences is performed recursively according to the well-known truth functions (negation, conjunction, disjunction, conditional implication), and according to contiguity relationships between situations (causality, finality, spatial and temporal relationships, etc.), by which factual constellations of situations are built up. In the evaluation of such contiguity conjunctions it is presupposed that the truth functional conjunction of the composing basic sentences is true, and it is claimed that the expressed factual relationship between the basic situations holds true. The latter means that the constellation of situations expressed in the complex sentence is real.

A short overview of the interpretation and evaluation of complex sentences, for which by now standard treatments are given in Formal Semantics, gives an idea of the system of complex sentences built on the set of basic sentences:

1. Recursive evaluation of complex sentences according to truth functions: negation, conjunction, disjunction, conditional

2. Evaluation of predicate-logic sentences with quantifiers

3. Dynamic predicate logic evaluation of sequences of sentences with noun phrases and pronouns: In principle, sentences in a text are interpreted as functions that change the information given in the text so far. Pronouns get assigned referents which have been introduced in the previous text by proper names or indefinite noun phrases, or have been taken up by definite noun

phrases. Only if a piece of text is closed, the whole can be evaluated as true or false. Dynamic interpretation of Tense works along the same lines.

4. Interpretation of other complex sentences which represent constellations of situations:

a. Contiguity between situations (causality, spatial and temporal relationships): Clause constructions with *because, before, after,* etc.

b. Relationships between situations and goals or wants (finality relationships): Clause constructions with *in order to.*

c. Relationships between situations and expectations (adversity, concessivity relationships): Clause constructions with *but, nevertheless,* etc.

5. Interpretation of sentences expressing expectations about the realisation of situations, whereby the expectations are based on norms, laws, normal courses of events (necessity, possibility, probability): Sentences with modal verbs *must, may, might,* and modal adverbs necessarily*, possibly, probably,* etc.

Such complex sentences represent organisational structures of information, erected on basic situational representations. This organisation over basic representations is syntactically represented on linguistic episodic maps in the form of inscriptions of complex sentences and texts. The organisational structure can be represented on the situational episodic maps as sequences of public procedures of evaluation. In fact, this whole organisational structure of information is a meta-level of control related to basic situational representations. The latter get embedded into the informational structure by evaluating the basic situational representations as being true or false, and thus taking them to be basic informations. These basic informations connect the whole informational structure to our perception of the world in terms of basic situations: true basic sentences represent basic informations. If these can get identified with real basic situations by means of perception, the information is true information.

Evaluation and interpretation, explicated as denotational semantics, is for us a conscious process, a procedure induced and endorsed by public evaluation. It is, as a conscious process, a manipulation and evaluation of representations, brought about and handled on episodic maps. The manipulation consists in using the evaluation of basic sentences and the evaluation of contiguity relationships as a basis within the interpretation of quantificational and truth functional structures. In the recursively organised evaluation of complex sentences the basic sentences and sentential functions are associated with basic situation types expressed in consciousness from out episodic maps. There the basic situation types can be embedded into perceptions of basic situations, and hereby they are evaluated as true, or as false in case such an embedding is not possible.

On the basis of the above considerations about evaluation we shall now discuss in what ways thinking is possible and whether and how far it can be independent of language use.

4.2 Ways of thinking. Thought, Imagery, and Language

There are brain processes which happen on and between conceptual maps and are involved in thinking, but are by themselves not thinking something about something. These happen through inclusions and intersections between activations patterns and via associations between activation patterns along connections on and between conceptual maps. By themselves, these processes are not conscious. They always happen automatically and play a role in automatic regulations of perception and motor activity. Since inclusion, intersection, and association on conceptual maps are, by themselves, not conscious they are no proper thoughts. This is so because they cannot be evaluated for truth or correctness. They are unconscious thought activities, which are not really about something in the outside world or in a model, because no representation of something outside is involved as long as there is no involvement of episodic maps in these processes. These activities of neuronal networks, which may be relevant to occurrent thoughts and actions, only result in proper, i.e. controllable thinking when they involve episodic maps and thus have results there, being representations, especially inner speech inscriptions, which then can be controlled and judged.

All proper or controlled thinking needs a medium of representation in order for intermediary results and final results of thinking to be represented and thus be controllable. They have to be represented because they have to be evaluated, not only as true or false, but also as desirable or undesirable, good or bad, correct or incorrect. For all evaluation, and therefore also for all propositional attitudes, representation is necessary. This is so because a representation is that which stands in the relationship of evaluation towards that which it represents and with respect to which it is evaluated or compared. The latter has also to be represented, as it occurs in perception, in order for a conscious experience of satisfaction, especially truth, to be possible. A representation has to be expressed in a conscious mood; otherwise it could not be evaluated in a conscious way.

All thinking is conscious only in so far as episodic maps are involved, linguistic or situational episodic maps. Activity between conceptual maps, which is by itself not conscious, only then has conscious results and thereby becomes controlled or controllable thinking, when episodic maps get involved in this activity. Hereby the activity is couched in a phenomenal medium and thus can be the content of an intentional act. In this way it can be thinking about something. We never are conscious of a concept, though we are conscious of the examples we classify under, or produce for, the concept. These examples are all episodes we are conscious of and we know that we possess a certain concept only because we are able to classify them. - Peacocke (1992) analyses the notion of possessing a concept by giving conditions under which the predication that he or she possesses a concept holds about an individual. - Likewise we know that we possess a certain individual concept only because we are able to identify a certain individual. We know that we possess a certain theoretical concept only if we can produce and understand the semantically characteristic set of general sentences which give an implicit definition of the concept. These sentences are represented on the linguistic episodic maps, and thus we are conscious about them. These representations on the auditory, visual, or motoric episodic maps are representations of certain sentences or of certain situation types only because of the role they play in the system of activation circuits that involves the conceptual maps. But concepts themselves are merely abilities or dispositions implicit in the distribution of connection strengths between neurones of the brain. Some connection strengths may be genetically determined as strong to begin with, others are due to a learning process in the course of which certain activation patterns get stabilised on conceptual maps. These activation patterns themselves are not concepts, though they are a result of the established ability to categorise examples in a certain way. The activation patterns can be seen and understood as "conceptual representations" only from the external point of view of the neurological scientist, and only in virtue of their relationship to the sensors and certain kinds of sensorial input established by the connections between the conceptual maps and the sensors. But they are not representations of a symbolic, conventional kind; rather they are a natural sign of a categorisation, whereby the sign is a symptom of that which it indicates because of the causal connection to what it signifies in the world, and to what has been providing the input into the learning processes of the concept forming individuals. Therefore, in the previous chapter, the term "concept indicator" has been introduced for the activation patterns on conceptual maps.

Concepts, which are implicit in a neural network architecture as dispositions of the system to classify and order data in certain ways, have no direct expression in consciousness; for the cognitive agent they have merely an indirect expression via their examples on situational episodic maps and their inscriptions on linguistic episodic maps. Accordingly they are publicly produced and perceived. Therefore, in Dynamic Conceptual Semantics, concepts are equivalence classes of stabilising (converging) sets of experienced examples, or the coherent sets of semantically characteristic linguistic inscriptions, i.e. coherent sets of sentences general with respect to the concept expressing term, which are held true. And each set from such an equivalence class is a representation of the concept (cf. Bartsch 1998). From these kinds of representations of concepts available to the language user, the linguist or semanticist can abstract frames or feature collections, which are partly schematic and partly linguistic representations, usable in science and computational contexts. However, as such scientific constructs they are not available and have no function for the language user.

Thinking is manipulating representations under evaluation, and hereby arriving at new representations. A thought is a representation considered under the point of possible evaluation. As such it is always conscious. Unconscious processes play a role, especially conceptual maps are involved for determining, i.e. categorising and identifying, activation patterns on episodic maps, which then have a representational expression in consciousness. Brain processes, as far as they merely play between conceptual maps, are without direct conscious expression and therefore cannot be part of the mind. The mind is the whole of representations and operations on these. As has been stressed above, representations are conscious and therefore they can be evaluated as correct or incorrect. There is no hidden mind, nor unconscious thought, though there are of course lots of cognitively relevant brain processes.

We distinguish ways of thinking according to what kinds of episodic maps are involved, and what the possibilities are on these maps.

1. Thinking is possible by means of representations rooted in situational episodic maps, with or without additional representation in accompanying language use. Manipulating representations, which are expressed in consciousness via the situational episodic maps, under the point of view of representational correctness, amounts to pictorial thinking, i.e. thinking by means of pictures. Manipulating representations on linguistic episodic maps under the point of view of semantic correctness or truth, amounts to thinking by means of language.

On the situational episodic level we can think in terms of propositional attitudes where we have a pictorial content instead of a propositional content. Thus we imagine a scene, a situation, within an attitude of believing or wanting this situation to be the case. Such a scene or situation is conceptualised by the interaction of sensorial areas and conceptual maps which is equivalent with embedding it into series of previous situational experiences. On this extensional level of conscious phenomena or representations, this embedding happens salva stabilisation of similarity measures under certain perspectives and salva coherence of contiguity relationships. Hereby aspects of the situation get categorised as properties and relationships, and they get identified as partial realisations of individuals. This is how it looks on the level of consciousness. In the architecture of neuronal maps, on the other hand, the processes are activation circuits between different neuronal maps, involving episodic maps and conceptual maps. By these processes, a situation gets analysed into the relevant parts. If the situational content has an isomorphic conceptual structure with a corresponding basic sentence, the situational content is the propositional content of this sentence. Therefore, if we have a suitable language we can express this pictorial belief or desire by imagining or uttering the corresponding propositional attitude sentence. But this is not necessary. Entertaining a propositional attitude with a basic propositional content does not need language. The attitude content can be specified by an imagined situational representation by means of the situational episodic maps, and it thus can be an instance of pictorial thinking.

There also can be narrative pictorial thinking by means of sequences of pictures. A film or a comic strip is an example of such a narrative in an external medium of representation. As far as language is not necessary for understanding, such sequences are understood as perceptions on the situational episodic maps. Also without an external representational medium we can imagine a strip of pictures or a film and thus realise narrative thinking in a purely pictorial medium on the situational episodic maps. Although pictures are richer in content than sentences, in pictorial thinking, and also in narrative pictorial thinking these must be conceptually analysed representations in order for us to be able to evaluate them as true or false under the perspectives which select the aspects from the picture that are at issue in the evaluation. Thus mental pictures are aspectualised and in this way contain a proposition-like structure.

This aspectualisation has been stressed by Nelkin (1996) for percepts, in distinction from pictures as such. Of course, there are switches possible in perspectives which modify which aspects are at issue. The concreteness of imagination does not consist in a fixed picture like a photograph; rather a process of imagination is at work, whereby aspects of the picture are emphasised in imagination and others, previously entertained, are pushed into the background. The full concreteness of a fixed picture is not available in a single moment of pictorial thinking. The imagined, but also the perceived picture rather is built up from series of partial pictures, situation types that are combined into a coherent whole of a picture by contiguity relationships. A real picture can be interpreted differently, depending on the perspectives taken. Different perspectives select different prominent features, between which we can switch. This is evident in psychological experiments about the duck-rabbit switch, the switch in the direction of a staircase, or the switch from the mirrored black profiles of a face to the white space between them, which is perceived as a vase. The basic lines and black and white surfaces of which the picture consists, admit to be collected into two different aspectualised pictures. It has been observed that this cannot be easily done in imagination, though it happens in the perception of pictures (Anderson 1998: 210, and O'Brian and Opie 1999: 141). This points to the fact that an image in imagination is already aspectualised by conceptualisation according to normal ontology, as it also happens in normal perception, in parallel with the primary visual perception of the white vase (= mirrored black profiles of two identically shaped faces) in terms of lines and black and white surfaces which build up the mirrored profiles or vase picture.

In perception, the normal ontology of thing, event, and action concepts is stronger in awareness than the primary visual categorisation, if we do not give special attention to the latter by making it the focus of inquiry into the picture. The primary visual categorisation gets suppressed in consciousness as soon as the normal ontological categorisation gets superimposed on it. If such categorisation cannot be found immediately, the primary categorisation stays there and gets re-partitioned in try-outs for new categorisations. Likewise a hidden part of a complex figure is much more easily found in looking at a real picture, where the details can be available selectively in multiple scanning under different perspectives and foci of attention providing different aspectualisations, than it is found in imagery, which is already aspectualised by its inducing conceptualisation (for examples cf. Anderson 1998: 210). In opposition to Anderson (1998), I would say that the switch or the finding of a hidden figure happens in the process of perception and not in a percept, because a percept is already aspectualised. That it happens in perception is due to the different percepts possible with regard to a series of input receptions, whereby the perspective or selective attention might change, and thus a different percept can come up. The percept is the duck or it is the rabbit, but not both together. The real picture is the same, the reception of data may be approximately the same in multiple scanning, but the selection from the data is guided by perspectives and contextually activated specific concepts. In terms of activation circuits this means that the activation patterns of the concepts that constitute the perspective or focus frame, for example looking for a face or looking for an artefact, strengthen a certain large circuit between activation on the sensorial fields and activation on the conceptual maps, which then becomes dominant in this context and determines what is seen on the episodic map, either a mirrored profile or a vase. These give rise to different percepts of the same real picture. The percept of a picture and the image or imagination of a picture is aspectualised by interaction between the episodic and the conceptual maps, as is every representation.

With respect to the close relationship between perception and imagination, Anderson (1998: 208) stresses Kosslyn’s (1994) point that imagery is an integral part of higher-level perception, a mode of perception that incorporates the knowledge of objects. Also Rollins (1989: 77) agrees with Kosslyn that "at some point in the data processing stream, images and percepts have a common format." Kosslyn argues that mental imagery must share common structures and processes with perception in the same modality. He presents as argument that forming a visual image disrupts visual perception more than it disrupts auditory perception. Rollins (1989: 91) claims that pictorialism assumes that images function like pictures in representing their objects, and that patterns of "cogent imagery" are formed, which constitute non-propositional attitudes" that by virtue of the analogy to picture perception can be called 'pictorial attitudes'." Here, more precision would be to the point: the images do not function like pictures, rather like percepts of pictures, which means that they are already aspectualised by categorisation and contain in this way a proposition-like structure. As intentional contents they are always typed. An external picture, by itself is not typed, i.e. it is not yet categorised. Only its percept is. Perceiving a picture consists of a series of percepts in which different aspects of the picture are selected. In order to function within an attitude the picture must be categorised in perception or in imagination, i.e. it must be a percept or an image. Of course, an artist who wants to communicate some message by means of his art, most likely will create a picture which by emphasis on certain aspects, by foregrounding and backgrounding, causes a certain aspectualisation in the perceiver. Still, the picture by itself will be richer than the aspectualised percept of the picture, and possibly admits a series of different percepts by laying the focus differently, for example by attending to certain details, or to form, balance, or the dynamics of colours and forms.

Even argumentative thinking in a rudimentary form can take place on the purely situational episodic level. We can imagine examples of certain actions and then intent to justify a newly performed action by putting it into line with these cases.

If the previous cases were approved of, the new case hereby is claimed to be approved of, too. By putting a new case in line with other cases it is seen as being of the same kind as the others cases, under a certain perspective of evaluation. That this action fits into a series of other actions is taken as an argument for the action. Instead of an explicit norm, which requires a linguistic norm formulation, a series of positively evaluated model cases is assumed to function in a way that the new case has to fit into it. The norm, to which the new case has to conform, is here given by the series of previous cases. This is non-linguistic categorising and as such an example of non-linguistic thinking. - In the novel *BOY* by Bernlef (1999) a deaf and speechless boy, who also cannot read or write, thinks in series of pictures. Hereby he uses pictures of objects and scenes which he has cut out from sales catalogues and journals. Also Sacks 1985 reports about patients who apparently think in terms of pictures. As pointed out above, I would rather say that thinking here takes place in the form of perception and imagination of pictures; percepts and images are aspectualised.- Argumentation about a single case by referring back to other cases can be done non-linguistically or linguistically, but argumentation by referring to a law or norm cannot be done without linguistic expression. The method of arguing by simply citing cases is prominent in argumentation by uneducated, mostly lower class people. - An illustration of this is given in Bartsch 1984: Normkonflikten in interactie en kommunicatie. In: *tijdschrift voor Taalbeheersing 6,* 1984, 2, 93-100. - This method is also important in juridical casuistic argumentation: a new case is argued about by putting it into a series of previous cases. Analogies are established in this mood of thinking: case H is similar to previous cases A,B,C under evaluation-relevant perspective **P**; likewise the evaluation of H should be similar to the evaluation of these previous cases. The alternative method is arguing by judging a new case via the application of a norm, general rule, or law. This method is only possible by using language in thinking, because a general rule and the implicational relationship can be represented only in form of a linguistic representation.

Boden (1994) characterises creativity as the detection and use of analogies, modulations in conceptual spaces and creative transformations. The depth of transformations attributes depth to the created product. I would object to the assumption of modulations and transformations in conceptual spaces or conceptual maps, because that would be an unconscious and destabilising process. Our concepts are not directly changed, rather the products of our conceptual system in imagination and perception can be changed by modulations or transformations and especially substitution of parts. Change in the conceptual maps is only indirectly possible via the products, which are conscious representations expressed by the episodic maps and which therefore can be controlled and evaluated, also in their aesthetic effects. These products, for example fictional persons, fictional animals, newly imagined kinds of things or situations, and new fictional constellations between situations, and fictional narratives can give rise to the formation of new individual and general concepts, like other experiences can. Contrary to Boden, I would place the creative process, like the process of conscious thinking, on the episodic maps, of course, in their normal interaction with the conceptual maps. The manipulations involve unconscious activation of conceptual maps, but they have intermediary and final representational products we are conscious about, and we deal with these in modulation.

Pictorial thinking, or better: thinking in percepts or images, may be accompanied by linguistic representations, but it need not be. Thus also speechless creatures with consciousness equipped for evaluation can think, namely in such a pictorial way. Although speechless, perspectives and situational or attitudinal/intentional contexts and other attention directing devices select those aspects of situations which have been conceptualised in repeated, relevant, or very strong previous experiences in the learning history of an individual. The aspectualised pictures are situations or episodes, as they are understood by the perceiving or imagining individual.

2. Thinking with intermediary or final results on situational episodic maps can be accompanied by language use, as inner speech or public utterances, both represented on linguistic episodic maps. The pictorial attitudes then get expressed by propositional attitude sentences, whereby the propositional content is expressed by a particular sentence, i.e. a non-general sentence. Likewise the pictorial narrative is expressed by a real linguistic narrative consisting of particular sentences, which can be richer in information compared to the pictorial narrative, by using also expression in general sentences, and all the logical and linguistic devices for organising information into discourse structures that take into account assumptions about hearer's knowledge, general presuppositions, and other relationships between background information and new information.

The dual coding approach of mental representation in propositional and in pictorial form has been introduced into Cognitive Psychology by Pavio (1986). Roediger and Goff (1998: 256) describe the consequences of Pavio's view for memory and thinking, namely "that people can represent information in both verbal and imagined form, and that if information is represented in both verbal and imaginal codes, its retention will be better than if only one code is used. Therefore, when people study pictures in preparation for a verbal recall test, they encode them both imaginally and verbally, whereas they encode concrete words in a verbal code (although some imagery may be evoked by vivid words). Because pictures are strongly encoded in both modes, they are better recalled than are concrete words. However, concrete words do weakly activate a non-verbal code.” Abstract words do not do this, and therefore words representing something concrete are better recalled than words representing something abstract. Also when forming images together with the verbal input material, the words are better recalled than without.

3. Like manipulations on the representations by paper and pencil, a great deal of human thinking is done by manipulations on the representations rooted in the linguistic episodic maps. In both ways of thinking, linguistic or pictorial representations of propositions and schemata are the essential vehicle of thought. Consciousness, as it is provided by neuronal activity on episodic maps in their interaction with conceptual maps, lets us see the representations such that we can

manipulate and evaluate them also under the points of view of correctness, especially semantical and logical correctness. Under these points of view we can make inferences purely on the basis of the syntax and of semantically relevant sub-categorisations of lexical items. In such logical thinking, linguistic representations are manipulated according to rules of inference. The logical ways of thinking by means of inference rules take place on linguistic representations which involve general, existential, and particular sentences. But such inferences are only thinking because we have a judgement about semantic correctness of these manipulations. The correctness can be shown by interpretation in denotational semantics, especially by truth evaluation and the recursive rules for the build-up of complex sentences. Thinking and knowing about the validity of this activity, like all judgement about semantic correctness, has to be a process that is conscious, or can be made conscious in its steps, which are the representations reached as intermediate or final results.

All propositional thinking, namely entertaining propositional attitudes, narrative thinking, logical thinking, and broader argumentative thinking, is thinking about something, some model or piece of world, with respect to which satisfaction, or semantic correctness, can be established. Thinking in the medium of linguistic representations must be valid thinking and generally semantically correct thinking. Therefore it has to take place in a way which incorporates episodic maps as the neural and mental medium for consciously expressed representations; and by such representations we can evaluate the intermediate and final results of our manipulation of representations for semantic correctness, especially for truth. Because of the perspective of evaluation, this manipulation of representations is what we call thinking.

Bechtel and Abrahamsen (1991: 169-172) propose that logical inferencing is learned by pattern recognition: evaluation patterns are learned, not rule formulations. In teaching sequences, pupils learned to recognise those patterns which result in truth. Model proofs of natural deduction are the learning material, explanations in terms of truth evaluation are given after the examples. "What happens when somebody knows logic construct proofs is that he or she simply *recognises* or *sees* what to do in particular situations ... on the basis of extensive experience"(1991: 174). Notice, however, that the patterns that have to be learned can only be detected if one is able to recognise propositional and predicate logic schemata in quite diverse concrete sentences. One needs to understand that, for example, sequences of quite different length and form represent propositions, and that all the different propositions have in common that they can be evaluated as true or false, and hereby can be represented in a schema by a simple propositional letter p, q, etc. The patterns cannot be detected just over sets of sequences of symbols; rather they are patterns over sets of sequences of truth-value bearing sentences. This, of course, is not patterning on the level of perceptive input. What we need here is an understanding of sequences of symbols as sentences, and an understanding of sentences as judgements. And over these a further abstraction is made such that what matters are their truth-values. On this level of patterning it does not matter what the sentence is, it can be simply represented by a propositional letter, or a predicate-logic form, which functions in the truth assignment. Such pattern recognition thus requires first some conscious steps of abstraction and appropriate representation of concrete linguistic material in order to get at logically relevant patterns, which then on this higher abstract level require perceptual, especially visual, consciousness in order to see the schemata of propositional and predicate logic representations on which the learnable patterns are manifested.

If Bechtel's and Abrahamsen's theory of logic learning is valid, it provides an argument for the claim that logical thinking is only possible thanks to the reflective consciousness of the thinking individual, or the designer and user of a logical machine. That here pattern recognition occurs over a material of a high degree of abstraction, which presupposes not just perceptive but also reflective consciousness, has to be added to Bechtel and Abrahamsen (1991), who suggest that their methodology of logic learning "raises the prospect that the very ability to engage in formal symbol manipulation may not be a basic cognitive capacity, but may be something we *learn how* to do by means of a very different sort of mental encoding (1991:174/175)." Bechtel and Abrahamsen see their position supported by Dreyfus and Dreyfus (1986) about whom they write (1991: 175): "In their analysis, expert performance does not rely on propositionally encoded information, but rather on the ability to recognise situations as similar to previous encountered situations and to rely on what worked in those situations." This, of course, is the general procedure of concept formation, which includes generalisation over situations in the trend of "if F in s than q": to the conceptualisation of a situation as being F it is connected that one can do or expect q. Since q is connected to F in s by association it will be activated together with F in s. If this were all, we would be restricted to associative thinking and associative behaviour. Rather thinking frees us from this automatic reaction by enabling us to take into account other premises. In the light of those other premises, the above conditional relationship enters as one of the premises such that we can evaluate its usefulness in special circumstances by drawing the conclusions from it together with other premises. In this way we are able to see undesired consequences which help us to block the conditional association. Concept formation, including associative conditional connections, is not what is typical for logical thinking. It only gives rise to the material for it, in providing premises.

Default reasoning is what the most typical everyday manner of thinking. If we know that birds can fly, and Twiggy is a bird, then we draw the default conclusion that Twiggy can fly. But suppose that Twiggy ist a penguin, and penguins can not fly. This more special piece of knowledge frustrates the default conclusion, and we have to conclude that Twiggy can not fly. Default logics are developed to model the correctness of this way of reasoning. In terms of circuit activation, on the other hand, we can say that imagining or perceiving Twiggy as being a penguin consists in a circuit of activation in which small circuits of associating its small wings with a relatively big and lump body inhibit or frustrate the establishment of a resonance circuit between the Twiggy activation and the activation of the conceptual indicator for flying, even though the latter may be somewhat activated by the indicator of the bird concept, which has been activated via the Twiggy representation. 'Flying' is a concept that includes a certain relationship beween body size or weight ,and the form, the size, and the strength of the wings, as it shows in the image of flying birds of different kinds. For preventing the default conclusion from birds, who generally can fly, to the more special case of a penguin, it is not even necessary that we have explicit knowledge of the statement that penguins cannot fly. Of course, explicit knowlege and understanding of such a statement will also suffice to prevent the above default conclusion.

A similar case is Crawky, the crow with a broken wing. Crawky is a bird, and as such would be expected to be able to fly. Also as a more specific kind of bird, a crow, it is expected to be able to fly. Here the image of Crawky with a broken wing, brought about by the respective circuit activation, inhibits the establishment of a synchronic or resonance activation circuit constituting the judgement or the situational image of Crawky flying, in which two broadly stretched out wings flapping strongly should be represented. Because Crawky is a bird and especially a crow, the indicator for flying is somewhat activated, but the competing circuit of Crawky with the broken wing prevents the built up of a full-blown resonance circuit for an image or judgement of Crawky flying. This competing circuit is stronger because in the perception or the imagination a specific scene, and more specific knowledge generally, gets more attention and is scrutinised more intensively, whereby the duration of the circuit activity is longer or more stable than the one due to general knowledge. We see that the model of circuit activation shows nicely the causal background of how we arrive, or don't arrive, at a default conclusion. Here the same principle is at work which Van Leuwen et al. (1997) have pointed out for destabilisation and switch in visual perception (cf. chapter 3). The justification of such a non-default conclusion, of course, does not depend on this level of explanation in terms of neurological circuits and their expression in consciousness. Justification rather depends on the relationship of our representations, namely perceptions, imaginations, beliefs, judgements, with the ways in which the world verifies them.

Logical thinking cannot and should not be reduced to concept formation on the level of situations and the connected unreflective associative behaviour. This, of course, does not imply that logical thinking must be a separate basic cognitive capacity. We still may think of it as building on the basic capacity of concept formation, but in a reflective and not purely perceptive kind of consciousness, which makes possible correctness judgements and higher abstraction over truth-value bearers, namely propositions and their predicate-logic structures. In reflective consciousness we are aware of the structure dependent relationships of sequences of symbols to evaluation in reality or models of reality, and by this we can understand such sequences as sentences expressing propositions, and especially as judgements about a piece of world. Only then we can devise representations under this point of relevance and detect patterns on this level, namely logical schemata or formulas. May be that there is no special cognitive component or module required for logical thinking. However, rather certain operations or manipulations on representations under reflective consciousness, i.e. under evaluation, are performed and seen or recognised as permissible salva correctness of judgement.

As has already been pointed out in passing, there are no hidden thoughts and there is no hidden or unconscious mind as a collection of hidden representations. Representations, per definition, cannot be hidden. They are conscious expressions of the episodic maps in their interaction with conceptual maps. What is unconscious is just brain. There are no hidden representations and no hidden schema's of procedures. Rather there are dispositions of the brain to come up with certain representations and procedures on the episodic and on the motoric maps, where there is the sensory and motoric material to make them experienced and controlled in consciousness. The unconscious, and the hidden mind have the same neurological character of being implicit in the strengthened neuronal connections, as everything learned on the basis of experience and training. The unconscious has the additional property that certain representations and behaviour get strongly associated with certain emotions. This means that the dispositions include strong associative connections between conceptual, behavioural, and emotional maps or areas. Something similar holds for the notion of a soul. It can best be understood as a disposition towards certain emotional and volitional reactions towards scenes, episodes, and behaviour. The soul, in such a view, is constituted by having conceptual maps, motoric maps, and their combination in behavioural maps, strongly connected to emotional maps. Then the emotional maps also can play a role in causing behaviour and creating images and motoric activities, and they determine the emotional character of the perception of episodes and of the proprioception of ones behaviour and actions.

Our mind is open to us. We can get to know our dispositions, hidden in the strength of the neuronal connections, only via their products. If we want to know our mind and our soul, we have to look at our behaviour and our imaginations and analyse them in their relationships to each other and to their emotional evaluations and preference rankings. Like we cannot analyse our concepts directly, but merely their perceived and imagined examples, we also cannot analyse our other behavioural and emotional dispositions directly. Analysis is not introspection into a hidden mind or into the unconscious, which traditionally has been thought of to work by hidden rules applied to hidden representations, as has been assumed in mentalistic linguistics (see chapter 1). Rather the analysis of our conceptual and other cognitive and linguistic capacities is an analysis of the brain’s conscious products and their interrelationships, which come about under appropriate internal and external stimulations. It is an analysis of the external, as we perceive and imagine it. Hereby we indirectly get to know about our hidden dispositions, i.e. about our mind and our soul.

Chapter 5

Controversial Issues in consciousness research

In this chapter the position on consciousness, neuronal network architecture, perception, imagination, and language, which has been developed in the preceding chapters, will be applied in the discussion about some controversial issues in consciousness research, namely 1. the question whether consciousness is an internal monitoring device of brain states, or of the external results of brain activities, 2. whether all conscious states do involve thought or judgement, 3. whether there are different kinds of consciousness, and 4. whether there is a 1-1 correspondence between brain states and conscious states. Contributions and positions concerning these questions will be discussed and the argumentation will be evaluated. Concluding, a stand point with respect to each question will be derived on the basis of the argumentations given and from the implications of the position developed in this essay.

5.1 Is consciousness an internal monitoring device of brain states, or rather a monitoring of the external?

Armstrong (1997) and Lycan (1997) both take consciousness to be an internal monitoring of neural processes. Consciousness then is a higher order perception of what passes in one's own mind. The problem with this assumption is that own neural states and processes can never be observed. We do not realise whether our neurones fire or receive an electric impulse; the brain is the only organ that cannot be felt, and therefore surgery on the brain does not require narcotisation of the brain itself. The brain apparently is not equipped for observing itself. The first part of this argument could be answered by insisting that our second order perception does not perceive in terms of third person observation which would mean observing the firing of neurones; rather the self-monitoring happens in a medium of first person perception which is of quite a different quality. If one would concede this, then the question arises why it is so that not all brain processes and states can be 'second order perceived' and what is the principle of selection of the ones which are monitored versus those which are not monitored?

Baars (1997) proposes the method of contrastive phenomenology, in which he contrasts conscious with unconscious processes and states in different mental modalities. Note that according to Baars' terminology not brain processes, rather mental representations can be conscious or unconscious. He states (1997: 187): "...by contrasting similar conscious and unconscious mental representations in belief, perception, selective attention, imagery, and the like, we can gather a set of sound empirical constraints on the distinctive properties of consciousness as such." -- What is rather problematic in this formulation is the notion of "unconscious mental representations", which is typical for traditional cognitive psychology. I shall come back to this inadequacy later. --

Baars distinguishes unconscious beliefs and conscious beliefs, whereby the latter are taken into a focus of selective attention due to their relevance in certain actions and further contexts of perception and thought. In perception we have percepts as conscious representations of input versus unconscious cases where the reception of stimuli has subliminal effects such as taking unconsciously into account certain road conditions when driving or walking, taking unconsciously into account contexts when disambiguating words in interpretation, or for example the fact that a very short presentation of the word *book* will influence the recognition time for a word like *paper*, though the first word was presented as a stimulus under threshold time and was thus too fast to be seen (consciously). There are cases of unconscious habituated input representations as contextual constraints of conscious percepts, and there are also unconscious expectations of specific stimuli that play a role in perception. Also on the level of imagery and inner speech, conscious cases such as retrieved images and inner speech, currently rehearsed items in working memory and in the sensory memory of different modalities are conscious versus comparable unconscious cases such as unrecalled memories of images and inner speech, currently unrehearsed items in working memory, and long term memory representations needed in order to come up with recurrent items in sensory memory. Further there are attended cases of stimulations and events versus unattended ones, the first being conscious and the latter unconscious. There are also conscious cases of learning versus unconscious learning in developing automaticity with practice. Baars (1997: 198) stresses that these contrasts "place strong constraints on any theory of consciousness". Although he does not really formulate the constraints, he points out that by becoming a matter of attention, conscious beliefs create access to new decisions and revisions of belief, and generally, conscious phenomena can be handled with control, correction, and flexibility. That brain states become conscious phenomena means, according to Baars (1988), that they become items in a global workspace, from where they can be broadcasted throughout the brain. Baars (1987: 223-29) speculates that the global workspace is not cortical, rather it is located in the reticular-thalamic system which connects to all sensorial and motoric areas and other cognitively relevant areas of the cortex. If this can be proven to be true, then the reticular-thalamic system, which humans share with other species, would be an area where binding occurs between different information processed on sensor-specific episodic and conceptual maps, resulting into the presentation of objects and situations. This leaves open whether or not the binding is done by certain common frequencies in neuronal activation patterns and circuits, as Crick and Koch (1990) propose.

It should be noted that Laberge (1990) also stresses the importance of the thalamus in directing attention towards a local focus, or to certain features which are important in the context of motor planning. Selective attention arises by circuits of activation in which reticular nucleus cells and thalamus cells are involved, providing extra activation to those areas which will be highlighted in attention, while causing inhibition to other areas. The enhancing activity is caused by expectations, taking the form of pre-activation of certain neurones which play a role in representing a local sub-area of the surroundings of the individual, or taking the form of other activated neurones which play a role in the activity patterns indicating certain concepts, or motoric gestures, or emotions connected to certain concept indicators or indicators of motoric gestures. These circuits he calls “thalamic enhancement circuits” and they support the selection of what receives expression in consciousness by paying attention to some aspects of situations and not to others.

According to Baars, there are many unconscious processes and representations occurring parallel in the brain, and those representations become conscious that perform a role in the special kind of computational process in the global workspace. This of course means that not brain processes become conscious, rather certain representations, i.e. results reached in these processes. Problematic is here the term 'representation' because it is used also for unconscious brain states. Representations stand in an relationship of interpretation to what they represent, which requires more that a causal connection of a brain state to some possible stimulatory input. To understand a representation requires that one recognises that it has an aspectualising relationship of correctness to that which it represents. It has to show certain aspects of that which it represents. This requirement makes that representations have to be in consciousness, and if there is such a global workspace as Baars postulates, certain brain states that come into the focus of attention, are hereby experienced as representations, casted in the material of the different sensory modalities. The subliminal receptions and other unconscious happenings in our brain do not give rise to representations proper but merely to the corresponding partial brain states that are processed further and made use of in guiding our motoric activity, in the unconscious steering of thought, interpretation, and perception, and thereby can have conscious consequences. However, according to Baars, a mental representation becomes conscious by entering the global workspace in which it can be part of certain computational processes.

In contrast to Baars, I rather would say that partial brain states become representations, and thus mental phenomena, by entering the global workspace of conscious phenomena. The status of unconscious beliefs, or non-rehearsed memory items, would then not be the status of a representation, but rather an implicit capacity of the brain, a disposition to make a representation in the global space of consciousness under appropriate stimulation. The disposition can be activated by external stimuli or (internal) representations in directed attention for recurrent contexts of such stimuli or representations. Long term storage of information, such as the capacity of conceptual ordering and memory generally do not consist of physically discrete representations stored as items somewhere in the brain. The brain, after learning, rather has the potential to form relatively stable activation patterns and circuits between these, after appropriate stimulation or cueing. These have a phenomenal, conceptually categorised expression in consciousness if sensory fields are involved in order to supply the necessary phenomenal content. The episode expressed on such a sensorial map within an interaction circuit with conceptual maps then is an (explicit) representation.

O'Brien and Opie (1999) put forward some ideas about what would make a representation come up in consciousness. A representation is explicit and can be used in controlled reasoning, arguing and planning, which are typical functions that can be performed by dealing with representations in consciousness. There are no implicit representations, rather there are potentials or capacities in neural networks to come up with activation patterns, some of which are stable in a certain sense. A representation is made by a stable brain activation pattern that can be the result of parallel processes steered by input and the potential of the neural networks coded in the strength of connections between neurones, similar to those in PDP-networks. The stability of the pattern consists in certain pikes of neuronal activation formed by strength and frequency of these firing of the neurones. Such stable patterns can "facilitate meaningful communication between PDP networks, and hence contribute to coherent schemes of action." (1999: 140). The "broadcasting" between different parts of the brain assumed by Baars takes place via such stable patterns. Consciousness here is a collection of several relatively independent parts which communicate via representations in a phenomenal medium ordered by space-time and conceptual structuring. In this medium the brain can see its own products, as far as these are stable activation patterns. A conscious, explicit representation is due to a stable activation pattern distributed over broad areas of the brain in interaction. O'Brien and Opie call their approach a "connectionist vehicle theory of consciousness", whereby the vehicles are the stable activation patterns, in contrast to a process theory of consciousness, where taking part in a special kind of computational process would make a mental representation to become a conscious one, a position advocated by Baars.

O'Brien and Opie, like Searle (1992), think that consciousness is absent in simulations of conscious systems on digital serial computers because there the neural states are not real activations but merely described symbolically in the code of the serial machine (1999: 140): "A simulated stable pattern of activity is no more a stable activation pattern than a simulated hurricane is a hurricane." If their hypothesis about stable activation patterns having an expression in consciousness would be correct, we could expect a working replica of a neural network to be conscious, if we ever could built such a thing. I think their position is not quite right in stressing that the stability of an activation pattern makes consciousness. An important factor is that the somewhat stable activation partly takes place in areas of the different perceptual modalities, such that we experience differences in colours, forms, between colours and forms, and between colour and sound, or sound and smell. Thus not all stable activation patterns need to have a conscious expression, only those that hit, so to speak, the sensory areas. The intra-modal and the inter-modal binding in their model seems to happen according the circuits-theory of Edelman (1989), providing for the different levels of abstraction in perception, possibly aided by Crick and Koch's suggestion (cf. a.o. 1997) of a certain common frequency of the pikes of activation.

O'Brien and Opie posit that consciousness is "not a single monolithic state, but a complex amalgam of distinct and relatively independent phenomenal elements." And they conclude (1999: 146): "Phenomenal consciousness is not an emergent product of complex information processing, nor of sufficiently rich and wide–spread information processing relations; rather consciousness is the mechanism whereby information is explicitly encoded in the brain, and hence is a fundamental feature of cognition." I think that this position, formulated against Dennett (1991) in the first place and Baars (1987) in the second place, is essentially right. Consciousness makes possible (discrete, explicit) representations and thus is the basis for having a real language with more than a causal semantics, namely with a denotational semantics, based on intentionality which makes representations to be about something and therewith representations. Being conscious of a dentational semantics makes possible notions of correct–ness of judgement and language use generally, and with that logical inference as a semantically sound way of reasoning. All reasoning is reasoning by using representations, relationships between these, and relationships between representations and the world we reason about. We cannot have propositional attitudes and thoughts, nor reasoning without such representations. But these representations are not just any relatively stable activation patterns in the brain.

There has been criticism of O'Brien and Opie (1999), presented in an appendix to their article. Church (p. 149-150) requires that the authors should be able to state which stable states are conscious and which are not. Here a suitable answer would be the one added above, namely that only those are conscious which hit the sensory areas, by which phenomenal material is provided. Other critics object to the much to simple notion of stability: Dennett and Westbury (p. 153) point out that there are more complex cyclic forms of stability between several networks and not just stable patterns in terms of constant simultaneous firing. This is certainly right, and we can add that circuits which hit the sensory areas get an expression in consciousness. In the circuits conceptual structure, time-space structure, and other phenomenal contents are composed. Lloyd points out that a "stable" state is problematic because every state is a moment in a process. Magan adds that stabilisation will not be fully reached; all states are involved in a dynamic interaction with other states; and Schröder (p. 167) remarks that the notion of stability is empty if no time scale is given, and furthermore the brain continuously gets new input which changes the states. Full stability is never reached. Also Taylor objects to O'Brien and Opie's notion of a stable state (p.168): a stable state would stay for ever. He suggests that one should rather speak of "local recurrences". "The bubbles of activity thereby produced by an input, and continuing their existence after the input is switched off"(p.169) explain several data of somewhat lasting or repeating percepts and imagery. Conscious events hereby are not instantaneous, but have a certain dynamics.

Lloyd thinks that the functional context of the vehicle of consciousness, the relatively stable activation state is important for consciousness. This point is made also by Clapin in another way (p. 151): He asks what exactly is explicitness? The stable pattern can encode more information, namely also superposition (hypernyms) and is thus not uniquely expressing one information. Explicitness comes in by relationships a pattern has to other patterns, corresponding to, for example, conceptual analysis. If the authors claim that consciousness is due to explicitness, they include a process theory of consciousness: "a stable state's explicitness depends on the processes it can undergo and its role in the information economy of the system it is part of." Clapin is right about explicitness, though he is not right in thinking that this kind of explicitness makes the state conscious, or that O'Brien and Opie claim this to be so. It is not so clear that O'Brien and Opie would accept this as their position. Anyhow, we should realise that the explicitness pointed out by Clapin rather is a consequence of the relatively stable state, or more correctly, of the recurrent circuit having an expression in consciousness by hitting sensory areas and thus creating episodes. On the level of consciousness we can become aware of the conceptual relationships by studying the relationships within and between the imagined examples of the concepts. Only by attending to these relationships we can explicate the concepts involved. However, the notion of explicitness of O'Brien and Opie seems to be a different on, namely that a representation is explicit, i.e. can be approached as a unit, or is a discrete phenomenon, compared to the implicit information within the architecture and the connection weights of neural networks. Thus a picture or a linguistic utterance and any mental image of it in consciousness is explicit information as opposed to implicit information hidden in the overall state of a network as an ability or disposition towards making explicit representations under appropriate stimuli or in an appropriate context.

The result so far is that consciousness might be due to some relatively stable activation patterns or, much more likely, recurrent activation circuits in architectures of neuronal networks that include areas for processing of sensorial data and thus combine conceptual categorising, space-time structure and phenomenal material. The latter is essential as a medium required for any mental representation. Consciousness is therefore not just an epiphenomenon of brain processes without any essential function in cognition, but it is the very basis of our ability to use representations of information about the world in a controlled and especially correct manner. All controlled cognition involves manipulation and evaluation of representations which, for this purpose, all have to be conscious. Other authors with similar views are Güzeldere (1997), Van Gulick (1197), and Flanagan (1997).

With respect to the position of Lycan (1997) and Armstrong (1997), who claim that consciousness is internal monitoring of neural processes, Güzeldere (1997: 797) argues that even if we could observe our brain states, this would "involve only the intrinsic properties of the brain states themselves, whereas figuring out what they are about needs to bring into the picture their extrinsic, intentional properties as well." According to him there is no internal second order perception, rather a thought makes a situation conscious. The automatically driving truck driver who suddenly thinks "Oh, there is the rail road crossing, I must have been driving for the past hour", has a thought and not a higher order perception. This position according to which a thought makes the driving conscious is supported by Carruthers (1989), who claims that even experience of pain requires a higher order cognitive faculty for producing the thought that one has pain. -- This position that consciousness always requires thought and thus some linguistic or language-like representation will be discussed and modified below. I shall argue that a representation, a percept or image is primary and causes a linguistically encoded thought which is also a representation. Furthermore we should notice that having an occurrent thought always requires a medium of representation, not necessarily a language. The thought can also be a percept or image that is evaluated as true, if it functions as a judgement. Of course, also the use of a language requires a medium in which the language represents. Also the latter is phenomenal material or content; sentences in what ever language are tokened orally, accustically, visually, whether in an external medium or in a phenomenal medium of our perception system, which is also used in imagining sentences. Occurent thoughts and linguist expressions both make use of the sensory fields that provide phenomenal material.

Van Gulick (1997) discusses why human abilities, like initiating actions, forming an integrated self-concept, learning new non-habitual tasks, forming plans of action require conscious phenomenal experience. They all involve meta-cognition in the form of a representation of the self in relationship with situations. Phenomenal awareness makes such meta-cognition possible because it creates the possibility for semantic transparency of representations. Phenomenal material is necessary for constructing representations which have an understandable relationship with what they represent in the world, and which have understandable relationships with other representations. Semantic transparency in representing an object, for example, means connecting a representation of the object with a network of representations concerning structure, function, and use of the object. The phenomenal representations in our mind have this property of semantic transparency (1997: 438): "It is the awareness of these transitions among representations in the seemingly continuous flow of experience that provides the phenomenal or subjective experience of understanding, our conscious feeling that we know what we are thinking about." Our subjective experience of understanding necessarily involves semantic transparency. Referring back to Kant, Van Gulick stresses that all conscious experience requires the presence of 'intuition' as a continuous, sensuous manifold within the structure of experience, having a spatial and/or temporal, and a conceptual organisation. Objects in their spatial and temporal extension are represented in the medium of different phenomenal material, i.e. qualia. These phenomenal representations, including the semantic transparency of objects, represent the external world and are not higher order perceptions, i.e.

representations of brain processes, though they are certainly brought about by brain processes.

Van Gulick also points out that the phenomenal representations fulfil functional roles, such that Baars' broadcasting activity seems to be a suitable metaphor for the role of consciousness. He stresses that qualia are indispensable for consciousness which consists of experience in the form of phenomenal representations. This is kind of circular. It describes our consciousness as it happens to be and as it functions. But his argument cannot exclude that there might be consciousness of quite a different kind without phenomenal material. He explicitly leaves it open whether the same functions can be fulfilled by non phenomenal representations, i.e. representations with absent qualia. These then would be representations in a hidden mind. I rather argue that such representations do not exist. They cannot be representations in the full sense of the word, because this includes that one is aware of the relationship of representing.

States in computers, produced according to certain input and a certain program, or brain states such as a recurrent activation circuit or a relatively stable activation pattern are called "representations" by computer scientists and many cognitive psychologists. But this is a somewhat loose use of the word "representation", or a secondary derivative use of the word created by concept broadening. These so-called representations are only representations of some X according to the scientist, who assumes that he has endowed the computer with a correct program such that the respective state can be used as a representation of X in further processes. The scientist knows in which way the so-called representation is causally connected with a certain input and with other representations. And if this all is correct, given our knowledge of the object area with respect to which the so-called representation has to represent, then such computer state can be taken by the outside observer or scientist as a correct representation. For him these representations are of course couched in phenomenal material, namely his perception of the electric circuits and the states of these he observes on measuring devices, or simply the print outs on the screen of the running phases of the program. For the computer itself the so-called representation is nothing since it is not experienced in any way in its relationship to what it should represent; and also seen from the outside it is just a state the computer is in. Only for the knowledgeable scientist, and in consequence also for the able user of the system, the computer state, or its print out, can be interpreted as a representation. Without this interpretation the so-called representation is not a representation at all. The interpretation requires that the so-called representation can be perceived, and the area or object which it represents can be perceived, and a denotational relationship between both can be established by an intentional act of reference in which the representation is used to represent the object. Therefore both, the representation as well as the object it represents, have to be couched in phenomenal material for the observer and interpreter. Of course we can also imagine possible, unreal objects as the interpretation of the representation, but also this is done in the phenomenal material provided by our perception apparatus. Even though the computer produces representations for us according to a program design, the computer running a program has for itself no representations, it cannot experience them as such, there is no sense in which we can say the computer has consciousness. Of course, it could be designed to register his own states, even in relationship to certain input. But that does not make the states to be representations. For that the relationship has to be judged as correct or incorrect, and not just be causally established in some way.

If a causal connection is regular and preserves desirable structures of the object domain, it can be called correct with respect to that domain and those structures. In order to judge about this correctness, both have to be perceived, the so-called representation and the domain with its structures. If we would construct a robot with two fields of states, one field with states resulting from sensory input and the other field with states which have to be taken as symbols for the sensory states, an isomorphic mapping or a merely homomorphic mapping (in case the symbols represent classes of sensory input and the structures hold between members of the classes according to class membership) could be established using a suitable program. The symbol system could then be called correct. But again: the computer does not know that it has developed a correct system of representations; we, with our ability for phenomenal consciousness, see that this is so, and with some good insight into the program used we can see why that is so. We can, possibly in the future, design all kinds of artificial systems which are correct in all functions that can be described, even in controlling for correctness under all points of view. Nevertheless they miss the most significant property of consciousness, namely to know what one is doing. For that one has to perceive oneself, the objects on which on acts, and what acts one is doing towards these. Perceiving something is not just registering some of its causal effects, rather it includes knowing that one registers it, and experiencing what it is like.

In the terminology of Flanagan (1997) we can say that robot systems have informational sensitivity. They store or process input in a regular manner, according to the programs enacted. But such systems have no experiential sensitivity, which involves awareness of the state one is in. This awareness is consciousness. The state one is in is a state of perceiving the outside world, states, processes and events in the environment (1997: 361): "Consciousness is multimodal, involving the discrimination of many different types of states, things, and events, and it is loud and noisy. It broadcasts what it knows." This broadcasting happens within the system itself, and also to the outside in form of speech or icons. He emphasises that this knowing the state one is in and this internal and external broadcasting certainly has an evolutionary value. One main function of consciousness is control, not directly of internal brain processes, but just before the onset of the motor activity which would be the output. He refers back to Libet (1985), who has stated (Flanagan 1997: 362) "that unconscious readiness potential for actions precedes conscious intention or urge, which proceeds muscle movement". The role of consciousness according to Libet (1985: 536-7) is the following: "Conscious control can be exerted before the final motor outflow to select or control volitional outcome. The volitional process, initiated unconsciously, can either be consciously permitted to proceed to consummation in the motor act or be consciously vetoed." Flanagan explains that the mind allows the action to proceed (automatically) to consummation unless it has reason to stop it. Let us remind ourselves that we all have experienced this controlling function quite a lot, especially in speech. How often, especially when we are disturbed or tired, are we not about to just say the opposite of what we mean, for example" large" instead of "small", "a cup of tea" instead of "a cup of coffee", "bread and bedfast" instead of "bed and breakfast", and all the other slips of the tongue which we, so to speak have already on our tongue and can just mend before having pronounced them, or which we decide nevertheless to pronounce, for the fun of some of them. We control and correct what we perceive we have already said, and we even more control and correct what we are about to say, i.e. have for a moment more or less ready made in our imagination, which involves the perceptional areas as well as the motoric areas in which speech is formed.

Note that it is not the case that the neural processes and states are perceived and possibly blocked because of incorrectness, as Armstrong and Lycan propose. Neural processes and states are never correct or incorrect. They are just what they are. Rather the imagination or the production of the external phenomenal result of such processes and states is perceived and possibly blocked and corrected. We have no power to block or control brain processes or hidden mental processes (if they would exist); we cannot even perceive them. Only the *results* of some of these brain processes can be controlled and blocked when they hit or enter sensory and motoric areas and hereby become experienced in consciousness. The phenomenal results can then be blocked from being involved in further controlled mental manipulations of representations or symbols on episodic maps, where they are in-between results. The phenomenal results are the perceived internal or external results. They are episodes which are judged by being played back through the neural system against other explicit and implicit knowledge they have associations with. If they fit, they are supported as correct and are controlled input for further brain processes. If they are in conflict with other explicit and implicit knowledge they get blocked from further involvement.

What can be conscious can be controlled, blocked, or corrected. Control hereby takes place on the intentional level where the relevant chunks of our knowledge can be represented and newly produced or perceived items can be measured against knowledge previously accepted as such, namely as correct information. When I say "please give me a cup of tea", but mean to have a cup of coffee, I have already a non-linguistic representation of a cup of coffee in my mind, including taste, smell, colour of the coffee. Against this conscious representation or explicit knowledge of what I want, represented on the situational episodic maps, which include circuits between conceptual and sensory and possibly motoric brain areas, the token "a cup of tea" is measured and detected as wrong by playing through the associative connections in understanding this expression. This links it via conceptual maps to imaginations of cups of tea that in this case conflict with the coffee representations, and hence it is blocked or corrected. That I was prone to say "cup of tea" for a cup of coffee to begin with was, of course, due to a miss caused by strong associations and closeness between the causal effects of cups of coffee and cups of tea in our experiences of these alternates. The phonological and morphological mistakes seem to be mistakes in binding the separately produced units into wrong orders within the phonetically stressed morphemes. Here the results can be controlled close before or at output by having available for comparison, due to the semantic and linguistic activation, the production circuits of the true word forms and morphemes.

Levelt (1999: 113-4)) treats self-monitoring in his model of speech production as consisting in a loop of stimulation from overt speech to the auditory fields in the brain, and from internal speech to the same auditory fields. Internal speech takes place in specialised pre-motor and motor areas, which are at other brain locations than the auditory areas. Thus self-monitoring most likely takes place through the speech perception system, via the loop. In his model, Levelt locates the representations that are self-perceived on a level of phonological/phonetic speech preparation, which includes already the phonetic word form, syllabification, word intonation, phrase intonation and sentence intonation. In self-monitoring, activations descending from these representations are transmitted in an internal loop. They don't yet contain the specific preparations of articulatory gestures, such as the exact steering of the articulatory muscles. I think that this view coincides nicely with the introspective observation that somebody speaking a foreign language and using it in thinking and internal speech generally does not self-perceive her special foreign accent, which she well realises clearly when she hears a recording of her own speech. The lack of the specially thick, more open, qualities of the /a/ and /o/ in Dutch are not self-perceived in internal speech by a speaker of Dutch with a strong German accent, who pronounces these vowels in a rather flat manner. She, however, recognises this accent immediately in the pronunciation of Dutch by other German speakers and in recordings of her own voice.

Concluding this section we can state that consciousness does not in general and not directly control brain states or brain processes. Being aware of the state one is in merely concerns those results of the brain workings which hit the sensorial fields of the brain, and it does not concern theses states themselves, rather the phenomenal expression of these, including categorisation on conceptual maps and space-time ordering in the form-motion areas of vision, hearing, and motoric activity. Sensorial fields in different modalities, which bind the results of specialised processing of the respective sensors and have them used in the circuits including conceptual activation, I have called episodic maps. On these maps we have conscious experience of our external surroundings in perception and we have imaginations and modifications of these in planning and thinking, and we have conscious experience of our body and its relationships to the surroundings. All that which we are conscious of becomes hereby a matter of control, modification, and planning. Consciousness provides us with orientation to which we can react in controlled and purposeful ways. Not that we always do that, but at least we are able to do this, and by this we know what we are and what we do, such that we can be at least partially responsible for what we are and what we do.

The next section shall deal with the question whether thought or judgement is part of all conscious states, and what role language plays in this. Are there conscious states without language?

5.2 Do all conscious states involve thought/judgement?

In chapter I have discussed Carruthers' (1996) claim that all conscious states include, sine qua non, judgement or thought in a language or language-like medium. The problem with this position was that he rightly wants to assume that higher animals have some form of consciousness, but it is not clear what language or language-like medium for thought representation might be available to a speechless animal. In the present essay, following Bartsch (1998) on concept formation and understanding, a tentative answer is given: Although a speechless animal has no language proper, it may have conceptual orderings in the form of some general concepts and individual concepts, which respectively characterise and identify situational impressions the animal gets as input. The corresponding neuronal connectivity, mainly connection strength and inhibition, serves to classify and identify certain situations and even serves to construct some B-intersections, i.e. intersecting and binding sets of experienced situations which form (quasi-)concepts into more complex (quasi-)concepts of basic situations (cf. chapter II). Situations are hereby characterised and identified, and they can be imagined as something desired or believed.

Another author who takes a strong position on the role of judgement in consciousness is Nelkin (1996). He proposed that "the senses are best defined by processes that end up with judgement-types and begin with organ stimulation" (1996: 30), and he claims that "judgement is no mere sequel to perception. Judgement is the sine qua non of perception"(1996: 35). There seems to be something right in this view, but it is formulated in a manner too extreme, if judgement is understood as being a judgement in a linguistic medium, a proposition. Nelkin seems to handle a weaker notion of judgement whereby it is typical for a judgement that it is a representation which aspectualises information. The problem with this is that the notion of aspectualisation needs to be explained such that aspectualisation can be possible even without using a language as a medium of representing the information. In fact this is the same problem as the one which arises for Carruthers.

Nelkin's specific contribution to this discussion is the differentiated structure he takes consciousness to have. Consciousness has a tri-partite structure consisting of CS (= sensation consciousness, including CN as phenomenal consciousness which consists of image-like states) as a basis, C1 (a first order propositional consciousness or awareness, including intentionality), and C2 (a second order consciousness about CN or C1, which he also calls 'apperception' (1996: 81 ff.), and which might, according to Nelkin, be located in Baars' global workspace, possibly situated in the reticular-thalamic system (1996: 227) . He argues for this tripartite differentiation by citing examples of dissociation of these levels of consciousness in pain perception, and in pathological cases such as blindsight and subliminal awareness. I shall discuss his examples in detail because they have to be dealt with, even if one does not endorse Nelkin's three-level theory of consciousness.

I first will discuss Nelkin's claim that even pain experience necessarily includes a judgement, and that pain is not classified as pain because of a natural class of pain phenomena, i.e. similarities in the feeling of pain, but because of a judgement that relates a feeling to the evaluation "harmful for the body"(1996: 75). There is no feeling of pain without such a judgement, Nelkin proposes. This "evaluative theory of pain" is a modification of the "attitudinal theory of pain", according to which pain is constituted by the attitudinal states of evading situations causing pain and having the appropriate emotional reactions. According to these positions an animal can experience pain, even though it does not experience pain phenomena. Nelkin modifies this position somewhat, by writing that phenomenal states are necessary for pains, but that they do not form a natural class; "no particular kind of phenomenal state is necessary" (1996: 82). There are no such natural classes of pain (1996: 83): "the phenomenal state of pain by itself does not constitute pain (unevaluated, or evaluated differently, there is no pain, even though the very same phenomenal state occurs)." (A pain is thus a complex state of the form (CS+C2), read as "What CS represents is harmful to the body." This position, unfortunately, is not compatible with the fact, that there are pains which are not harmful for the body, rather promote its health, as we all know. Such pains we experience in stretching muscles, after jogging for a certain time, where after some more time the pain disappears due to the production of endorphins. There are certain medical treatments which are painful, but we nevertheless evaluate them as beneficiary for our health. In all these cases we have pains, fitting into certain subcategories of pain experience, which are natural kinds of pain experience for each of us. On the other hand there are phenomenal experiences we have which we certainly judge as being harmful to the body, but which are no pains at all, such as the fine experience of a smoker when inhaling the smoke of tobacco. Since, according to Nelkin there is no natural class of phenomenal pain states, rather merely the evaluative judgement "harmful for the body", this C2-consciousness is defining for us what pain is, and thus we have here a case of dissociation between the two levels of consciousness, CS being variable, though indispensable in some phenomenal experience or other, and C2 being constant and thus defining what pain is. This position is not convincing, seen the counter examples just presented. Also the cases of patients under morphine, who "say they feel pain but it no longer hurts" (1996: 61) are not convincing because these people still feel some nagging, which indicates where the pain originates when it hurts. So they feel and know that the origin of pain is still active, though it now does not feel really painful, i.e. does not hurt. I would not say that this shows that there is a considerable cognitive input into our *feeling* of pain (1996: 68), rather there is this cognitive input in our judging that the cause of pain is still there and pain would make itself felt in full strength if no medication was used.

We will now examine other cases of dissociation between Nelkin's three levels of consciousness.

The blindsight cases are treated by Nelkin in a special way: He claims that blindsighted people do perceive in their blind field, though they do not experience phenomena. He takes this position because he holds that blindsighted people have proposition-like cognitive states with respect to the object in the blind field, arising from the senses, though they do not have a phenomenal state. In terms of Block's distinction between phenomenal consciousness and access consciousness (cf. section 3), he would say that they have some rudimentary access consciousness with respect to the objects in the blind field. Nelkin claims that blindsighted people have some CS without CN, namely sensational states without phenomenal consciousness, and they have C1, i.e. propositional consciousness and intentionality with respect to objects in the blind field. But they lack C2-awareness. According to him, these are examples of a dissociation between C1 and C2 and between CS and C2. The second order awareness about ones own state of awareness ('apperception' or self-awareness) is lacking (1996:125). C2 with respect to CS would require a CN-representation, i.e. phenomenal awareness. According to Nelkin C2-awareness, apperception consciousness, is a second order state which has CS/CN or C1 states as its content. By itself it has no phenomenal properties (1996: 82): "Apperception has no phenomenology to it (nor does C1 ...). In the sense that there is something it is like for an organism to experience phenomenal states (Nagel 1974), there is nothing it is like for an organism to apperceive. ... If one is apperceptively aware of S nothing phenomenally is added to the phenomenality of S itself."

Now, according to Nelkin, blindsighted people have no CN and no C2 with respect to objects in their blind field, though they have C1, since they assent to propositions. Nelkin does not take into account that these people do not really assert the propositions. Rather, when presented with a choice, for example, whether there is an X or there is an O in the blind field, they will guess with a high probability the right proposition. Therefore I rather would explain the blindsight reaction in another way: Blindsighted people are caused to guess the right proposition because of two factors, first the propositions presented before them as an alternative, and second the sensory input in their blind field. The sensory input in the blind field is partly processed at least so far that it fits better into the semantic processing of the one proposition than into the processing of the alternative proposition. The semantic processing of the proposition hits the visual maps and thereby can get an imaginary visual content by activating unimpaired visual areas outside the blind field. On the visual maps the unconscious sensory input to a large extent fits with the results of understanding one of the presented propositions. This fitting is not judged by the blindsighted person, because he is not aware of the input. Rather only some of the neural processes and states in understanding the sentence and processing the visual input of the blind field are fitting with each other. Hereby the person feels an urge to assent to the one proposition, and not to the other. But he does not know why he assents. Thus he does not judge, but merely feels an urge to prefer one presented sentence to the other, due to the partial fit he is not consciously aware of. He experiences himself as just making a guess. He is not representing sensory data by means of a proposition, which he could construct if he had perceived the data in the form in which they can get represented in propositions of our language. Forming propositions as judgements is expressing linguistically the structure of situations given in perception as the structure of the phenomena in our surroundings. The latter structure is due to concept formation, i.e. to the ordering of experiential data by similarity and contrast under perspectives and by contiguity relationships. The guess of the blindsighted person by itself has no visual intentionality with respect to the objects in the blind field, though he, of course, understands and interprets the presented and chosen sentence via the unimpaired visual capacities he applies with respect to his consciously available visual field. He then can use the sentence as a judgement referring to an object in visual space by turning his head such that the object gets into his visual field, or he can use the sentence intentionally with respect to the space of his tactile perception, and if the object makes noises he can refer to it also in his auditory space.

I would even claim that the blindsighted person cannot directly visually imagine the object in his blind field, though he can imagine the object in space by imagining turning his head and having the object in his perceiving visual field, or in the spaces of the other sensorial faculties. This strong claim is made, because imagining something in a space with respect to oneself uses the perceptual and motor faculties as directed according to the body's position in space. My conclusion is that the blindsighted person is not using the proposition as a judgement about the object in the blind field and thus also does not have a belief about what is in the blind field based on unconscious sensory input. He therefore has no C1 consciousness with respect to objects in the blind field, as long as the object is in the blind field. The guess about a right choice is not a judgement proper. A judgement can indirectly be made in the way spelled out just above, but that then does not refer to an object in the blind field, but refers via the seeing field or spaces of other modalities and motor faculties. Of course, we can say that the blindsighted person is informationally sensitive in his blind field, but that is totally unconscious. It is certainly wrong to claim, as Nelkin does, that blindsighted people have visual perception because they have CS sensory input, together with C1, namely together with a judgement about what they see (1996:120). They merely guess blindly in a situation of presented choice, but they do not have a perceptual judgement with respect to objects in their blind field. As long as the object remains in the blind field, they don't know by themselves whether their guess is true or false. This of course means that they cannot judge on the basis of received sensory input in their blind field.

Another blindsight experiment, according to Nelkin, shows CS without C2 (1996:178): A half circle in the blind field having a certain hue has to be completed by a half circle in the seeing field with the matching hue. In the latter field several half circles with different hues are presented after each other just next to the half in the blind field. The blindsighted person has to select the hue in his seeing field which matches the hue in the blind field and thus completes that circle. Here again he comes up with right guesses in a probability which is more than chance. Here consciously perceived sensory information is processed and the one that fits best to the unconsciously received sensory information, which is partially processed, is chosen without knowing why it is chosen. Instead of claiming that the blindsighted person sees the hue, we just should say that the person's system is informationally sensitive to the hue, by simply receiving sensory input without seeing it. Nelkin uses this example in arguing that only those CS-states which are CN, i.e. have phenomenal content, are also C2, i.e. are apperceived. But still he claims that the pure CS-states without C2 are image-like, while perception is proposition-like in that it is combined with judgement and thus aspectualised, such that only certain aspects of the image are phenomenally represented. This seems wrong: images and percepts of situations or of external pictures are all aspectualised; they are partial compared to the real situations or pictures. An image or perceived picture for us is a percept or a series of percepts, representing different aspects of a complete situation or external picture. An image or a percept is always a partial picture of something which it more or less represents. Likewise every situation for us is only partially represented, be it by a description, or in perception. A consciously entertained image or percept of a situation shows the situation categorised under one or more perspectives and identified by contiguity relationships with objects in space and time. These categorisations and identifications can be linguistically expressed.

Both, our perception of a situation and our linguistic description of it are aspectualised. Aspectualisation does not mean that a representation is simply a proposition or proposition-like. Rather it means that it is categorised under perspectives by means of our conceptual system, and this it has in common with the result of understanding a sentence. Of course, most of these categorisations can be linguistically represented. A picture consists of analogue information, as all image-like representations do, but this does not preclude that at the same time it is aspectualised under several perspectives by our conceptual system. What is wrong is that Nelkin takes it that CS, i.e. unconscious sensory information merely in the form of neuronal excitement by received sensory input, is a picture or is image-like because it is analogue, while CN is aspectualised. Of course, sensory information as it comes into our sense organs is analogue, but that does not mean that it is a picture or an image. Pictures and images, and even very short lived momentous colour sensations, rather are on a par with perceptions; they are consciously entertained representations and not just sensorily effected neuronal states, which is what CS states are. Images are produced by internal input, coming from the conceptual maps or memory to the same sensorial areas that also produce perceptions on the basis of sensory input and circuit activation involving conceptual maps. Pictures and images are representations such that we are aware of how they represent what they represent; this is totally different for CS states of which we are not conscious and which therefore are, for us, nothing and certainly not representations, contrary to pictures and images which we are aware of as representing what they represent. This, of course, does not exclude that from CS states information is played through in a one-way fashion to conceptual maps and motoric maps such that our body can adjust itself unconsciously a certain sensory input, like in automatic driving or in raising our voice unconsciously if there is noise around, even when we do not experience the noise consciously.

Another example of dissociation between different kinds (or rather aspects) of consciousness is subliminal reception of sensory data in driving a car or walking on the street, whereby one reasonably secure steers through the traffic, but is not consciously aware of all the input of road and traffic data. With respect to the non-conscious data this is, according to Nelkin, a case of lacking C2 and therefore lacking awareness of our having CS, though we in fact have CS, which is processed unconsciously. I would say our system here is informationally sensitive to a lot more data then those we have a conscious representation of. We might even have some representation in the visual, auditory , and tactile and motoric spaces without explicitly putting attention to these. We are sensitive to the data, but not consciously sensitive. We are not aware of getting these data. CS by itself, i.e. without CN, is no consciousness at all, not even a level of consciousness, though CS may function in directing motor activity via a one-way causal chain without conceptually enriched circuit activation, which would lead to an expression in consciousness.

In his dissociation arguments Nelkin also claims that there are C1 states without apperception, i.e. without C2. These are unconscious intentional states with a propositional content, for example unconscious beliefs or desires. I would rather say that these are not propositional states at all, and certainly not intentional states. Rather we here have dispositions of a neuronal network which do not consist of discrete propositions; the network has the potentiality to produce, under suitable stimulation, the propositions in linguistic or imaginal form on the episodic maps, on which we have representations in form of inner speech or imaginations of situations. I want to stress that all propositions need a medium of representation in which they are individualised or discrete as propositions, and this happens on the (linguistically specialised) episodic maps which are part of our sensory modalities connected in circuits of activation with conceptual maps providing for grammatical categorisation and conceptual maps providing for the semantic aspects of linguistic expressions. Thought therefore always happens in this medium. There are unconscious thought processes. But a thought itself as the result of such processes has to be represented on the episodic maps. And there we can deal with it in evaluation and symbolic manipulation. So we always think in one or other of the languages we have learned, or we think in sequences of images of situations, often aided by linguistic representations. In pointing this out, I oppose Nelkin's claim that thoughts do not require phenomenality (1996: 159), and that they can be apperceived, as C2 over C1, also without phenomenality. And therefore intentionality, which for him is bound to C1-states, i.e. to propositional states, can be entertained without CN, i.e. without phenomenality. This just is not true. Only thought processes can take place without phenomenality, and this being so, we are unconscious of thought processes. But the results of these processes, the thoughts are conscious only because they are couched in some medium of representation, either in sentences of inner speech, i.e. in speech images, or in representations on motoric episodic maps of articulation, or in auditory, or visual written representation on the maps of the respective sensorial modalities, in circuit interaction with conceptual maps. The results of thought processes can also be represented as images of situations, of sequences or constellations of situations. However, they are only existent thoughts in these forms, and intentionality with respect to thoughts, as we entertain it in propositional attitudes, is bound to such representations. It therefore requires phenomenality.

Nelkin claims that all perception includes judgement, what he also calls aspectualisation. That seems right at least in so far that perception includes conceptualisation. However, against Nelkin I argue that also picture-like representations can contain aspectualisation in selecting or outlining forms in drawings or in sequencing of pictures. Especially satirical cartoons are very much aspectualised. There are means of aspectualising in external pictures, such as composing sequences of pictures, heightening aspects of pictures by foregrounding and light effects and by representing details only to a certain extent. Conscious composition of concepts is only possible by using representations in the form of linguistic signs, or sequencing of pictures, drawings, or iconic symbols. Semantic composition understood as such, in whatever representational medium, is a property of consciousness by which semantic transparency of representation of our world is achieved, and by which the contents of our intentional attitudes towards the world are structured (cf. also chapter I).

Nelkin, like for examples Chalmers, also takes CS-states and C1-states to be states of awareness (1996:188), although the individual is not aware of them without C2, i.e. without apperception. I rather would say, though this seems a purely terminological issue, that these states are not states of awareness at all. Rather they are states of informational sensitivity. A state can be called a state of awareness, or a conscious state, only if we are aware of this state, i.e. are in a state of apperception with respect to this state. I also claim that every state of awareness, every conscious state, requires phenomenality, because only hereby we have available representations, be it thoughts, emotions, proprioception, perceptions, or images. Nelkin holds that proposition-like states, C1-states, and phenomenal states, CN-states, get apperceived and this directly causes a apperceptive representation, i.e. second order representation (1996: 207), which we know, but which we do not feel in any sense. Why then stop at second-order representations? We are aware of being aware that we are aware of such and such a fact. I think the notion of a separate second order representation is useless and does not make sense, especially because it would give rise to an infinite regress of representations of representations. If we apperceive a state we are just consciously in that state. We do not need to produce an extra representation of it and be conscious of that. We just know we are in that state because we experience or 'feel' a phenomenal representation as the result of thinking, perceiving, imagining, or of running through pre-motoric processes with results on motoric episodic maps.

Concluding, we can say that consciousness deals in representations; and nothing else than a representation is conscious, especially not the neural processes and states that lead to these representations as their phenomenal results. All representations on episodic maps contain a circuit interaction of sensorial and motoric maps with conceptual maps, in which the causal effects of previous input have had their impact on the neuronal connections such that on these maps evolving indicators of quasi-concepts, and finally concepts, are implicitly encoded in a converging learning process. Therefore all representations are categorised, i.e. conceptualised in so far as the sensory input will interact with the results of previously received similar input. Also contiguity relationships, especially spatial and temporal locations will be represented in a combination between sensorial, especially visual form and movement maps, and motoric maps. In this sense one can say that all conscious states, i.e. all representations contain some conceptual characterisation and some identification in space and time, and thus are couched in the structures of thought and judgement. The conceptual structures, including the generalisations over contiguity structures, evolve in the learning process of an individual and thus the conceptualisation will be rudimentary in the beginning, but it is continuously developing further with growing sets of experiences. The first experience of something which is not similar to something else previously experienced under a relevant perspective, must be felt in an undifferentiated manner and it must be localised in the visual field, though it cannot yet be categorised. Only the second input of the same (kind of) thing is rudimentarily categorised by being linked to the causal effects of the first input, and with more input of this kind, indicators of quasi-concepts and then concepts arise by stabilisation of identities and contrasts. Then new input of this kind will be categorised in an advanced manner, as it is in thoughts and judgements of experienced perceivers. In terms of neuro-physiological development, the connections have grown and have been strengthened in the learning processes in which the routes of circuit activation have been established that are typical for full-grown perception and understanding.

5.3 Are there different kinds of consciousness?

Block (1997 and elsewhere) makes a distinction between two kinds of consciousness, phenomenal consciousness (P-consciousness) and access consciousness (A-consciousness). Though, according to him, they mostly appear together in one conscious state they can be dissociated in some cases. And these dissociation cases are his argument for upholding his claim that there are two kinds of consciousness, though they mostly appear together. He takes P-conscious states to be devoid of cognitive, intentional, and functional properties. This means that they do not involve thought; they are not about anything; and they are not definable in terms of a computer program. Thus a person not possessing the concept of a donut could have a phenomenal experience of space being filled in a donut-like way. A P-conscious state has a content which is the 'feel' of what it is like to be in that state. This feel he calls the experiential properties of the state. According to Block, we can experience in a P-conscious state sensations and even types of sensations, without A-consciousness. A-conscious states, on the other hand, are cognitive, intentional, and functional. An A-conscious state is being used in reasoning and in rational control of action, and for humans it makes possible representation in speech. It is 'representational' in the sense the term is used in computer science, and only representational states can be used in reasoning. But note that states which are representational in this sense are not per se conscious states. A typical example for A-conscious states are, according to Block, propositional attitudes, such as thoughts, beliefs, desires, etc.

Against Block let me stress again that the propositional content of such attitude states must be phenomenally represented, either in inner speech in some or other natural language one is engaged in, or in sequences of imagined situations. Both require phenomenal material in which the representations are couched, which are provided by one or the other sensory modality. Thus a Helen Keller would have linguistic representations in terms of tactile phenomena, in reality and in imagination, and situational representations also in terms of tactile phenomena, and further in taste phenomena, olfactory phenomena, and motoric and other proprioceptional feelings. Thought only is conscious if couched in phenomena of one or other sensory and motoric modality. Thought processes are never conscious, only in-between and final results of these processes are representational in consciousness. Against Block I therefore hold that without the material provided by the so-called peripheral modules there is no consciousness. In this I agree with Flanagan (1997: 372), who argues against the position that phenomenal consciousness is merely an epiphenomenon without any function. Block also claims that phenomenal consciousness has no function, in distinction to A-consciousness, which is functional. Flanagan stresses that phenomenal consciousness is "out of touch with so much of what goes on in the brain, but is hooked up to the sensory modules and to declarative memory, that is, memory of facts and events." According to him, phenomenal consciousness is playing several causal roles for orientation in the external world and inferences from that; it is a kind of attention. I would even sharpen this conclusion by saying that all consciousness is phenomenal consciousness; even consciousness of thoughts is only possible by being conscious of representations couched in one or other phenomenal medium.

One of Block's examples of P-consciousness without A-consciousness is hearing a pneumatic drill without realising it because one is totally involved in an interesting conversation. When one suddenly becomes consciously aware of the noise, one is not only P-conscious of it but also A-conscious. Against this example we can object that one has not even been P-conscious of the noise, though one might have raised one's voice automatically. We have here a case of what Flanagan (1997) calls 'informational sensitivity' without 'experiential sensitivity'. The noise phenomenon is so much suppressed by the activity of listening attentively to speech, which takes place in the same sensory modality, that we are for a while not aware of the phenomenon of noise. Of course, the noise can become so hard that the speech data cannot be grasped anymore and thus the noise becomes a matter of attention. In Block's terms, contrary to

Block's own interpretation of the example, we rather might speak of A-consciousness, because the noise-information was processed such that the system adjusted its behaviour by speaking louder. The noise reception thus was functionally relevant, changing the speech behaviour in a rational way. But notice that this is an unconscious information processing, a case of pure information sensitivity of a system, not a case of consciousness. Therefore the best account of this example is to suppose that the person, for a while, was not conscious of the noise at all, though he was informationally sensitive to it.

It is problematic to claim, as Block does, that experiencing *kinds* of sensations can be purely P-conscious without being A-conscious. This is not convincing because to experience a kind of sensation involves already a categorisation of the sensation. A certain feeling, for example a certain pain may be unique, but if it is repeatedly experienced it is automatically categorised and by this recognised automatically as a certain kind of pain in distinction to other kinds of pain, or other kinds of feelings. If we see a colour, for example red, for the first time, it is not yet categorised for us, though it has an impact on the connection strengths between our sensory and conceptual maps, and with that on the occurring activation patterns on the conceptual maps. This impact is the starting point for forming the neural basis for knowing the concept Red. If we see something red repeatedly and also see other contrasting colours we automatically link a new red experience to the previous red experiences, because the input causes most activation at those neurones on the conceptual maps that had been activated by red input previously. This is categorisation, or conceptualisation, and it cannot be separated from something like pure undifferentiated sensation. The latter is only there at the very first experience and might not even be conscious for certain simple phenomena like colours. It might be that a conscious feel of the sensation already requires a feed back from the conceptual maps, which helps to delineate the quality of the feeling or sensation. Contrasts and recognisability might be needed for consciousness and focusing attention on a feeling or sensation.

If we have no concept of a donut, another of Block's examples, and see a donut for the first time, we have of course a representation which includes categorisation as a somewhat flat round thing, delineated by concentric circles, of the form of a ring. Thus this is no example of pure P-consciousness. Of course we have no complete concept of a donut which has to contain knowledge of the use and the material of the thing. If we even don't have the concept of a ring, we will categorise the object as something round. And if we see something round for the first time, we might not even experience it as round. Maybe the round experience comes about after having had some input of round forms in contrast to other forms. We don't know what our first experiences are like, and how much is conscious about them. An inborn bias for attending to some basic colour and form concepts may already trigger some experience, but it becomes an experience of a certain kind only when we can recognise the experience as similar with a previous one, and in contrast with other experiences. A first experience of a certain kind of thing minimally is categorised as being similar to and being different from other previous experiences. Thus the donut is not categorised as a donut but, may be, merely as something round. This already is a cognitive process. Only the very first experience of an individual cannot be categorised at all. Here we don't know what it is like by itself. It cannot be like something because there is nothing yet to which the organism can compare it. The question "What is it like?" is only meaningful if comparison is possible with something else. Then it can be like something else. A very first sensory input cannot be like anything, because there is not yet anything. We might have a very undifferentiated feeling that there is something, but it is not like anything and not unlike anything. A special recognisable feeling is only effected in us by experiencing identity and contrast to other feelings, and this is already conceptualisation, in Block's terms a matter of A-consciousness. My conclusion is that there is no P-consciousness without at least some A-consciousness, because every content experienced as such and such a content has to have a character and is identified by spatial relationships to other objects and parts of space.

Block admits that phenomenal content mostly is in some way or other representational, but it need not be. I would argue that this is so only for the very first experience, and there we cannot yet speak of a phenomenal content but rather there may be a totally undifferentiated feeling without a character. A phenomenal content would have to be experienced at least as distinct from other contents and should be recognisable as being of a certain kind by being similar, i.e. partly identical to other contents.

According to Block, A-consciousness without phenomenal consciousness can be found in the propositional attitudes of a super-blindsighted person, an unrealistic construction of a person who would be totally blindsighted, but still would react to what is put in front of him by taking the initiative to say what kind of thing there is, and he would perform actions which take into account what is in his visual field, although he does not see it. According to Block, such a person would have A-consciousness without P-consciousness. The well-known blindsighted person cannot employ this initiative; he merely can make a choice from presented possibilities, and therefore there is, in the interpretation by Block, not P-consciousness and also not A-consciousness with respect to objects in his blind field. Against this interpretation of the blindsight cases Flanagan (1997) argues that the blindsighted person has some passive informational access which is unconscious and which prompts the mostly right choice from a presented alternative of propositions purported to identify an object in his blind field. Conscious awareness of the environment, he points out (1997: 366), "facilitates semantic comprehension and adaptive motor action in creatures like us." The blindsighted person will not make a move to the water fountain in his blind field, even if he is thirsty; he has to be pressed to it. On the other hand, a system without conscious awareness but merely informational access to the external, which may be automatically connected to the appropriate motor activity, has no representation of this information in consciousness and therefore could not take a position towards it such that he could control whether or not to activate the appropriate motor activity. Flanagan argues that if access to information is conscious it is so because of its boundedness to phenomenal (experiential) sensitivity. As a case in support of his assumption Flanagan states the incapability of Luria's (1972) wounded soldier Zazetsky, who has a linguistic memory of events without being able to experience them as memory of his own life history. He misses the phenomenal sensitivity of the events, which he has remembered as stories; thus the memorised and remembered stories are merely distant stories, but not events from his own history. The stories are not felt to be his stories. Experiential sensitivity involves awareness of the state one is in, namely consciousness.

The whole conception of two different kinds of consciousness, which we also found in Nelkin (1996), who distinguishes cases of (CS+C2) and (C1+C2) consciousness, is inspired by the model of a modular mind, where there are sensory and motoric peripheral modules, and central cognitive models in which information processing is executed. In this model, phenomenal consciousness is purely due to the sensory modules and can occur separately from access consciousness which comes about by the cognitive processing of the central cognitive modules. In special cases, especially in pathological cases the route from the sensory modules to the cognitive modules is not activated or it is damaged. Then we have phenomenal consciousness without access consciousness. This traditional picture of the one-way modular relationship from sensory input to cognition, and consequently to motoric output, is typical for traditional Cognitive Science and Artificial Intelligence. It is not a good model for how the brain works. And it is not a good model for the mind as a fairly abstract model for organising the conscious results of the intelligent workings of the brain in a structure that relates to intentional objects and knowledge of the world. By assuming that information is just processed in a series from one module to the next, the traditional model neglects the circular activity between sensory fields and conceptual maps, and its role in binding, selecting features, and directing attention.

Also in Philosophy we find positions which assume that there are pure sensory contents without conceptualisation. One of them is Davies (1997). He rightly claims (1997: 309) that "a perceptual experience has representational properties; it represents the world as being [in] a certain way." An object is represented in space in a certain direction and distance from the observer, having a certain shape. On the other hand, he also claims (1997: 309) that "the perceptual content of an experience is a kind of *non-conceptual* content. What this means is that the subject can have an experience with a certain perceptual content without possessing the concepts that would be used in specifying the content of that experience." This must be so for human infants and "certain other creatures, who arguably are not deployers of concepts at all. Enjoying experiences with conceptual content does not require the possession of concepts." He refers back to Peacocke's (1992: 61-62) 'scenario content' as the non-conceptual content: "I suggest that one basic form of representational content should be individuated by specifying which ways of filling out the space around the perceiver are consistent with the representational content's being correct. The idea is that the content involves a spatial type..."

It certainly is right that there are spatial contiguity relationships processed by the visual form and motion fields, motoric fields, and also auditory fields. In binding these activations together by strengthening circuit connections, an internal model of space is built up by means of directions and distances, acquired by experiencing space in motor activity connected with visual and auditory experiences. This already requires binding between different sensory and motoric modalities. However, spatial types, i.e. spatial characterisations, do not evolve without any categorisation. Spatial relationships are categorisation of instances of relative locations of objects to each other and to the self. A relationship is a concept, a two or more-placed concept, and so is a spatial relationship. In order to have a spatial representation one needs spatial concepts, i.e. relationships of distance, movement and direction. Also objects have at least to be recognised as objects, even if they are not yet identified as special objects. Thus, minimally, the concept of an object as something bounded in space has to be formed in contrast to other objects and to empty places in space. Therefore I conclude that there is no perception without some conceptualisation, and of course human infants develop the knowledge of spatial relationships and a notion of an object in contrast with other objects and empty places. The notion of an object, or minimally a filled space, will be enriched and differentiated by more experiences.

The same is true for the cognitive development of other animals. There is no reason to assume that they do not have concepts. Quasi-concepts are equivalence classes of series of situational impressions ordered on the basis of partial identity (similarity) and contrast under partly inborn and partly developed perspectives, which select or put under attention certain aspects of the data which function in identity and contrast relationships. Concepts are the stabilised quasi-concepts, and linguistically guided concepts are those that are formed with respect to satisfaction situations for linguistic expressions. Animals, of course, can form those concepts necessary for their daily orientation in their natural and social surroundings and their eating habits. They have no linguistically guided concepts. However, perception such that something is recognised as being of a certain type (character) and being identified by contiguity relationships in space and time, i.e. perception as not merely reception of sensory data, but as perception of the world as it appears for the creature given its perspectives, inborn and formed by needs and desires for pleasure and avoidance of pain, makes up the consciousness not only of humans, but also of many kinds of animals. There also might be living creatures which purely process sensory input information without perception and automatically map it on motor activity, as we also do for a great deal of received input which we do not perceive, but to which we are nevertheless informationally sensitive. However, for the higher animals consciousness is their life, as they experience it. It consists of conceptualised representations coming about by interaction circuits between conceptual maps and sensory fields.

Also Davies points out that the perception of the external, its natural and social surrounding, is the basis for consciousness of a creature and he rightly stresses that for this the creature need not be a thinker, like Evans (1982) assumes, and also Carruthers (1996), we might add. But against Davies, I argue that animals, even though they are not thinkers in their full glory, form concepts under those perspectives that are important for them. Concepts are just orderings on growing sets of data. And even a fairly rudimentary conceptual system containing spatial relationships as types of contiguity relationships and containing the distinction between empty and filled places, resulting in a rudimentary concept of an object, is a conceptual system, though quite small as a beginning. Our conceptual system is built up of contiguity structures and similarity structures on sets of data, leading to the formation of individual and general concepts. What kind of concepts and how much of a conceptual system can be learned by different kinds of living creatures, given certain biases in the learning brains, is a matter of the needs and the intellectual capabilities the creatures are endowed with. But whatever a conceptual system looks like, it provides the contents of the world represented in the respective individual's consciousness. It provides the contents, which are never without some form given by conceptual structuring. Form is in the content, as it structures the content and distinguishes it from other contents or makes it similar to others. Form is perceivable in content by interaction of sensory input with a conceptual system which develops in learning from data taken up under certain inborn or developed perspectives. The phenomenal character of, for example round and square, is categorisation or conceptualisation under the perspective of spatial form, the phenomenal character of, for example red and green, is a conceptualisation under the perspective of colour. Without such conceptualisations built up in learning neuronal networks, there would be no differentiation and identification and thus no phenomenal character.

Employing concepts is a way of thinking; it is a way of thinking without having a language. The representations of the thought are not in a linguistic medium but in sequences of imagined situations. In a simple case, wanting a piece of food of the kind as has been eaten the day before, which by the way was a certain kind of sausage, means to imagine this piece of food or a series of modifications of it, by preference bigger pieces. The (quasi-)concept here is an equivalence class of such series of imaginations and perceptions of this kind of food. The equivalence here would be determined by the smell, the taste, the texture of the substance, and possibly the size and form. Aspectualisation and selection comes about by keeping certain features constant and making others variable by modification in series of imagined examples. Thinking means instantiating (quasi-)concepts in imagined instances; the propositional content of the attitude is represented by an imagined situation of eating the stuff and the attitude is represented by remembering the pleasure of eating, together with the feeling of longing for a repetition. We will credit our dog with thinking this far. We may correctly assign to him this propositional attitude when he scratches at the box from where we yesterday took the sausage.

Burge (1997) supports Block's distinction between two kinds of consciousness, though he makes some additional points, namely (1997: 427) "that access consciousness - indeed any sort of consciousness- in an individual presupposes the existence of phenomenal consciousness in that individual." With this, of course, we can agree. But he claims that "although phenomenal qualities are individuated in terms of what it is like to feel or be conscious of them, one may have phenomenal states or events with phenomenal qualities that one is unconscious of." So one can be in a state of pain without feeling the pain, for example because one is diverted from feeling it by other events and feelings. For being conscious they must be felt. Or one may feel a piece of metal as cold after having touched a much warmer one, though it is as warm as it was when one felt it before. Burge claims that here one is in a state with the phenomenal quality of a certain warmth, though one does not feel it.

Burge's description of what goes on in these cases is not convincing. With respect to the pain example we can object that one is not in a state with the phenomenal quality of pain if one does not feel it, even if the signals from the affected body part are as strong. The brain states which these signals effect, just do not get the phenomenal expression they would have if attention would or could be given towards the source of pain. Probably there is missing some strengthening power of attention which would be necessary against diverting perceptions and feelings, in order to produce the pain phenomena with sufficient strength. Either the pain is felt at least a little, or there is no pain phenomenon. With respect to feeling the temperature, we can object that probably already the signal from the tactile nerve endings in the skin is not there or is very weak, because just before something hot has been felt with the same hand. If it had been felt with the other hand, then another explanation might be to the point. The effects in the brain could relatively be very weak because of the strength of the previous feeling of the much hotter metal, the effects of which are still recycled in the tactile field of the brain. Therefore no feeling of warmth is produced with respect to the piece of metal that would still be felt as warm in another context. So we have no reason to assume the same measuring result and the same brain state for the same tactile input in two different contexts. Herewith we also have no reason to assume that there are the same phenomenal qualities expressed.

Burge argues against Block that access to central cognitive powers does not by itself cause access consciousness. Also this kind of consciousness is not fully explained in functional terms. There could be a perfectly rational zombie, without any form of consciousness. I think Burge is right in making this point, but he cannot explain what is added to rational processing in order to make a conscious thought. He misses an explanation because he still adheres to Block's view that there can be rational consciousness with respect to intentional states "without any phenomenological types as essential elements." He goes on (1997: 428): "Conscious mathematical beliefs or thoughts are commonly examples." And "Rational-access consciousness need not have any essentially phenomenal aspects." The question here is what he means be 'essentially'. To me, who has worked in Mathematics for years and has conversed with many mathematicians and logians about mathematical and modeltheoretic questions, it certainly seems essential that a mathematical thought, or whatever thought, is represented in some phenomenal medium, by its symbolic or linguistic form, or by an image or a drawing depicting it. Mathematics has never been done without writing or drawing, be it in external material or in imagination. Both are in visual or auditory or articulatory form, and in extreme cases in tactile or motoric form. A special phenomenal medium is not really necessary, but some such medium is necessary. This means that some sort of phenomenal qualities are essential even to mathematical thoughts, as they are for all representations, though these need not be phenomenal qualities of a specific sensory modality; there is a choice between several of them. Only what is represented, i.e. is perceived or imagined, can be manipulated in controlled ways. Even if we infer in retrospection that we must have had some thought in the process of thinking in order to arrive at a certain resulting thought which is represented symbolically or by some image, we still cannot say that this intermediary thought was presented in an access-conscious state. It has not been conscious in whatever sense, if we had not consciously represented it in some phenomenal medium. Of course it later on is conscious as the result of a process of backwards reasoning, by which we seem to find out that we must have been thinking this thought. I would rather say we have not been thinking this thought as a separate thought, as far as thinking is manipulating representations. Rather it has been an unconscious state of brain activity, or even more likely, it is merely implicitly encoded as a disposition in processing, which did not even occur as a brain state of a separate thought. Anyhow, for us it has not been a thought, though in a later rational reconstruction of the process of thinking in intentional and propositional terms it appears as a thought. All this reconstruction, of course, takes place in a phenomenal medium, and then there appears represented in that medium a separate thought.

Burge also claims (1997: 134) that we often "verbalise our thoughts unreflectively and immediately in a public manner, without any prior phenomenally conscious inner articulation." These thoughts he counts as examples of rational-access consciousness, without phenomenal qualities. Here we can object again that these thoughts as represented are only there the moment they are articulated to a sufficient degree in imagination. The articulation is the medium of their representation. Before such articulation, they are simply not there as thoughts, though there are dispositions implicitly encoded in the strength of the neuronal connections in our brain, and there is an unconscious brain activity leading to these thoughts. But as separate mental entities the thoughts are not there if they are not entertained in some phenomenal medium, in at least some rudimentary form of articulation in some sensorial or (pre-)motoric medium.

My conclusion again is that there is no access consciousness pure. All conscious states require phenomenal content. Otherwise they could not represent anything. Representation requires a medium which is phenomenal content if the representation is to be conscious and hereby controllable, and in this way is an representation in the full sense. So-called representations in computer science are only in an indirect way representations; they can be interpreted by us as representations of certain situations and constellations, be it of surroundings or of linguistic representations in speech or otherwise. For us, when we refer to them, these representations are also couched in a phenomenal medium, which can be in the form of computer outprints or representations of computational states on the screen. The computer states themselves are likewise in principle perceivable by following electric currents and the position of open or closed connections in logical gates. So everything that we can understand as a representation is couched in phenomenal content.

Concepts themselves are not representations for us. Rather a linguist can make a partial representation of a concept according to examples of use of that concept, observed or imagined. Concepts for us can only be conscious in series of examples. What occurs on our conceptual maps in our brain are not representations in some mentalese. Neither are the connection strengths between neurones in which our knowledge of concepts is implicitly stored as dispositions of the system to react in certain ways. Rather they are the cumulative effects of the examples provided in concept learning, relative to the learning mechanisms, possible biases, and also randomness in begin situations. Further there are the linguistic cumulative effects of the relationships of the words expressing concepts in sentences held true. Especially theories and other coherent sets of general sentences held true define conceptual relationships between words, and hereby secondarily by their connections to the conceptual maps they constrain the concepts by constraining their applicability. In this way the general sentences constrain what are possible episodes and what can be expected on the episodic maps. However, to entertain a theoretical text (a theory) or a story is a complex episode of a special kind, and hereby it is a conscious mental activity, suited to control.

5.4 Is there a 1-1 correspondence between brain states and conscious states?

We have already pointed out repeatedly that there cannot be an one-one mapping between brain states and conscious states, because most brain states are not accompanied by consciousness. Even brain states which have intelligent functions are not all accompanied by consciousness. Baars (1997) has made a point of this fact by contrasting unconscious and conscious states in the same area of cognitive and perceptual activity (see discussion in section 1). Thus we have delineated a set of brain states with some specific characterisation which then have an one-one mapping on conscious states. A not fully satisfactory attempt to do that was made by O'Brien and Opie (1999), who claimed that stable states in neural networks were the candidates, since they can be interpreted as representations (see discussion section 1). The problem here was their notions of representation and of stability, and the deficiency in their argumentation why brain processes should not contribute to consciousness. A general problem for such a position is how a brain state should be identified as being the same as a previous one, or even as being the same as one in another person. Should the assembly of activated neurones define the identity? In a strict sense, then no two brain states, even of one person, are the same. Every occurrence of a single object happens in a different external setting or a different mind internal context, and thus it will give rise to a different state of activation on the brain. Every instance of a general concept will occur in different circumstances and thus show differences from the other instances. Due to contextuality the effected brain state in the perception of the instance could not be the same. One might filter out the differentiating sub-assemblies of neurones and thus concentrate on an intersection of the brain states, whereby the intersection is taken as corresponding to the class concept, or in the first case, the individual concept. This intersection, which is stable across the different presentations of instances, one might take as a representation of the concept evoked by an instance. But note furthermore that concepts are not conscious at all and therefore cannot be representations, only instances of them in perception or imagination are conscious and are representations because they are couched in a representational medium. These instances might effect relatively stable brain states for a short time during the process of perception. But how short a time needs to be taken into account for defining the identity of such a brain state?

Another problem is that the same cognitive functions might be performed by brain states which are different in the sense defined above. An individual in its life time changes his conceptual network by learning new concepts, new relationships between concepts and especially new contrasts between concepts. A concept that had been indicated by a certain assembly of activated neurones at occasions of suitable stimulation might later in life be represented by a different assembly. Such changes, of course, happen during the learning processes. The only condition for being a brain state that indicates a concept on suitable occasions is 1. that the brain state, i.e. the activated assembly stands in the right relationships to other concept indicators, such that previous semantic contrasts, implications, partial identities and compatibilities, which are not corrected, are preserved in a growing network. And 2. the relationships to the sensory inputs and suitable motoric activations are preserved, such that reference to the same external objects and actions remains possible. This consideration shows that the identity of a brain state must not be defined in a material way, but in a structural, or functional way. In this structural definition there is one item that involves a material identity, namely the relationship to sensory inputs. The sensors stay the same and they measure the input in the same way. It is reasonable to also assume that their connections to the sensory fields in the brain stay the same, at least after they have grown out fully aided by external stimulation in the first year of infancy. The partial identity of structural relationships between concept indicators and there relationships to partially identical sensorial and motoric material, i.e. to certain groups of neurones on sensorial and motoric fields, guarantees the partial constancy of the phenomenal character of objects. We have arrived at a combination of a structural, functional notion of identity, combined with a material notion of identity for the sensory organism, consisting of the sensorial fields and their sensory organs, and a material identity for the motoric ability, consisting of the fields for motoric gestures and the executing organs.

Of course, it is likely that different individuals built up different neural activation patterns for the same concept, since they might start from a different random distribution of connection weights between neurones in the conceptual maps and between those and the sensory fields, possibly aided by some in-built biases. Furthermore the learning history is different from individual to individual due to the natural and social surrounding he lives in, and to factors of simple accidence of the situations he is confronted with. The different conceptual systems of individuals have to be structurally isomorphic to a great degree, will people act in a co-ordinated way, perceive the world in a similar way, and be able to understand each other. But this structural identity does not include that there are assemblies of neurones identically localised in the corresponding brain areas which have an identical function. An activation pattern representing a certain concept in situations of appropriate stimulation may be located at different places for different people, though it has to be found in functionally corresponding fields of the brain architecture.

Kripke (1997) has criticised the identity thesis which claims that there is a one-one correspondence between certain kinds of brain states and states of conscious awareness. He uses the general argument, also used by Putnam, that a certain function, and especially a mental function, can be realised by quite different means, here by quite different brain states. Thus one kind of pain sensation could correspond to different brain states, which in some individuals have not the corresponding mental state while in others they have. Apparently he assumes that at least in the same individual there will be a correspondence. Chalmers answers this criticism to the one-one hypothesis by pointing out that all depends on how sameness of brain states is defined. Inclusion between neuronal assemblies would not be right because it could be that when an A-state has a corresponding feeling of pain, an (A+B)-state might not have the corresponding feeling of pain, since the B-state could cancel the effect of the A-state. The problem with defining sameness for brain states in a material and local way does not occur for a functional definition of sameness of brain states.

Mundale (1998: 138) also points out that "functionalism was proposed by philosophers, such as Putnam, as an alternative to proposals for identifying mental states with brain states. Putnam and others argued that, since instances of the same mental state (e.g. hunger) can be realised by different brain states, and instances of the same neurological state can give rise to multiple mental states, mental states could not be identical with brain states." Mundale points out that it is not even clear yet what counts as the same physiological and neuroscientific state. These are defined differently by different scientific methods, such as measuring blood circulation, warmth, electric activity, and electro-magnetic fields in brain areas. He stresses that "brain parts, let alone states, are not just given - they are identified through a variety of criteria that are still evolving." It is therefore too early to just embrace the functional view. "But, given the use of functional, including psychological criteria in demarcating brain parts, there is reason to be pessimistic about the strength and importance of the conclusions that multiple realisability can be made to support." We shall see in what follows that defining brain states by criteria of psychological functions and then claiming an one-one correspondence between structures or systems of brain states and structures or systems of mental states, especially conscious states, is a circular endeavour. The claim of the one-one correspondence becomes a tautology if there is no independent definition of what brain states and structures or systems of these are.

We have seen above that Block only assumes access-consciousness to be defined in functional terms, namely as states that play a role in intelligent information processing, while phenomenal consciousness is not defined functionally, but rather referred to as some state that has a 'feel' to it. One problem with the position of Block was that he cannot distinguish processes and states of intelligent processing of information which are conscious, from those which are unconscious. The brain states, described in a certain coarseness are at the same time mental states, but some are hidden mental states, and others we are aware of. What makes for this difference is not explained functionally. Here Nelkin (1996) comes in who assumes that there is an extra level of consciousness, namely C2-awareness which is a second order awareness of the states we are in. The C2-awareness has as contents the phenomenal properties of CN-states, i.e. of those sensational states (CS-states) that have phenomenal properties, and/or of C1-states, i.e. of proposition-like states. This means that we also have phenomenal states we are not aware of, and we have propositional states we are not aware of, a position also criticised with respect to Burge (1997) in this essay. Nelkin takes it that all three aspects of a full state of consciousness are defined in functional terms.

We shall now present Chalmers' approach to consciousness and discuss his arguments for a moderate functional view about the relationship between brain states and the occurrence of conscious states, namely the position of Chalmers (1996). He does not want to defend a position in which consciousness is reduced to functional organisation, but rather a non-reductive functionalism which is compatible with certain kinds of property dualism. Hereby conscious experience is determined by functional organisation without being reducible to it. I shall show that his arguments for (this moderate form of) functionalism are not convincing. He uses them against defenders of partly materially based accounts of consciousness, such as the view that consciousness is due to a certain material biological endowment. He rather wants to defend the view that experience in a physical system arises from its functional organisation. Thus, if two systems have exactly the same functional organisation, they would have the same experiences, the same consciousness. This principle is referred to as the principle of *organisational invariance*. He defines functionalism (1996: 14) in the following way: "On this view a mental state is defined wholly by its causal role; that is in terms of the kinds of stimulation that tend to produce it, the kind of behaviour it tends to produce, and the way it interacts with other mental states."

Chalmers introduces a notion of awareness, functional awareness, as a psychological notion different from our everyday use of awareness (1996: 28): "Awareness can be broadly analysed as a state wherein we have access to some information, and can use that information in the control of behaviour." This, of course, is Flanagan's 'informational sensitivity', or Block's notion of access-consciousness, which by itself does not imply consciousness. A robot or a zombie has this kind of awareness. Now, according to Chalmers, this is the psychological notion of awareness or consciousness, which is functionally based on how a system behaves, especially on what it reports linguistically about its surrounding but also about introspective accessibility of information. From this he distinguishes the phenomenal consciousness whereby mental states are characterised by the way they feel, rather than by what they do (Chalmers 1996: 11). According to him, the psychological notion of consciousness is easy to understand; we see how it works. But this is not so for phenomenal consciousness. (Functional) awareness need not to be accompanied by phenomenal consciousness. Chalmers, of course, wants to attack the difficult problem of phenomenal consciousness. He does not attempt a functional reduction, but rather wants to show the essential role in the functional organisation also for this kind of consciousness. Reductionism, like cognitive models generally, only can explain the functional aspects of consciousness, not the phenomenal aspects. Also supervenience is not a solution because it leaves the main problem unexplained, namely why is it that B-properties supervene on A-properties, i.e. why is it that no two possible situations are identical with respect to their A-properties, while differing in their B-properties (i.e. A' = A'' => B' = B"). According to Chalmers, also neurobiological positions, such as the one of Crick and Koch or the one of Edelman, though interesting contributions, do not bring us closer to solving the hard problem of how all this processing does give rise to consciousness. Chalmers (1995: 203) states the hard problem like this: "Even when we have explained the performance of all the cognitive and behavioural functions in the vicinity of experience - perceptual discrimination, categorisation, internal access, verbal report - there may still remain a further unanswered question "Why is the performance of these functions accompanied by experience?... this further question is the key-question in the problem of consciousness. Why doesn't all this information processing go on 'in the dark' free of any inner feel?"

Chalmers tries to do some steps into the right direction. He first makes the way free for a property dualism, namely that brain states have causal and functional properties and also may have phenomenal properties. Then he gives some hints to what the phenomenal properties add to the functional capacity. Firstly he argues that dualism is logically possible, witness the thought experiment about zombies, which are exact physical replicas of ourselves and can do exactly the same things we do, but have no consciousness. A naturalistic dualism is supported by Jackson's (1982) argument and quite similar by Nagel's (1997 and earlier) argument against reductionism, concerning a Martian who thoroughly investigates the human mind in relation with observable causal and functional relationships and still does not know from his investigation what it is like to see a rain bow. In Jackson's argument we are called to imagine a neuro-scientist brought up in a black and white room, who knows everything about colour processing in the brain. However, later when she is confronted with red objects she learns some new fact, namely what it is like to see something red.

Chalmers (1996: 176) investigates how far functional awareness goes. The Zombie can make phenomenal judgements, when programmed and trained in appropriate ways for making statements based on sensory input. He can be trained to make the statements of different order correctly: 1. order: That is red; 2. order: I have a red sensation; 3. order: Sensations are mysterious; or even Consciousness is difficult to explain. These judgements can be explained functionally by neural net activation without any consciousness coming into play. But here we should correct Chalmers' terminology: The zombie does not make judgements, because he does not know what he is talking about; he has no notion of correctness or incorrectness of these sentence utterances with respects to the things they refer to. The zombie does not even refer to anything; he simply reacts due to causation. He just produces utterances causally connected, by the programming, to certain other states he is in. The same is true with respect to a zombie having a belief. A current belief state is an information state with the addition that one consciously holds the information true about that what one refers to. We distinguish false and true thoughts, and only thoughts with a truth claim we take to be beliefs. For a judgement or belief it does not matter whether it is caused by (conscious) experience; what matters is that it is taken to refer to something and is evaluated as true or false on the basis of reference and predication. For checking that, one needs own or other peoples experience. Of course, also a zombie could be trained to check for truth; he could just repeat the causal process. However, he would not know that he is checking anything. Chalmers takes a similar route in his arguments (1996: 148) by a thought experiment about a zombie twin: "In response to the arguments of my zombie twin, we note that justification of my beliefs about experience involves more than the mechanisms by which the beliefs are formed: it crucially involves experiences themselves. Because my zombie twin lacks experiences, he is in a very different epistemic situation from me, and his judgements lack the corresponding justification....I have evidence." Though indistinguishable from a zombie under a third person view, we know that we have these or that experiences, that we have this or that evidence, and that we are conscious. Chalmers concludes (1996; 208) this argument with "there is good reason to believe that the epistemology and semantics of experience cannot be essentially causal, and should instead be understood in other terms." To show just this was the topic of chapter 1 of the present piece of research, where normative notions were shown to play a role in a denotational semantics; this in contrast with a merely causal semantics.

Chalmers (1996: 213) gives a frame for a non-reductive and still moderately functional theory of consciousness by trying to formulate psycho-physical bridge laws which should govern the relationship between consciousness and physical systems. This is supposed to provide some substance for the supervenience relationship; it then is not just a logical relationship, rather a psycho-physical one, based on some plausibility constraints. There has to be a plausible coherence between conscious experience and cognitive structure (1996: 218). The coherence between consciousness and (functional) awareness is such that, for example, the visual experience of a red book is accompanied by a "functional perception of the book". He states the principle of structural coherence as follows (1996: 222): "Where there is consciousness there is awareness, and where there is (the right kind of) awareness there is consciousness." We note that the formulation "the right kind" hides what still is a great deficiency of his theory. He also gives a weaker formulation of the coherence principle (1996: 123): If there is consciousness, as it is for us, then "structural features of consciousness correspond directly to structural features that are represented in awareness." Though this seems reasonable, the following criticism has to be made: this either is a vacuous claim because the structural features of awareness are constructed by the cognitive scientist on the basis of structures in consciousness, i.e. structures of conscious experience such as those reconstructed from a rationalistic point of view in knowledge analysis, intentional steps in behaviour, and intentional steps in reasoning give rise to structures in proposed psychological cognitive models. Or it is not an empty claim; then there should be an independent way of constructing or recognising psychological and neurological structures of the right coarse grainedness as parts of awareness. But that is impossible because the right degree of coarse-grainedness of the neurological structure and its activities can only be projected from out the psychologically defined functions and functional steps in these.

A further problem for Chalmers coherence principle is, of course, blindsight. He immunises his position against blindsight cases (1996: 227) by doubting that in these cases there is (functional) awareness without experience. He claims that either there still might be some experience, or there is also no real (functional) awareness in the sense of availability of information for control. The latter is supported by the fact that the blindsighted patients merely guess (mostly rightly) from presented alternatives. "The lack of experience directly corresponds to the lack of control", he remarks. With respect to Block's examples of zombies with super blindsight, he remarks that they are not examples against his position, because they are just thought experiments and not actual cases; and in normal blindsight cases, access consciousness is also impaired, as Block realises himself. He discards as arguments against his coherence principle also cases cited by Block for phenomenal consciousness without access consciousness, by claiming that access consciousness was very well dispositionally present, but could not be accessed or activated. Note that Block's examples were the following: a loud drill was not heard because the person was engaged in an interesting conversation. Here Block claims that there was phenomenal consciousness without access-consciousness, because the person did not report hearing the drill, but only did that later when he suddenly realised the noise. This interpretation does not seem correct. Rather the person merely had sensory (functional) awareness without phenomenal consciousness. If phenomenal properties are not realised then they are not there. Or they are very weakly expressed, as in subliminal cases, and we then can not attend to them. They are suppressed by other phenomenal properties, and thus do not fully come into existence. Here the phenomenal properties are only produced with so little strength relative to the interfering situation, that they are not really there. Their expression is frustrated, so to speak. In this case we have functional awareness, an informational sensitivity to the sensory data, and they even are processed such that there is access-consciousness because the behaviour rationally changes to automatically speaking louder. In Chalmers' terms we can say that there is full (functional) awareness present, though there is no phenomenal consciousness of the drill noise. Block's other case was that nine letters were flashed fast and all seen, but only the first three were reported. This shows, according to Block, that there was access consciousness only with respect to the first three letters, though there was phenomenal consciousness with respect to all letters. Chalmers here claims that there was access consciousness with respect to all letters, but reporting the first three hindered reporting about the other six. But why should that be so? If it was a matter of reporting, then it should also be the same in cases of nine slowly performed flashes. What Chalmers wants to uphold, inspired by his coherence principle, is that there was a full functional awareness present since there was phenomenal awareness of all nine letters.

We shall now examine Chalmers' arguments in the chapter 'Absent Qualia, Fading Qualia, Dancing Qualia' (1996: 247-275). The principle of *organisational invariance* is stronger than the coherence principle, since it says that experience is invariant across systems with the same functional organisation expressed in a representation of the same degree of detail. It thus states that systems with the same organisation, in the same coarse-grainedness, have qualitatively identical conscious experiences. This, of course, would be false if consciousness would depend on the biological make up of an organism, such that neuro-physiological properties determine the quality of consciousness, or even whether there is consciousness at all. So opponents of a functional view on consciousness, even if it is not reductionist, point out that functionally organised system can be realised in the strangest ways, even by a system of water pipes (cf. Searle 1992), which certainly would not produce qualia and other phenomenal properties. So the system by itself could not be responsible for producing consciousness. Chalmers calls this the 'absent qualia' argument. Against this argument he wants to show that it leads to the empirically absurd possibility of fading qualia, and since such a thing is very implausible, the argument is very implausible, and thus cannot hit the functional position. The other argument is directed against functionalism with respect to the specific qualitative properties of consciousness. It states that two systems with exactly the same organisational structure might both produce consciousness, but that could be quite different kinds of consciousness, for example such that, although they would have structurally the same identities and differences in the colour system, one would have blue experiences where the other would have red experiences. From this 'inverted qualia' argument, Chalmers wants to derive that then the systems could also have changing qualia, dancing qualia, which would be absurd. Having this very implausible consequence, the argument cannot be used against his principle of organisational invariance.

According to the 'absent qualia argument' a system like ours, with our functional organisation and non-organisational biological properties of neurones has qualia, while a system with the same functional organisation made from silicon chips has no qualia. What would it be for a system in between, where part of the neurones were replaced by silicon chips? Replacing one or more neurones could result in suddenly disappearing or in gradually fading qualia. The first seems quite implausible if only one or a few neurones are replaced out of a great number. But there also would not be an abrupt limit from where the qualia suddenly are no more produced. Thus the possibility of fading qualia seems more plausible. Since the person with the partly replaced neurones and the correspondingly fading qualia still is functionally the same, he would say about a shirt with a bright red colour that it is bright red, before and after the replacement, even though his phenomenal experience would be faded after the replacement to a weak pink. He will still claim that he experiences all kinds of shades of red though in fact his experience is of a more or less homogenous greyish pink. In his judgements about his perceptions this person would be out of touch with his experiences. In the extreme case he might not consciously perceive any difference and thus anything at all and still come up with all kinds of judgements which we judge as true. This is, in fact, the evolving zombie depicted by Searle (1992: 66f.), who can be imagined to arise by replacing neurones by silicon implants. Although his experiences fade away and even are reduced to nothing the system still entertains all the beliefs it would have if it still was substantially and not just organisationally the old one. The strange thing about this thought experiment is, that Searle accepts its consequences and takes it to reject a functionalist basis of consciousness, while Chalmers thinks it consequences to be so absurd that it cannot reject functionalism, but rather supports it: his way out is just to assume that nothing changes in the experiences by replacing neurones by silicon chips with the same organisational functions.

This argument against the 'absent qualia' position, however, really is not worth anything because, depending on the position one starts out from, one finds it appealing or finds it implausible. The role which the argument can play depends on the evaluation of the result as plausible or implausible. However, the evaluation of the result of the thought experiment is circular, since it is not independent of but heavily rests on the anti-functionalist (Searle) or functionalist position (Chalmers) from which one departs and at the same time wants to defend. That a system can behave rational with respect to the outside world and sensory input it gets, and still be completely wrong in his judgement of his fading experiences, or even talk about experiences which are not there at all seems very implausible to Chalmers. Thus he concludes that unless we are prepared to accept this massive dissociation between consciousness and cognition, the system consisting of silicon chips must be taken to be conscious like the original system was. But note that the statements of the evolving silicon based system are not statements about its experiences at all. Rather they are about its internal states couched in the vocabulary of experience. The system works such that it monitors its internal states and reports on them in the current language. So these are different statements because the mean something different. Furthermore, since the system does not experience its internal states, what the system says about these are for the system itself not even statements about its own states, but just it is utterances produced according to the causal connectedness of the system. We can interpret them as being about his inner states. This, of course, is very plausible, a position Searle does embrace. Note also, that this system does not even realise that its experiences have faded, since it cannot remember the old colour experiences. It cannot remember these, because for that it would need from out its conceptual map for colour concepts and its memory of certain colour episodes the activation of the neurones on his sensory field of colour vision, which have been replaced by chips. He would still have his implicit colour knowledge, made use of in transforming sensory input to utterances of colour reports. But he would not have a current memory of colours, and thus no colour representations.

The 'inverted qualia argument' does not fare any better. It is directed against the position that two structurally identical but materially different systems might have experiences that differ in phenomenal qualities. The thought experiment is a modification of the previous one: If part of the neurones or neuronal circuits would be replaced by silicon chips with the same functions then, according to the position under critique, the experience of a red colour might get replaced by the experience of a blue colour. What was red before would now be blue. And if the silicon circuit would be installed parallel with the neuronal circuit which it can replace we could switch within seconds from one circuit to the other, and thus the individual would experience dancing qualia, for example the same object would be red or blue depending on how the switch is flipped between the two alternative circuits. My experience would change "before my eyes" from red to blue and back. The linguistic report of the colour would of course stay the same. In both cases I would say "the apple is red", although in the second case I now experience blue. However, this account does not seem correct, because it is based on an idea of memory that is unrealistic. The individual is thought of as remembering that the apple was just a moment ago red while now it appears blue. This indeed would be very confusing. The individual thus would have a memory like a library, in which the previous red-experience is stored as being red and can now be compared with the new blue-experience. This is not how a memory works. Rather the memory is like a neural network in which concepts, and also memories of single episodes, are kept implicit in the strength of the connections between the conceptual map and the sensory maps.

Episodic memory (cf. Chapter 3.4) must contain indicators of episodes and their parts which are connected to maps with indicators of space-time relationships and to maps with other conceptual indicators. In current memory, the episode indicator and the relational and conceptual indicators activate parts of the sensorial and motoric fields, and by establishing circuits of activation a remembered episode arises on the episodic maps. Current memory, i.e. acts of remembering situations and objects, is only possible by activation spreading from the conceptual maps to the sensory maps involved in the previous experiences. On the sensory maps it gets the phenomenal quality. Thus, a change within the conceptual maps by replacement with silicon chips, upholding the same structural properties, i.e. the connections and their strength, would not be felt at all and would have no phenomenal consequence. A change within the sensory fields by replacement with silicon chips could possibly change the quality from red to blue, but the individual would not realise this because it cannot remember that the apple was red. The apple memory, a current activation pattern in the episodic memory getting into interaction with the conceptual maps along the connections strengthened in memorising the previous episode, would activate the neurones on the visual map or, after a switch, the replacing silicon chips there, and thus the same-colour apple would be remembered as what it appears now. Remembering is only possible via the sensory fields involved.

Otherwise there would not be any phenomenal quality to the memory. Remembering the colour of certain object could either happen in terms of remembering the describing sentence, and that indeed stays the same in this case; the expressed judgement remains "The apple is red." Or the remembered episode is realised on the sensorial map with silicon replacement. Also then the sentence reporting the new perception would stay the same, because the individual has kept the relationship between the outside stimulus and the reporting sentence identical. The silicon chips have the same relationship to the input stimulus and to the output sentence utterance as the neurones have which they replace. Chalmers finds the consequence of dancing qualia very implausible because it involves an "extreme dissociation between consciousness and cognition"; when one's experiences change significantly one should be able to notice the change. But notice that this puts the whole situation wrong: the experiences do not change at all because the individual cannot remember that it had other experiences. What he remembers and what he reports is totally coherent; it is the same about the same subject. Only we outside could suggest, according to our inverted qualia hypothesis, that he must have a different experience; but he himself will never notice this. Thus he, subjectively has no different experience. Though the qualia might have been inverted, he will never experience it due to the fact that memory is bound to the sensory field, and thus goes with the material changes made there. Of course, this whole scenario is similar to the well-known argument of logical empiricism which shows that it is not meaningful to say that we all experience the same qualia, or that we possibly do not. This just can never be verified or falsified. So also our individual can neither know, nor not know whether he experiences the same qualia as before, because his current memory is always qualitatively determined by the sensory fields as they are now.

Note, by the way, that if also the motoric articulatory fields and the auditory fields would be replaced by silicon structures our imagined person could not even remember sentences or handle sentences as representations, because he could not experience sentences. Of course, he still could automatically utter them in the situations we would find appropriate. If all sensory and motoric modalities would be replaced by silicon structures we would end up with a complete zombie.

Like the argument against absent qualia, also the argument against inverted qualia is not convincing. Qualia-inversion is not implausible at all, rather it is a meaningless claim because it cannot ever been proven or disproven. But this, of course, is no reason to claim that the qualia are the same between two functionally identically but materially different systems. This claim is simply meaningless.

What we end up with is Searle's versus Chalmers' position, namely that the phenomenal quality of our experience is not (sufficiently) caused by the functionally defined cognitive system, but needs the biological quality of the material it is made of for producing qualia. This does not preclude that also other material might be able to produce qualia. Chalmers, however, wanted to show that this position is empirically implausible, and in this he did not succeed. That his arguments are not conclusive he realises himself. Therefore he formulates a weaker version of functionalism, non-reductive functionalism, namely that conscious experience is determined by functional organisation without necessarily being reducible to functional organisation. But note that this weaker position is not what is stated in his coherence and invariance principle. Those, especially the invariance principle, are formulated much stronger, and are the basis of his further speculations about the possibility of panpsychism (1996: 293-310), namely that the whole cosmos could be structured like an information processing system; and there could be in this way an overall consciousness. But if that would not be so, Chalmers makes us asks the following question (1996: 308-9): If not all information processing systems are conscious, what then are the properties that make some of these experiencing/conscious, while others are not? This is something those have to answer who do not accept his coherence and invariance principle. In the present essay this question is answered by pointing out that the sensory and motoric fields play the essential role in creating experience and making, hereby, consciousness possible and also real.

I want to conclude this discussion by pointing out that our experience certainly depends very much on the functional or structural properties of our cognitive system. Thus we experience objects having properties and standing in relationships to other objects and situations. These all are structural properties of conceptual categorisation and contiguity-ordering. But they can only be experienced in phenomenal material, like any form needs a material out of which it is made. The sensorial fields are the sources for the phenomenal material. They somehow produce it. Our perception of objects and situations, and our current memory or imagination of these, is due to an interaction between the conceptual and memory maps with these sensorial maps. The visual maps of form and movement provide, together with proprioceptic motoric and acoustic maps, for the contiguity ordering of space, and also time; the visual map for colour, the ones for taste, sound, smell, and touch provide further qualia in which structural relationships and categorisations can be couched. What has been called episodic maps in the more abstract model of cognition developed here is in neurological reality very likely a complex of interplay between sensorial and motoric maps with conceptual maps. This then also settles the question whether consciousness is brought about by a vehicle, such as a relatively stable activation pattern, or whether it is due to processes. The vehicle and the process theory of consciousness are both by themselves not convincing. Rather consciousness is found in the interaction described above, with an essential role for the sensorial fields in the brain which provide the phenomenal material in which the structural and categorial properties can be couched. Vehicles, i.e. activation pattern on the sensorial and motoric maps and activation patterns of the conceptual maps both play a role connected to each other by the processes of activation running between them. A functionally defined system of conceptual maps is rooted in the medium of sensorial and motoric fields. The phenomenal properties and organisation of objects or situations, i.e. episodes, come about only as phenomenal expressions of an interaction between structurally and functionally determined conceptual aspects, and materially determined sensorial and motoric aspects.

Conclusion

The main thesis of this book has been that consciousness arises by an interaction between primary sensory fields and conceptual maps in resonance circuits. These synchronised circuits of activation constitute and episodic map. This thesis was reached as a result firstly via a discussion about blindsight phenomena, causal and denotational semantics, and the functions of consciousness in judgement and other acts involving evaluation (chapter 1), such as planning, designing, and employing free will. Secondly the result was reached via a discussion of what understanding basic sentences amounts to, and what the shortcomings of simple connectionist models are with respect to modelling understanding (chapter 2), and it was worked out in some detail for perception and understanding of situations and basic sentences (chapter 3).

The position reached about the nature of consciousness takes into account that all consciousness is episodic. A primary sensorial, motoric, or emotional field, when involved in such a circuit activation is called an episodic map. We can then say that consciousness arises on such episodic maps. The episodes are episodes of situations or episodes of linguistic inscriptions, spoken or written or signed otherwise. The constituent structure of situations and of linguistic inscriptions was modelled as a temporally induced architecture of smaller and larger activation circuits, which when hitting a sensorial field, give rise to representations, which are conscious. The notion "representation" is restricted to conscious phenomena. At the same time, this result is strongly supported by research and modelling in neurological cognitive theories, to which reference is made throughout the book (especially in chapters 3 and 5). Ways of thinking and their relationship to consciousness are described within the model developed. Thinking, as far as it is conscious, is seen as manipulating representations under the point of view of evaluation. Thoughts are the results of such processes; they are representations, and herewith conscious phenomena (chapter 4). Thus, there are no thoughts without consciousness. Thoughts are either representations on linguistically specialised episodic maps, or on situational episodic maps, or on both together. The essential feature of all thought is that it is couched in aspectualised representations, whether in outer or inner speech sentences of some language or in aspectualised pictures or images.

Memory (cf. chapter 3 and 5) is either conceptual, i.e. generalising, or episodic. The so-called declarative memory is episodic, whereby situations or linguistic inscriptions are remembered. When we remember something, an indicator of an episode, being a certain neuronal assembly located in a temporal order with other indicators within the memory map, is activated from out other such indicators, or from out indicators on conceptual or emotional maps, which in turn might be activated by some linguistic or situational input that reminds us on something we experienced in the past. The activation on the memory map in turn activates those neuronal assemblies in the sensorial, motoric or emotional fields which where involved in the original episode, and it activates those indicators on conceptual maps to which it is strongly connected, which are indicators of spatial and temporal relationships and of other relationships, and of one-place concepts, which were involved in the original episode. These enter into activation circuits with the neuronal assemblies in the respective sensorial fields, addressed directly from out the memory. In this way we have reconstructed the remembered episode, more or less accurately. Remembering is a way of imagining; but it is not just initiated from out the conceptual maps with maybe some emotional maps involved. Rather it also includes the space-time indicative pattern of a specific episode which was fixed by the original temporary connection of the original episodic activation on a sensory field, including its spatial and temporal order, to some neurones in the memory field.

The thesis about the episodic character of consciousness has then been confronted with several controversial issues in consciousness research and served there to take a position within these ongoing discussions (chapter 5).

The first issue pertains to the question whether consciousness is an internal monitoring device of brain states, or rather a monitoring of the external. The answer tends to the later: the external as we perceive it and reconstruct it in theories is monitored. Likewise designs, plans, and other imaginations, as well as linguistic representations and thoughts are monitored and evaluated. They are all of the same kind as our representations of the external, namely representations on the level of consciousness. There is no use in trying to look at brain states, which we cannot do, except as something seen as external in neuro-physiological research. We cannot monitor our own brain states, though we can monitor our state of mind, i.e. what is in conscious. We then don't look into our heads, rather we are just experiencing our surroundings and our theoretical reconstructions of our natural, social, and cultural world, and possible worlds of these kinds. We evaluate and monitor these, and not states of our brain.

The second controversy is whether all conscious states involve thought or judgement. The answer depends on what is meant by thought and judgement. In a narrow sense, thought and judgement are in terms of propositions, which are understood sentences, i.e. linguistic representations together with the circuits containing the conceptual activations that go with these representations. As judgements these representations are evaluated for truth or falsehood. According to this narrow sense, we must conclude that there are conscious states that are not thoughts or judgements. These are situational perceptions, images, designs on the one hand side, and emotions, motoric gestures, bodily feelings or proprioceptions, on the other hand. If we take thought and judgement in a broad sense as including all these kinds of representations, the answer would be that all conscious states involve thought or judgement. Since all these representations include some categorising or conceptualising activity, they are hereby aspectualised, and this they have in common with propositions. However, instead of overextending the notions of thought and judgement too far, we rather take as an position regarding this issue, that all conscious states involve representations, and herewith are aspectualised.

The third issue is whether there are different kinds of consciousness. There has been claimed that there are phenomenal consciousness and so-called access-consciousness. This position is typical for the modular picture of the brain as a central computer for cognitive processing, which results in access-consciousness, and perceptive and motoric modules for the input and the output. What is merely perceived is in this vision not yet cognitively processed and thus is part of phenomenal consciousness. The position developed in this book is a negation of this modular computer metaphor. It rather is argued that all perception is automatically categorised to some degree and thus contains the features of access-consciousness. All perceptions, imaginations, feelings have a representational character in that they are aspectualised by lower or higher categorisation. On the other hand there never is pure access-consciousness, because every conscious state needs a phenomenal medium in which the respective representation is couched; the phenomenal qualities can only be provided by sensorial, pre-motoric and emotional fields, which become episodic maps by their interaction with categorising maps. An important point is that the notion of a hidden representation, used by many cognitive psychologists, is seen as senseless. On conceptual maps we have no representations of concepts or propositions, rather merely indicators of concepts, i.e. indicators for growing and stabilising sets of examples, that make up a concept. We are never conscious of these indicators or of concepts; we are merely conscious of their examples, which have phenomenal qualities and are presented and represented on episodic maps. Thus there is only one kind of consciousness, and this is episodic consciousness.

The last issue treated is whether there is a 1-1 correspondence between brain states and conscious states. The final answer has been that there is a certain kind of brains states, including processes, which corresponds with conscious states. These brain states are resonance circuits that involve both, sensorial fields and conceptual maps. These circuits have a constituent structure of smaller and larger circuits. Consciousness hereby is not merely a question of a certain system or structure of a cognitive system, but requires also certain substantial properties, as we find them in sensorial, motoric, emotional and proprioceptic fields, which are able to provide for the phenomenal qualities.

Concluding I want to address the question of whether this model of consciousness which is conform to the neurological background, has any bearing to philosophy? Certainly, no reduction of consciousness or specific conscious phenomena to neuro-physiological processes or states has been arrived at, or even has been intended. Philosophy has its place it always has had; description and analysis of conscious phenomena is primary to any description of brain phenomena. Also these are, of course phenomena, as we see them according to our scientific questions asked and methods used. And furthermore, the brain phenomena that correspond to conscious states and phenomena represented in consciousness are only understood as such corresponding brain states and processes by seeing them in their relationship to the conscious phenomena to which they are taken to correspond, on the basis of contrast analysis in brain scanning research. Thus, phenomenal analysis is primary to what we can learn from brain research. On the other hand, we should be aware of not thinking of phenomenal analysis as looking into our heads, or introspection. We cannot really look into ourselves; we can't see anything there. So-called introspection or analysis of our internal states is not more than rational reconstruction and theorising about our overt behaviour, our surroundings, and our conscious modifications and manipulations of these under the point of view of evaluations, feelings, desires and preferences. All these are represented and thus open to us on the episodic maps. To turn ones view onto ones inner life is not more than looking at the external and its modifications which we can imagine and design following unconsciously our dispositions, and controling and reconstructing these according to evaluations with respect to conscious beliefs, feelings, desires, and preferences.

References

Allen, C. 1998. Review of Peter Carruthers, Language, Thought, and Consciousness. *Philosophical Psychology 11:* 91-95.

Armstrong, D. 1997. What is consciousness? *BFG* : 721-728.

Anderson, R.E. Imagery and spatial Representation. In: Bechtel and Graham (eds.) 1998: 204-211.

Baars, B.J. 1987. Biological implications of a global-workspace theory of consciousness: Evidence, theory, and some phylogenetic speculations. In: G. Greenberg and E. Tobach (eds.) 1987. *Cognition, Language, Consciousness: Integrative Issues*. Hillsdale: N.J. Erlbaum: 209-236.

Baars, B.J. 1997. Contrastive phenomenology: A thoroughly empirical approach to consciousness. *BFG*: 187-202.

Bartsch, R. 1987. *Norms of Language*. London and New York: Longman.

Bartsch, R. 1995. *Situations, Tense, and Aspects*. *Dynamic Discourse Ontology and the Semantic Flexibility of the Temporal System of German and English.* Berlin-New York: Mouton de Gruyter.

Bartsch, R. 1996. The Relationship between Connectionist Models and a Dynamic Data-Oriented Theory of Concept Formation. In: *Synthese* 108: 421-454.

Bartsch, R. 1998. *Dynamic Conceptual Semantics. A Logico-Philosophical Investigation into Concept Formation and Understanding.* Studies in Logic, Language and Information (= SiLLI-series) Stanford: CSLI and FoLLI Publications. Distributed by Cambridge University Press.

Bechtel, W. and A. Abrahamsen. 1991. *Connectionism and the Mind. An Introduction to Parallel Processing in Networks.* Oxford: Basil Blackwell.

Bechtel, W. and G. Graham (eds.) 1998. *A Companion to Cognitive Science.* Oxford: Blackwell Publishers.

Bod, R. 1995. *Enriching Linguistics with Statistics: Performance Models of Natural Language.* ILLC Dissertation Series 1995-4. Institute for language, Logic, and Computation. University of Amsterdam.

Boden, M. 1994. What is Creativity? In: Boden, M.(ed.), *dDmensions of Creativity.* Cambridge: MIT Press, Bradford Books.

Block, N. A confusion about the function of consciousness. *BFG* : 375-416

Block, N., O. Flanagan, G. Güzeldere (eds.) 1997 (referred to as *BFG*), *The Nature of Consciousness. Philosophical Debates.* Cambridge, Mass.: MIT-Press.

Brown, C.M and P. Hagoort (eds.) 1999. *The Neurocognition of Language.* Oxford: University Press.

Burge T. 1997. Two kinds of consciousness. *BFG*: 427-434

Carruthers, P. 1996. *Language, Thought, and Consciousness.* Cambridge: Cambridge PUniversity Press.

Chalmers, D. 1995. Facing up to the problem of consciousness. *Journal of consciousnss Studie*, 2, 200-219.

Chalmers, D. 1996. *The Conscious Mind: In Search of a Fundamental Theory.* Oxford University Press.

Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origin, and Use.* New York-Westport-London: Praeger.

Chomsky, N. 1995. Language and Nature. In: *Mind* Vol. 104. no. 413. Jan. 1995: 1-62.

Churchland, P. M. 1986. Some Reductive Strategies in Cognitive Neurobiology. In: *Mind* Vol. 95, no.379, July 1986: 279-309.

Crick, F.H.C. and C. Koch. 1990. Towards a Neurobiological Theory of consciousness. In: *Seminars in Neurosciences* 2: 263-275.

Crick, F. and Ch. Koch. 1997. Towards a Neurobiological theory of consciousness. *BFG*: 277-292.

Davies, M. 1997. Externalism and Experience. *BFG*: 309-327.

Dennett, D. 1989. *The Intentional Stances.* Cambridge Mass.: MIT Press.

Dennett, D. 1991. *Consciousness Explained.* Boston: Little, Brown, and Company.

Dennett, D. 1995. *Darwin's Dangerous Idea. Evolution and the Meaning of Life.* Londond: Allan lane. Penguin Press.

Dreyfus, H.L. and S.E. Dreyfus. 1986. *Mind over Maschine: The Power of Human Intuition and Expertise in the Area of the Computer.* New York: Free Press.

Edelman, G.M. 1989. *The Remembered Present: A Biological Theory of Consciousness.* New York: Basic Books.

Edelman, G.M. 1992. *Bright air, Brilliant Fire*. New York: Basis Books.

Elman, J.L. 1990. Finding Structure in Time, *Cognitive Science* 14, 179-211.

Elman, J.L. 1993. Learning and Development in Neural networks: The importance of starting small. *Cognition* 48: 71-99.

Elman, J.L. 1995. Language as a Dynamical System. In: Port R. and T. Van Gelder (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge Mass.: MIT Press: 1996-225.

Elman, J.L. 1998. Connectionism, Artificial Life, and Dynamical Systems. In: Bechtel and Graham (eds.) 1998: 488-505.

Evans, G. 1982. *The Varieties of Reference.* Oxford: University Press.

Fadiga, L. and Gallese, V. 1997. Action representation and language in the brain. *Theoretical Linguistics 23*: 267-280.

Flanagan, O. 1991. The Mystery of Consciousness. In: *Consciousness Reconsidered.* Cambridge, Mass.: M.I.T. Press.

Flanagan, O. 1997, Conscious inessentialism and the epiphenomenalist suspicion. *BFG*: 357-373.

Fodor, J. 1976. *The Language of Thought*. Hassocks: The Harvester Press.

Fodor, J. 1987. *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*. Cambridge, Mass.: MIT Press.

Fodor, J. 1997. Connectionism and the problem of systematicity (continued): Why Smolensky's solution still is not right. *Cognition 62*: 109-119.

Fodor, J. 1998. *Concepts. Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.

Fodor, J. and McGlaughlin, B. 1990. Connectionism and the problem of systematicity: why Smolensky's solution does not work. *Cognition 35*: 183-204.

Garfield J.L. 1997. Mentalese not spoken here: computation, cognition and causation. *Philosophical Psychology 10*: 413-435.

Garson, J.W. 1994, Cognition without classical architecture, *Synthese* 100: 291-305.

Grossberg, S. 1996. The attentive brain. In: *American Scientist* 83: 438-449.

Grossberg, S. 1999. The link between brain learning, attention, and consciousness. In: *Consciousness and Cognition* 8: 1-44.

Grossberg, S., Mingolla, E., and Ross, W.D. 1994. A neural theory of attentive visual search: Interaction of boundary, surface, spatial, and object representation. In: *Psychological Review* 101: 470-489.

Gulick van R. 1997. Understanding the phenomenal mind: Are we all just Armadillos? PartII: The absent qualia argument. *BFG*: 435-441.

Güzeldere, G. 1997. Is consciousness the perception of one's own mind? *BFG*: 789-806.

Hagoort, P., Brown, C.M, and Oosterhout, L. 1999. The neurocognition of syntactic processing. In: Brown and Hagoort (eds.): 273-316

Hebb, D. O. 1980. *Essay on Mind*. Hillsdale, N.J.: Llawrence Erlbaum Ass. Publ.

Horgan, T. and Tienson, J. 1992. Structured Representations in Connectionist Systems? In: St.Davis (ed.) 1992, Connectionism: Theory and Practice. New York-Oxford: Oxford University Press: 195-228.

Horgan, T. and Tienson, J. 1997. Précies of Connectionism and the Philosophy of Psychology. *Philosophical Psychology 10:* 337-356.

Jackson, F. 1982. Epiphenomenal qualia. *Philosophical Quarterly* 32:127-136.

Kandel, E.R., Schwartz, J.H., and Jesell, T.M. 1995. *Essentials of Neural Science and Behavior.* London: Prentice Hall International.

Kosslyn, S.M. 1994. *Image and Brain. the Resolution of the Imagery Debate.* Cambridge Mass.: MIT Press.

Kreiman, G., Koch, Chr., Fried, I. 2000. Imagery neurons in the human brain. *Nature,* vol. 408, 16 November 2000: 357-361.

Kripke, S. 1982. *Wittgenstein: On Rules and Private Language*. Oxford: Basil Blackwell.

Kripke, S. 1997. The Identity thesis. *BFG*: 445-450.

Laberge, D. 1990. Thalamic and cortical mechanisms of attention suggested by recent Positron Emission Topographic experiments. *Journal of Cognitive Neuroscience*, Vol. 2, no. 4: 358-372.

Leeuwen, C.van, Steyvers, M., and Nooter, M. 1997. Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. In: *Journal of Mathematical Psychology* 41: 319-344.

Levelt, W.J.M. 1999. The cognitive architecture of language. In Brown and Hagoort (eds.) 1999: 83-208.

Levelt, W.J.M., A. Roelofs, and A.S. Meyer. 1999. A theory of Lexical Access in Speech Production. In: The Behavioral and Brain Sciences (BBS) 22: 1-75.

Luria, A.R. 1972. *The Man with the Shattered World: The History of a Brain Wound.* Cambridge: Harvard University Press.

Lloyd, D. 1996. Consciousness, connectionism, and cognitive neuro-science: A meeting of the minds. *Philosphical Psychology* 9: 61-79.

Matthews, R.J. 1991. Psychological Reality of Grammars. In: A. Kasher (ed.), *The Chomskyan Turn.* Oxford: Basil Blackwell: 182-199.

Mc Ginn, C. 1991. Can We Solve the Mind-Body Problem? In: *The Problem of Consciousness*. Oxford University Press.

McGinn, C. 1991. *The Problem of Consciousness*. Oxford: Oxford University Press.

McGinn, C. 1993. Consciousness. In:*Problems of Philosophy. The Limits of Inquiry*. Oxford: Basil Blackwell: 27-45.

Lycan, W,G. 1997. Consciousness as internal monotoring.*BFG*: 755-772.

Mundale, J. 1998. Brain mapping. In: Bechtel and Graham (eds.) 1998: 129-139.

Nagel, Th. 1997. What is it like to be bat? *BFG*: 519-527.

Nelkin, N. 1996. *Consciousness and the Origins of Thought* . Cambridge: Cambridge University Press.

O'Brien, G. and J. Opie 1999. A connectionist theory of phenomenal experience. *BBS* 22: 127-196.

Pavio, A. 1986. *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

Peacocke, Ch. 1992. *A Study of Concepts*. Cambridge, Mass.: MIT Press.

Piaget, J. 1953. *Logic and Psychology*. Reprinted 1965. Manchester: Manchester University Press.

Putnam. H. 1981. Brains in the vat. In: *Reason, Truth, and History*. Cambridge University Press: 1-21.

Quine, W.V.O. 1974. The Nature of Natural Knowledge. In: Guttenplan, S. (ed.), *Language and Mind*. Oxford: Clarendon Press.

Roediger, H.L. and L.M. Goff. 1998. Memory. In : Bechtel and Graham (eds.) 1998: 250-264.

Rollins, M. 1989. *Mental Imagery.* New Haven and London: Yale University Press.

Sacks, O. 1985. *The Man who Mistook his Wife for a Hat.* London: Pan Books, Picador Edition.

Sacks, O. 1985. *The Man Who Mistook his Wife for a Hat.* Picador edition, Pan Books. London.

Saffran, E.M. and A. Sholl (1999). Clues to the functional and neural architecture of word meaning. In: Brown and Hagoort (eds.): 241-172

Scha, R. 1990. Taaltheorie en Taaltacnologie; Competence and Performance. In Q.A.M. de Kort en G.L.J. Leerdam (eds.), *Computertoepassingen in de Neerlandistiek.* LVVN-jaarboek. Almere: landelijke Vereniging voor Neerlandici.

Schnelle, H. 1996. Reflections on Chomsky's "Language and Thought". In: *Theoretical Linguistics*. Vol. 22, No. 1/2: 105-124.

Searle, J. 1983. *Intentionality*. Cambrideg: Cambridge University Press.

Searle, J. 1992. *The Rediscovery of Mind*. Cambridge, Mass.: MIT Press.

Searle, J. 1997. *The Mystery of Consciousness*. London: Granta Books.

Smolensky, P. 1988. On the Proper Treatment of Connectionism. In: *Behavioral and Brain Sciences* 11: 1-23.

Smolensky, O. 1990, Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence 46: 159-260.*

Van der Berg, M., Bod, R. and Scha, R. 1994. A Corpus-based Approach to Semantic Interpretation. In: *Proceedings of the 9 the Amsterdam Colloquium.* IILC. Institute for Language, Logic, and Computation. University of Amsterdam.

Vision, G. 1998. Blindsight and Philosophy. In: *Philosophical Psychology* 11: 137-160.

Weiskrantz, L. 1986. *Blindsight: a Case Study and Implications*. Oxford: Oxford University Press, Clarendon Press.

Wittgenstein, L. 1960. *Philosophische Untersuchungen.* (Engl.: *Philosophical Investigations*). Frankfurt a. Main: Suhrkamp Verlag.

Zeki, S. 1992. The Visual Image in Mind and Brain. In: *Scientific American*, September 1992, Special Issue "Mind and Brain": 43-50.

Ziff, P. 1960. *Semantic Analysis*. Ithaca, N.Y.: Cornell University Press

Index